

Paying for results? Rethinking probation reform

August 2013

Ian Mulheirn

Summary

In May this year the Ministry of Justice (MoJ) produced the blueprint for its planned reforms of probation. This briefing note examines the detailed proposals for the payment by results element of the scheme. Payment by results, by which non state providers are supposed to be rewarded according to how much progress they make in cutting re-offending, is an essential part of the Ministry of Justice's plan to improve the effectiveness of rehabilitation services as they face deep budget cuts.

We use a financial simulation to examine the MoJ's proposals as they look to an investor considering the chances that they will make a return on their investment in rehabilitation services. This demonstrates the problems posed by statistical uncertainty around measured levels of reoffending. Our results, using generous assumptions, show that:

- Under the MoJ's proposed payment structure, providers risk making losses if they spend money on rehabilitative services.
- Providers can only be confident of being rewarded for their efforts if they achieve reductions in re-offending that are larger than the available evidence suggests is achievable, even with greater resources than are likely to be on offer under the new scheme.
- The payment mechanism encourages providers to cut spending on services and allow reoffending to drift marginally upwards.
- The proposed regime therefore creates strong perverse incentives: the opposite of how a PbR scheme should operate.
- As they stand, the plans would offer poor value for money for taxpayers and should be radically revised before the scheme is rolled-out.

We propose a number of solutions that would resolve the problems identified, improve incentives on providers, and offer much better value for money to the taxpayer.

- The reward scheme should be simplified to pay (or penalise) on a per-person basis for any observed reduction (or increase) in re-offending below the baseline level of reoffending. This would ensure that providers have unambiguous incentives to invest in services that cut reoffending.

- To maximise the level of investment in services, outcome payments should be concentrated over a reasonably narrow range of re-conviction rates.
- The Department should drop its annual 'cohort' payment model and implement a 'cumulative' payment model in its place. This would minimise the statistical uncertainty problems about which the Department is most concerned.

Introduction

Since 2010, and particularly since Chris Grayling stepped into the role of Secretary of State for Justice, the Ministry of Justice (MoJ) has been pressing ahead with plans radically to reform the probation service. The agenda has two central aims. The first is to help the Department save money to meet its tough budget settlement, under which it will be spending around one third less in 2015-16 than it was in 2010-11.¹² The second aim is to introduce a radically new way of commissioning services to reduce reoffending and therefore cushion the blow of falling spending.

In May, MoJ produced the command paper *Transforming Rehabilitation: A strategy for reform* which outlined the way in which these changes were to take place.

- Probation services are to be opened up to providers from the private and voluntary, community and social enterprise (VCSE) sector. Providers will operate in 21 separate Contract Package Areas across the UK.
- The prison estate is to be reorganised to allow more joined-up resettlement help for offenders as they move from custody to the community. This will enable providers to engage with offenders before their release.
- And, in the most eye-catching and intriguing part of the reforms, new payment incentives are to be designed to encourage non-state providers to innovate and take risks to reduce reoffending. While the bulk of spending will be made on a 'fee for service' basis for providers to deliver interventions mandated by the courts, a proportion of the contract value will be made on a 'payment by results' basis.

This briefing note examines the detailed proposals for the PbR element of the scheme. Much store has been set by the capacity for PbR substantially to reduce the level of recidivism. In our 2010 publication, *Prison Break*, the SMF described an integrated PbR model that we believed could work. The MoJ's latest proposals bear a close resemblance to those ideas. We therefore welcome many elements of the proposals. Efforts to reconfigure the prison estate, in particular, will be essential if incentives to reduce reoffending are to bite. But while the structure of the scheme seems sensible, there are many other elements to a successful PbR.

¹ Gemma Tetlow, "Post-spending round briefing, 27 June 2013", Institute for Fiscal Studies, www.ifs.org.uk/budgets/sr2013/gemma_tetlow.pdf.

² The Ministry of Justice disputes the assertion that savings are part of the reform agenda.

This paper focuses on the payment mechanism proposed by MoJ in its May document *Payment Mechanism – Straw Man* and explores whether it will in fact give providers incentives to reduce reoffending. We begin by outlining the case for PbR in principle, before going on to assess how far the current proposals fulfil these goals in practice. We then use some financial modelling to take account of the uncertainty of payments that providers will face under these plans. For a number of reasons we conclude that the proposed regime will not give providers incentives to cut reoffending, and that providers could in fact make bigger profits by cutting costs and allowing re-conviction rates to drift upwards.

