Payment by Results in public service reform: silver bullet, dangerous weapon, neither, both?

Fraser Battye, ICF GHK and Mike Daly, DWP
July 2012

Abstract:

Policies based on Payment by Results (PbR) aim to strengthen the connection between inputs (payments for a service) and desired outcomes. PbR is a key element in the public service reform agenda, as articulated through the Open Public Services White Paper, and is currently being tested in a wide range of policy / service areas – including reoffending, health services, social care and labour market programmes.

This joint ICF GHK / DWP paper begins by examining the thinking underpinning policies based on PbR, the mechanisms by which it could be expected to lead to improved outcomes and the evidence for the effectiveness of those mechanisms. It then moves on to describe and assess the different models being tested, examining issues raised in the light of DWP’s experience, before concluding with ten key considerations for both policy makers and policy analysts.

Payment by Results (PbR) is a form of outcomes based commissioning; it is a central element of current reforms to public services

Traditionally, public services have been funded according to inputs or outputs - regardless of how effective they were for users. Outcomes based commissioning offers an alternative. It asks commissioners to be clear about the needs of the population and the long term outcomes required, rather than the detailed set of services to be delivered. It encourages commissioners to focus on ends, not means, and is seen as a means of promoting improvements in public services – especially during a period of cuts in funding. In essence, it requires a focus on results rather than process, moving further down the intervention ‘logic chain’, as illustrated in Figure 1.1.

Figure 1.1 Outcomes-based commissioning encourages a focus further down the logic chain

PbR aligns funding arrangements with this intention, paying for interventions / services, at least in part, on the basis of the outcomes achieved. PbR is therefore a potentially powerful tool. It can change the incentives of service providers; it can alter the balance of opportunities and risks between commissioners and providers; it can, in conjunction with other approaches, radically alter the ‘provider landscape’.

This theoretical power has been persuasive in policy, where PbR is promoted as a central means of improving public services. It is mentioned fifteen times in the ‘Open Public Services’ White Paper (July 2011), which states:

“Open commissioning and payment by results are critical to open public services. This is not just about opening up services to competition; it is also about empowering all potential providers, from whichever sector, with the right to propose new ways to deliver services, and linking payment to results so that providers are free to innovate and eliminate waste.”
“...it is not enough to pay someone to provide a service with the only recourse being that if they fail they will not be re-awarded the contract. In these cases it makes sense to build in an element of payment by results to provide a constant and tough financial incentive for providers to deliver good services throughout the term of the contract.” [our emphasis]

PbR can therefore be seen as part of a set of developments in commissioning. This is illustrated in Figure 1.2, which shows three differing approaches to service delivery. These range from approach ‘A’ whereby needs are assessed and provided for by a single (presumed state) institution; through approach ‘B’, whereby needs are assessed, commissioned for and provided by multiple service providers (presumed from different sectors); to approach ‘C’, which links payment for providers to the outcomes they achieve for users.

Figure 1.2 PbR can be seen as a development of outcomes-based commissioning

Approaches B and C also hint at possible mechanisms for achieving better results, in terms of competition ‘for the market’ (to become a provider) and possibly, under PbR, ‘in the market’ (to reach service users). Advocates for these mechanisms argue that:

- Diversity in the delivery of public services results in better outcomes than public sector monopoly; that a mix of public, private and voluntary providers stimulates greater innovation, choice and responsiveness in service delivery. The thinking here is partly that providers are given space to devise the service response; to think differently and experiment with different means of achieving outcomes. Knowledge about effective delivery (what works) is thereby assumed to be diffuse – held among a diverse range of providers in a market, rather than by a single ‘all knowing’ official / institution;

- Competitive pressure among providers can be used positively, since – all other things being equal – more effective providers will thrive. Again, this might be competition for, or within, a market; and

- Providers have stronger incentives to perform. Giving commissioners the ability to decommission poorly performing providers, and introducing more competitive pressures into the system, provides for a sharper focus on delivery.

Opponents of these arguments have tended to focus on both ‘philosophical’ issues (notably on the displacement / erosion of other incentives, such as commitments to public service as a good in itself)
and practical issues (such as the increased problems and cost of measurement). Here we focus on practicalities, treating broader considerations as prior to, and outside of, the scope of this paper.

**Different approaches to PbR have been developed and are being tested**

A range of services are now being commissioned using PbR. These include: Sure Start Children’s Centre pilots; drug recovery pilots; smoking cessation services; programmes to reduce the numbers of children being taken into care; re-offending schemes; employment programmes; and physical activity programmes. Again, Cabinet Office is supporting promoting the spread of these pilots (see the recent update: ‘Open Public Services 2012’).