The elements of the scheme design that result in these perverse incentives stem from a mixture of political priorities and concessions to providers to make the scheme sufficiently attractive for them to be involved. The unfortunate result is that this payment mechanism fails in its over-riding aim to create financial incentives to reduce reoffending. In the paper, we propose a number of changes that would resolve the problems we have identified.

Why payment by results?

For decades, the provision of public services has been hampered by central direction from Whitehall. In offender rehabilitation, as much as in employment services, professionals at the front line of service delivery are often unable to deploy their superior knowledge of their clients and their circumstances to help achieve the outcomes that everyone wants: desistance from offending behaviour in the former case, or sustainable employment in the latter. The result, in many cases, is a mediocre service that comes to be characterised by a disempowering tick-box culture for service users and professionals alike.

If the outcomes are so undesirable, why is the central planning approach to public service delivery so persistent? Is it that Whitehall fails to recognise the shortcomings? No, policymakers are all acutely aware of the problems. Some argue that centralisation is the result of Whitehall's tendency to hoard power and resources for their own sake. But this is unconvincing. The real reason for the hardiness of the status quo lies in the difficulty of designing a cost-effective alternative.

Devolving much greater autonomy to front-line professionals carries two great risks. The first is that, as in any walk of life, there is variation in the skill, capability and application of people delivering the services. Centralisation is, in effect, an insurance policy against excessive variability, albeit one with the unhelpful side-effect of limiting the upside as well as the down-side of professional autonomy. The second risk is that autonomy without accountability is a recipe for spiralling costs – either in the services themselves or elsewhere in the public budget – since those spending the money have little responsibility for the outcomes achieved.

Conventional competitive tendering - where would-be providers bid to operate the service in question – offers no solution to these fundamental problems. 'Fee for service' contracts may cut costs in a static sense,³ but the need tightly to prescribe the services that are to be delivered offers no greater autonomy to front line professionals.

³ DeAnne Julius, *Public services industry review* (London: Department for Business, Enterprise and Regulatory Reform, 2008).

Indeed, commissioners' concerns about profit-making may mean that there is, if anything, even less autonomy than under direct public provision.

Significant operational autonomy at the frontline is therefore only possible if it is combined with appropriate financial accountability. For Whitehall policymakers, wherever the autonomy lies, accountability must go with it. This reframes the challenge for public service reform into one of seeking ways to devolve financial risk: if it can be devolved effectively then so can autonomy of provision.

In some areas of public services, such as people managing personal budgets for long-term conditions, the devolution of choice and accountability to the individual level can work well. In an area such as offender management, such an approach is clearly inappropriate. Here policymakers must identify a service delivery agent to devolve the risk to. But the other challenge is defining what we mean by risk.

'Risk' in this sense, is the possibility that the state won't get what it wants from the service in question: too many reoffenders, or, in the case of employment services, too many people left without help to get back to work. 'Bad' outcomes such as these carry a significant financial cost, both to the state and to society. The financial consequences are less important than the human consequences of failure, although the two are usually closely associated with one another since more reoffending means more victims and higher costs.

Assessing the financial consequences of interventions is therefore a useful way to quantify the risks policymakers worry about. Attaching prices to outcomes – where outcomes can be clearly defined and measured – therefore offers a powerful mechanism to achieve the devolution of operational autonomy and financial risk together. This is the theory behind paying by results. If such models can be made to work, their potential for boosting the productivity of public services could be substantial.

The Ministry of Justice reform plan

As the Ministry of Justice takes further steps towards outsourcing probation services, paying by results to reduce reoffending is an important part of the package. The bulk of providers' contracts will involve a 'fee for service' (FFS), the two primary aims of which will be: to extend supervision to offenders receiving sub-12 month custodial sentences, for the first time since 1996; and to make substantial savings in order to meet the Department's challenging spending settlement.

But the new arrangements will also involve a significant PbR element to try to drive innovation and better outcomes for a given amount of money. An initial outline of the design of the PbR payment mechanism was published by the Ministry of Justice in May.⁴ The purpose of this paper is to explore some of the proposals, making suggestions for improvements to the design.

⁴ Ministry of Justice, "Rehabilitation programme - Payment Mechanism Straw Man", <http://www.justice.gov.uk/downloads/rehab-prog/payment-mechanism.pdf>.

The PBR 'straw man'

Box 1.1: Key Terms

Underlying reoffending is the rate of reoffending that would be observed if providers had such large cohorts of people as to minimise the random variation that occurs with small numbers of observations.