Analysis of the approaches currently in use shows that there is no single model of PbR. Moreover, models are evolving over time. Figure 1.3 shows that it is possible to develop a typology based on the ways in which commissioners can alter the design depending upon the context and aims of their scheme. Variables include:

- The level at which results are specified, e.g.: population / neighbourhood / family / organisational / individual;
- The extent to which payment is based solely on outcomes, or whether a blend of outputs and outcomes is used;
- The significance of PbR within the contract. Contracts range from providing small increments (e.g. a 2% bonus / withheld payment) up to, *in extremis*, no outcome = no payment; and,
- Whether PbR is structured to alter incentives at the level of a system (e.g. a local health economy) or individual staff member.

Moreover, blends of these configurations are also possible. These differences suggest a need to be clear about the precise definition / approach being referred to when we use the term ‘PbR’: it can (correctly) mean different things to different people.

**Figure 1.3 ‘PbR’ describes a range of different models (which can be blended)**

Lastly, some of these schemes have combined PbR with innovative forms of ‘social’ financing – notably Social Impact Bonds. This introduces further layers of complexity that are outwith the scope of this short paper, although the issues inherent in PbR also apply (perhaps in amplified form) to these cases.

**PbR contains a series of possible mechanisms for improving outcomes**

As noted above, PbR is an attempt to change the incentives facing the providers of services by linking their rewards to the outcomes they achieve. The essential theory is that incentives matter. The main mechanism is that improvements can be achieved by changing incentives along the delivery chain. This is shown in Figure 1.4, which illustrates the theory of promoting PbR in policy leading to incentive

---

changes and thereby improved results / better value for money. While not typically made explicit in policy, this can be seen as an attempt to address the principal-agent problem by aligning interests and incentives around the achievement of an outcome that both parties regard as desirable (such as employment or a reduction in drug abuse).

**Figure 1.4 The main mechanism within PbR is a change in incentives for service providers**

Changing incentives can therefore be regarded as the principle mechanism underpinning PbR. As the Open Public Services White Paper states:

“...payment by results creates new challenges for commissioners in setting and monitoring appropriate outcomes. These need to be set to align incentives correctly between the provider and the public interest; they need to be complex enough to prevent gaming but not so complex as to undermine the flexibility of providers.”

Underneath this, analysis of Open Public Services and the current set of PbR pilots suggests a series of ‘sub-mechanisms’ or sub-theories: many of which overlap; some of which offer differing accounts. They are that PbR is:

- **A means of promoting evidence based practice.** Because providers need to be effective in order to get paid, they have stronger incentives to follow the evidence and do ‘what works’;
- **An alternative to ‘targets and terror’.** Paying providers by results can reduce the need for performance management (providers now have a financial incentive to perform, rather than coast), this substitutes a blunt instrument (‘traditional’ performance management) for a sharp tool (financial inducement / penalties);
- **A way of using competition to stimulate innovation and drive technical efficiency.** Commissioning from diverse providers = more competition = more innovation = better VFM;
- **A way of harnessing the powers of creative destruction.** All things being equal, good providers thrive (better results = more money) and poor providers exit (worse results = less money). Over time, services improve;
- **An approach to transferring risk.** In its ‘purest’ form, PbR means no outcome = no payment; Government then only pays insofar as outcomes are achieved, shifting some of the risk onto providers;
- **Performance management with teeth.** ‘Targets’ can be redefined as ‘results’ and payments attached, thereby adding to traditional approaches to contract / performance management;
- **A useful approach to improving the availability and quality of data.** PbR requires more data; making this available allows a more open view of performance, which could introduce ‘reputational effects’ analogous to league tables; and,
- **A means of improving targeting.** Payments can be structured so as to incentivise provider to focus on particular groups. This can help to address the problem of ‘cream skimming’ by paying more for achieving results with ‘harder to help’ populations.

Evidence for the effectiveness of these mechanisms within PbR agreements is currently lacking. It is not yet clear how far (and under what circumstances) the theories implicit in PbR will achieve results in practice. In its 2012 briefing ‘Local payment by results’, the Audit Commission noted that:

“Our review of UK and international research evidence found few rigorous evaluations of PbR and no complete, systematic analysis of its effectiveness.”
ICF GHK recently completed one of the few studies in this area, looking at PbR’s use in smoking cessation services in the West Midlands. We found that the approach was successful (albeit with significant challenges associated with implementation), but that contextual factors - associated with local institutional arrangements - and the relatively simple ‘subject’ of the PbR agreement were central to explaining this scheme’s success.