Measured or observed reoffending is the actual level of proven reoffending suggested by the data for a given probation trust or CPA. For reasons of random variation this could be above or below the 'underlying' reoffending rate that is determined by factors such as the case mix and the quality of services deployed.

The baseline is the average re-offending rate in previous years in a given CPA or probation trust, against which improvements or deterioration in reoffending will be measured.

Statistical significance is attained when the measured level of reoffending is sufficiently far below (or above) the baseline level that we can be confident that an underlying reduction (or increase) in reoffending has occurred.

The statistical zone of uncertainty is the range of reoffending rates for which providers will see no change in their payment because they represent reoffending rates that are not significantly higher or lower than the historical rate.

Contract Package Areas are the 21 areas of the country that providers will bid to operate in under the MoJ's proposals. These 21 areas are based on existing – sometimes merged – Probation Trust areas.

The target level is the degree of improvement in measured reoffending for providers to exceed the statistical uncertainty zone and start to get paid.

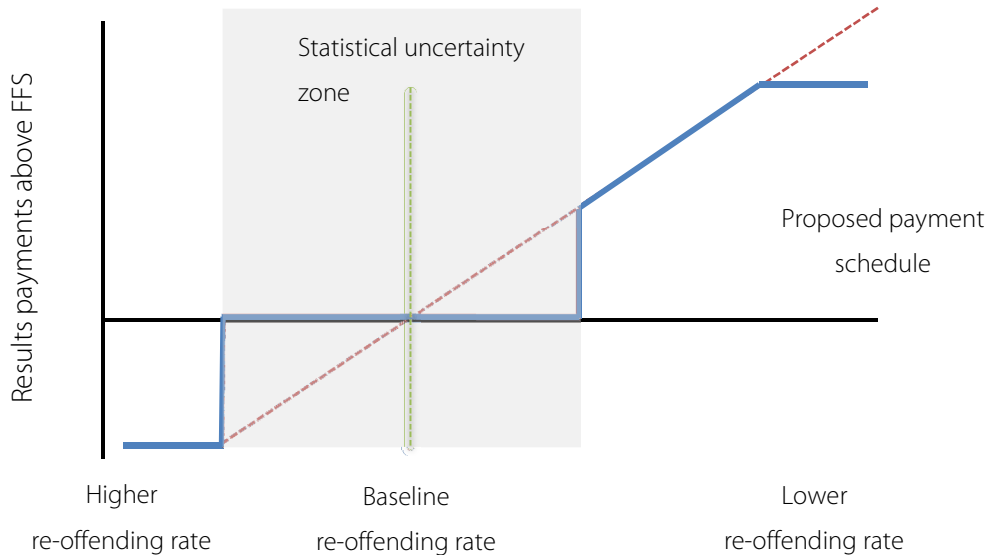
The Ministry of Justice's detailed proposals on the payment mechanism for the PbR element of the Transforming Rehabilitation reforms illustrate the challenges of designing such a scheme to ensure that government pays only for real performance success.

The diagram below outlines some key parts of the proposal. If reoffending rates remain at their historical level in a given CPA (the 'baseline'), providers will receive no additional payments above the fees for service. If reoffending outcomes are marginally better than the baseline, providers will not receive any success payments since improvements could be the result of random variation (or 'statistical noise'), rather than superior underlying performance. The same is true if performance dips somewhat below past levels: no change in payment.

Only once providers cut reoffending by a statistically significant amount will they see a jump in revenue as they receive outcome payments in respect of each additional non-offender above the baseline level. Equally, if re-conviction rates rise by a significant amount, providers will face penalties. Additional payments are made for each extra non-offender above the statistical 'zone of uncertainty', before reaching a cap for providers who perform very

well. The range of reoffending rates involved in the diagram remains to be determined: for example, the statistical zone of uncertainty depends primarily on the cohort size in the specific Contract Package Area (CPA).

Figure 1.1: The proposed payment curve



There are a number of interesting aspects of the proposed payment scheme that merit closer analysis to determine whether they will create the desired incentive for providers to cut reoffending. Here we consider four: statistical risk and incentives; financial resources and investment incentives; the ‘cohort’ approach; and the unstable baseline problem.

1. Statistical risk and incentives

Statistical uncertainty poses a big challenge for payment by results regimes. If cohorts of offenders are small, we would expect to see a high degree of random variation in recidivism rates between different providers offering exactly the same quality of interventions. The larger the cohort, the smaller the degree of random variation we would anticipate, making it easier to identify good or bad performance with a high degree of statistical certainty.

If government were to pay a provider for every additional person who doesn’t reoffend, above the historical reoffending baseline, this would create a financial risk for the taxpayer. This is because random variation in reoffending creates the possibility that government could pay for a reduction in the observed rate of reoffending in one CPA that *looks* like good performance but is in fact nothing more than the result of statistical good fortune.

However, on that payment approach, it is equally likely that government will recoup payments from other providers who *look* like they have done a bad job, when higher observed recidivism in that CPA is in fact also the result of statistical volatility.

To avoid paying for, or penalising providers for, this statistical ‘noise’ in any given CPA, the MoJ’s proposed payment schedule involves a flat payment zone either side of the baseline reoffending rate. Over this range of reoffending outcomes, payments to providers are fixed. The width of this zone is determined by the level of statistical uncertainty. So how wide is this statistical uncertainty zone likely to be in most CPAs?

The Straw Man document suggests that for annual cohorts of offenders a three percentage point reduction (or increase) in the reoffending rate, compared to the baseline, would be needed to be reasonably confident that an underlying improvement had been made (or deterioration had occurred). In other words, over a six percentage points zone around the baseline – three points either side - no additional payments will be made or penalties levied on providers for their performance.

This degree of uncertainty appears to be very roughly corroborated by some simple analysis of the variability of past re-conviction rates by probation trust, shared with us by the MoJ. However, with this analysis being based only on five years’ historic data for each probation trust, it is far from certain what the size of the statistical uncertainty zone will be in each CPA.

The probation trust data suggest that in Staffordshire and the West Midlands, for example, a 6.3 percentage point reduction in re-conviction rates would be needed to provide certainty that reoffending had been cut, implying a statistical uncertainty zone of around six percentage points either side of the baseline. Meanwhile the same figures imply that in London, only a 1.3 percentage point reduction would be necessary.⁵

In the light of this variation, our central case for the purposes of this paper is the assumption that the statistical significance zone is three percentage points either side of the baseline, as set out in the Straw Man document. This poses big problems for the scheme and provider incentives. Even if providers do make an underlying improvement of, say, four percentage points, the measured figure could come in above or below the three percentage point threshold above which they begin to receive outcome payments.

In seeking to protect taxpayers against paying out for spurious reductions in reoffending, this approach also means providers that do succeed in improving the underlying service often won’t be paid accordingly. On the other hand, providers who perform poorly, seeing underlying reoffending rise, are protected from having to repay money by the uniform payment in the statistical uncertainty zone. All of this has the effect of damaging incentives for providers to improve.

⁵ These figures, at the 95% confidence level, can be calculated from Table 24 of Ministry of Justice, “Proven reoffending statistics quarterly bulletin: October 2010 to September 2011”, <https://www.gov.uk/government/publications/proven-re-offending-statistics-october-2010-to-september-2011>.

With 13 of the 21 proposed CPAs expecting caseloads of 10,000 offenders per year or fewer, it is very questionable whether the PBR incentive will have any real bite. Even solid underlying improvements in the reoffending rate may not be rewarded, while substantially poorer outcomes will not be penalised under the plans.

If reducing reoffending was a costless activity, this dynamic in the payment regime would simply mean weakened incentives for providers to cut reoffending. But since providers will actually have to spend their own money to achieve such improvements, the proposed payment structure means that doing so could leave providers worse off than if they had spent nothing at all. This payment structure effectively offers providers a high-stakes gamble on whether to invest large amounts of money to achieve unprecedented reductions in recidivism with very significant risk that they will see no outcome payments for that investment.

Box 1.2: Simulation assumptions

- A cohort of 10,000 offenders – the average annual cohort size across the 21 CPAs will be around 11,000.
- A standard deviation of 1.5 percentage points' reduction in the measured re-conviction rate. This assumption is based on the figure used in the Straw Man document for an annual cohort. MoJ officials have provided some historic reconviction data by probation trust, which although rudimentary, tend to support this figure. Of course, it is likely that the standard deviation will be smaller in CPAs with large cohorts, and larger in areas with small cohorts.
- Each additional outcome above the baseline is assumed to result in a £5,000 payment to providers. This equates to an assumption that just over £120m is available for outcome payments and targeted at the marginal ten per cent of offenders – around 24,000 - above the baseline. It should be stressed that this assumption is little more than an educated guess at the level of outcome payments that may be offered. Nevertheless, the assumption is not central to the point illustrated here.
- Each additional outcome above the baseline is assumed to cost £3,500. In other words it is assumed that: outcome payments are set high enough for providers to be able to make a profit; and that the cost of successful interventions does not rise as providers get further into the caseload. Both assumptions are arguably generous to the MoJ proposal.
- It is assumed that there is no cap on the outcome payments on offer. In fact MoJ has said it will set a cap, limiting the up-side for providers. This is therefore a generous simplification of the MoJ proposal, made for ease of calculation.