This suggests a need to consider the conditions under which the above mechanisms might lead to the outcomes intended. We need to know not just whether ‘PbR works’ - but how, why, for whom and under what circumstances. While waiting for these results to come in, the ‘realist review’, developed in the UK by Ray Pawson and others, holds promise in that it provides a framework for the ex ante examination of context and mechanisms, with the ability to draw in evidence from other policy areas (e.g. (how) did making information on learning outcomes widely available affect school performance?).

**DWP’s experience of PbR raises a series of important issues and lessons**

The Department for Work and Pensions, and its predecessor departments, has a relatively long history of paying for outcomes - going back at least to programmes such as Employment Training in the early 1990s. These earlier programmes have yielded some important lessons which have contributed to the development of more recent programmes which we define as PbR. Yet we would not see these programmes as ‘payment by results’ as we currently understand the term.

More recent programmes, such as the Work Programme, have three important characteristics:

1. **The proportion of the total contract value tied to results is high.** The intention is to ensure that providers cannot survive on service fees alone, with outcome-based payments providing an extra top-up. In the Work Programme for example, after three years, payment will be solely on the basis of outcomes.

2. **The trigger for payments is as close as possible to the eventual desired outcome.** In some earlier employment programmes, payments were based on simply entering a job. In the Work Programme by contrast, the maximum payment in respect of any participant among the harder to help groups is based on them remaining in work for two years.

3. **Effort is made to reduce the scope for paying for deadweight outcomes.** In some cases, such as the much-discussed re-offending pilot being run by the Ministry of Justice, this can involve an explicit calculation of the difference between achieved outcomes and the counterfactual; in others, such as the Work Programme, it can involve careful design of the outcome payment structure.

DWP’s experience of PbR programmes as thus distinguished is relatively recent; we do not yet have results which tell us how well they have worked. What we do have is reflections based on experience to date of design and implementation, some of which link back to the key features of a PbR scheme as described above.

Firstly, it is not always possible to measure the ultimate outcomes of interest owing to the time taken for these to emerge. An example of this is the Innovation Fund mentioned above, which is focused on young people who are, or are at risk of becoming, NEET (not in education, employment or training). The rationale for this as described in the prospectus is that: “Being NEET is associated with negative outcomes later on in life, such as unemployment, low pay and depression, whilst participating and gaining qualifications has a positive impact in terms of employment and wages”. In principle it would be possible to measure these outcomes directly, by tracking over a sufficiently long period; in practice this is not a useful basis for payment of providers, and it is necessary to use a proxy outcome. The risks of doing this are clear, but are reduced if there is a strong evidential basis for a causal link from the proxy outcome to the eventual desired one. For example, in the smoking cessation scheme evaluated by ICF GHK, final payments were based on beneficiaries having given up smoking for a period of three months. The full health benefits from quitting are dependent on longer-term outcomes -

---


but there is very strong evidence that those who quit for three months are at reduced risk of starting smoking again, meaning that the proxy outcome can be used with confidence. In general, the further down the logic chain one can track participants, the better.

Secondly, the full value of outcomes may emerge over a longer period than it is feasible to track them for. In the case of employment programmes, the longer that somebody remains in work (than they would not otherwise have had), the greater the value; in the case of re-offending programmes, the longer somebody goes without re-offending the better. But there are again obvious practical limits to how long we can track people for. Risk to commissioners is minimised if the direct observation period for outcomes is long enough to justify a level of payments which is acceptable to providers. If not, again it may be necessary to forecast outcomes beyond the observation period.

Thirdly, directly measuring outcomes net of deadweight is very challenging. The MoJ re-offending pilots tackle this directly, by comparing the outcomes of a cohort of participants with those of a comparison group, and paying on the basis of this difference. Such an approach is not always possible – particularly where the programme in question is a national programme.

Fourthly, while payments may be based entirely, or almost entirely, on outcomes, there are likely to be other aspects of the service delivery in which we are interested. (A common instance is the possibility that providers will offer no help in cases where their expected gains are outweighed by likely costs). It is inevitable therefore that contracts with providers will be more complex than simply paying for defined results. It is necessary to balance the desire for these additional conditions and the potential simplicity of a PbR contract.

Conclusion: PbR is not a silver bullet, but it is potentially powerful. Accounting for factors inherent in PbR will help maximise benefits and minimise risks

There are reasons to think that PbR holds at least some of the potential that its policy proponents cite. Yet there is a paucity of empirical evidence. Moreover, there are many challenges to realising hypothesised benefits in practice: implementation is more likely to be a rocky road than plain sailing. These challenges are not insurmountable, but policy makers and commissioners need to think carefully about where risk may arise, and consider whether potential risks and costs outweigh the gains. Analysts too, in approaching interventions based on PbR, should be aware of the issues inherent in these agreements.