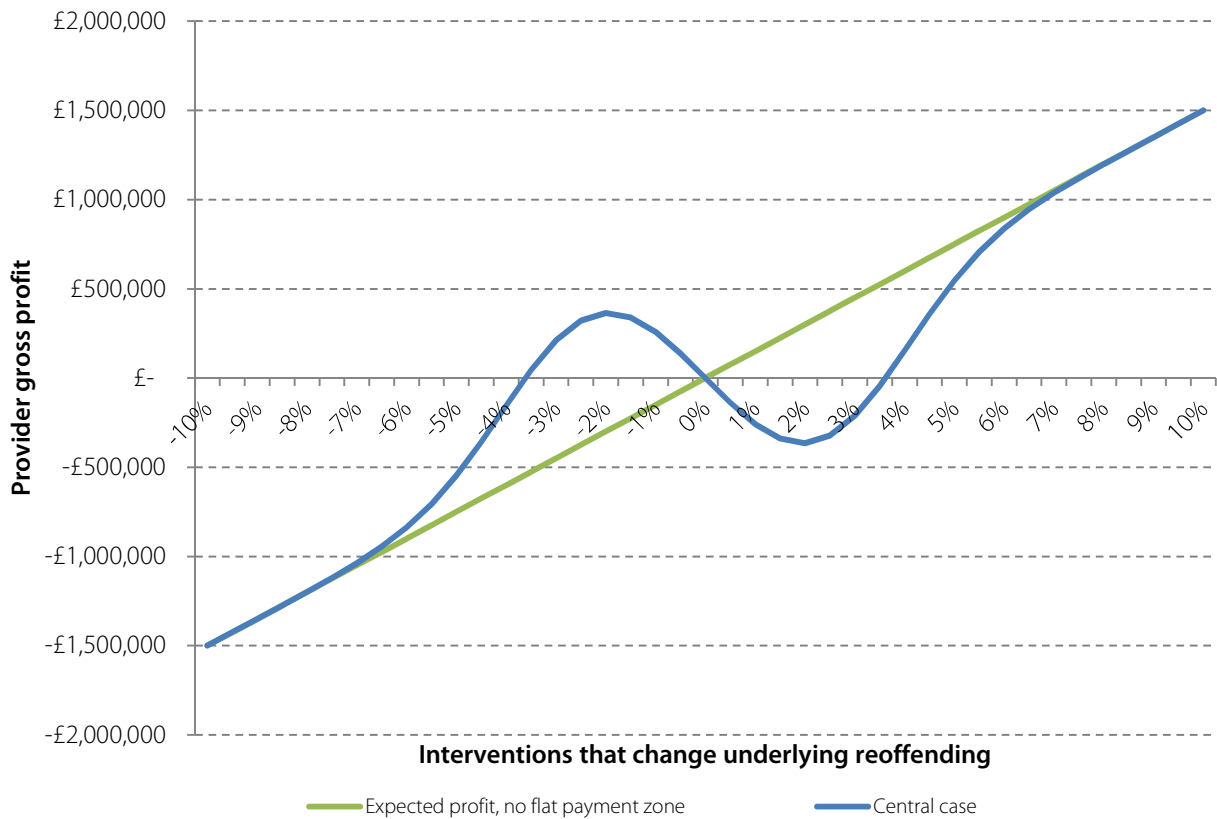
We modelled the financial incentives for providers and investors, taking into account the impact of the statistical uncertainty on the outcome payments they might receive for a given level of spending on services. Doing so reveals the real incentives facing providers under these proposals. The other assumptions contained in our model are shown in Box 1.2.