Our analysis suggests ten main considerations:

1. **Defining and measuring ‘results’**. Desired outcomes must be clearly specified and quantifiable, in order demonstrate whether a result has been achieved. The target population also needs to be defined in a way that cannot be manipulated by the service provider. A tightly defined cohort, in combination with carefully structured payment regimes, also reduces the possibility for perverse incentives. There needs to be a process for measuring outcomes independent of contracting parties.

2. **Articulating the theory**. Being clear on the ways in which PbR is expected to work is useful in several respects. It provides an ex ante ‘sense check’ (does this account seem plausible in this case?); it thereby helps planning and contract design; it also eases the process of evaluation (to what extent did the theory hold in practice?). Moreover, it would show the expected means by which better value is to be attained (reducing costs / improving outcomes?).

3. **Attributing results to the intervention in question**. An evidence-based link between the service and the result needs to be established to ensure that providers are paid for outcomes they are responsible for. Interventions with a single and well-defined outcome are therefore most immediately amenable to use of PbR. ‘Simple’ interventions with strong evidence connecting activities to outcomes offer an easier case than complex interventions with ambiguous evidence. This links to wider methodological debates around the use of control / comparison groups and the degree of analytical precision needed to support the agreement.

4. **Specifying and administering payments**. Once the ‘result’ has been defined, the ‘payment’ must then be specified. Commissioners need to set an appropriate level of payment for agreed outcomes. This can be arrived at through a variety of approaches, e.g. historic average pricing (what the service has cost); normative pricing (what it ‘should’ cost); costs arrived at through a competitive tendering process; or – perhaps most challengingly - the level of savings (‘cashable’ or
otherwise) that the intervention generates. The appeal of linking outcomes to cashable savings is strong in the current financial climate, yet the question of cashability is thorny: it requires disinvestment in services where use is reduced, increasing (and most likely spreading) the challenge to local commissioners.

5. **Balancing risk and stimulating the market.** The movement of risk must be balanced: too little transfer of risk undermines the benefits for commissioners; too much will dissuade providers from entering the market. Providers will not take on a level of risk that they deem unlikely to enable them to achieve the rewards linked to the achievement of outcomes. This can be a particular issue for smaller independent and voluntary and community sector organisations that may not have the resources to invest ‘upfront’ and bear the risks of not achieving the desired outcomes. Defining and staging payments is also therefore critical to attracting and rewarding high quality providers. Here payment could be made entirely on the basis of outcomes or could combine outcome based payments with an element of fixed payment (and/or payment on outputs).

6. **Avoiding perverse incentives.** Without careful design the PbR model can create incentives for providers to target people most likely to achieve a result, rather than those most in need (‘cherry-picking’ and ‘cream-skimming’). Also, specifying a single outcome may lead to incentives for providers to narrow the focus of their work. The distinction between ‘that which counts and that which can be counted’ is important here: performance measures must be meaningful, rather than simply available.

7. **Considering the cost of failure.** PbR contracts pay for success and – in their ‘purest’ form do not attract costs if the outcomes are not achieved. Commissioners need to satisfy themselves that unsuccessful performance would not result in excessive harm or a situation where they have to pay for a different intervention to remedy the failure.

8. **Assessing commissioner skills and capacity.** PbR is an emergent approach and requires new skills and relationships between commissioners and providers. This is perhaps one of the most important (and often overlooked) practical considerations in making PbR work.

9. **Minimising transaction costs.** PbR typically increases transaction costs. Changes to the monitoring and payment regimes, alongside sometimes significant set-up costs and commissioning time, often increase the cost of establishing and overseeing PbR contracts. These costs then need to be weighed against the net outcomes of PbR relative to alternative commissioning models.

10. **Considering PbR relative to alternatives.** Many of the issues described above are unique to PbR; others are amplified by it; others are more generally applicable within outcomes based commissioning. Decisions on the application of PbR should not therefore be made by the test of how well it is likely to perform in and of itself, but how appropriate it looks relative to the alternatives of ‘standard’ commissioning and performance management. To the extent that empirical evaluation evidence becomes available, it will in the first instance show how well an individual intervention has worked; it will not show whether a similar intervention delivered along more conventional lines would have been more or less effective.

It’s difficult to avoid the conclusion that more research is needed. This needs to focus on whether PbR works (how, when, for whom and under what circumstances) relative to current alternatives. Moreover, there is a need to learn from current pilots and trials. Results are important, but so too is process: analysts especially must emphasise the importance of learning in this emerging area. In short, there is a way of viewing PbR as a black box (payments go in, results come out); this is not a helpful perspective if the aim is better policy development. PbR will never be a silver bullet - such ammunition is unavailable in social policy - but careful development and a focus on learning will minimise its potential as a dangerous weapon.