Chart 1.1 below illustrates the results. It simulates providers' expected (i.e. risk-adjusted) gross profit (i.e. revenue less the costs of delivering the interventions necessary to secure that revenue), given the MoJ's proposed payment

structure for the PBR part of the contracts, under the assumptions outlined above. The expected gross profit for a provider is determined by three factors:

- the cost of the interventions made – more intensive interventions will cost more but should also reduce underlying reoffending by more;
- the level of outcome payments that will result from cutting reoffending by a given number of percentage points; and
- the probability that statistical noise will leave the *measured* desistance rate above or below the *underlying* rate.

Chart 1.1: Investor perspective on PBR offer: risk-adjusted gross profit from different interventions



The horizontal axis represents a menu of intervention levels that the provider can expect will shift the *underlying* desistance rate by the percentages shown. For example, if the provider chooses to hold services as they are, not investing more to cut reoffending, they can expect the reoffending rate to remain at the baseline 0%. There is a possibility that, by good or bad statistical fortune the measured reoffending rate will be higher or lower than the baseline, but those possible outcomes are equally likely and therefore cancel each other out in terms of what

revenue this provider can expect (in probabilistic terms). Hence at the 0% intervention, expected gross profit for the provider is nil.

If, on the other hand, the provider decides to invest in interventions that are thought to cut reoffending by three percentage points, for example, the expected gross profit will change as shown.

In the chart, the green line illustrates how that expected gross profit would vary in the simple case where a per-person payment (or penalty) is made for each offender above or below the reoffending baseline, with no flat payment zone of the type shown in Figure 1. On this alternative payment schedule, providers could also reduce their spending on services, resulting in higher reoffending rates and penalty charges levied at the same rates, but this behaviour would push their expected gross profit into negative territory. Under this very simple payment regime, without the statistical significance zone that features in the MoJ proposal, the expected gross profit unambiguously improves for every extra pound invested in rehabilitation services. This alternative (and much simpler) payment schedule therefore gives clear positive incentives to providers.

However, the payment schedule set out by the Department looks very different to this from the investor's perspective. The incentives it introduces for providers are perverse. Since no payments or penalties are made for providers whose measured reoffending rate lies within the zone of statistical uncertainty, only a measured reoffending reduction of more than three percentage points will result in a payment. This risk-adjusted gross profit for the MoJ's proposals is illustrated by the blue line on Chart 1.1. It shows that a provider has to aim for an almost four percentage point reoffending reduction if they are to expect to avoid making a loss.⁶

In this world, no investor would target a reduction of reoffending of, say, two percentage points, since the expenditure involved in achieving it is very unlikely to be rewarded with outcome payments, leaving the investor with a loss. In other words, efforts to cut reoffending will effectively be penalised.

Only by paying for interventions that will shift the underlying desistance rate up by around four percentage points can the provider be reasonably confident of making a small margin. This is a very large hurdle for providers to overcome.

The flat payment rate in the statistical significance zone also has the perverse effect of encouraging providers to cut spending on reducing reoffending. Again, because increases in reoffending are not penalised until they reach the point of being statistically significant, this encourages providers to strip out interventions, thus saving money, and raising expected gross profit. In the example here, the profit maximising strategy is to cut services to the point where reoffending rises by two percentage points. Efforts by providers to save money, pushing up reoffending, will be rewarded under this proposed payment structure.

Needless to say, this makes no sense from the taxpayer's perspective. At the margins where we would hope to see a reoffending reduction – and worry about a rise in reoffending – from these reforms, the proposed scheme is not

⁶ Note that this is a higher degree of improvement than the 3ppt Target Level required for payment. Were the provider to target a 3ppt improvement exactly, they would face a 50% chance of being paid nothing (if the *measured* outcome slipped below 3ppts) despite having spent money on interventions. For this reason, on a risk-adjusted basis, it only makes sense for providers to target improvements well above the statistical significance zone.

a system that pays for results, but one that does exactly the opposite. Clearly the sensible strategy will be for providers to make their margin on the FFS part of the contract and spend (at best) no additional money on rehabilitation interventions.

The green line in Chart 1.1 shows how the expected gross profit would look if MoJ abandoned the idea of having a flat payment in the statistical significance zone, and instead offered to pay by results according to observed outcomes. While there would be cases in which payments were made to providers who only *appeared* to have cut reoffending, these would be off-set by savings on providers who appeared to have increased it, leaving MoJ no worse off. More importantly it would set clear incentives that the more a provider cuts underlying reoffending the more likely they are to make a profit. This is surely central to any PBR scheme worthy of the name. By contrast, the MoJ proposal would cost taxpayers the same but give providers unambiguously worse performance incentives.

Varying the assumptions

Smaller cohorts – bigger risks for providers

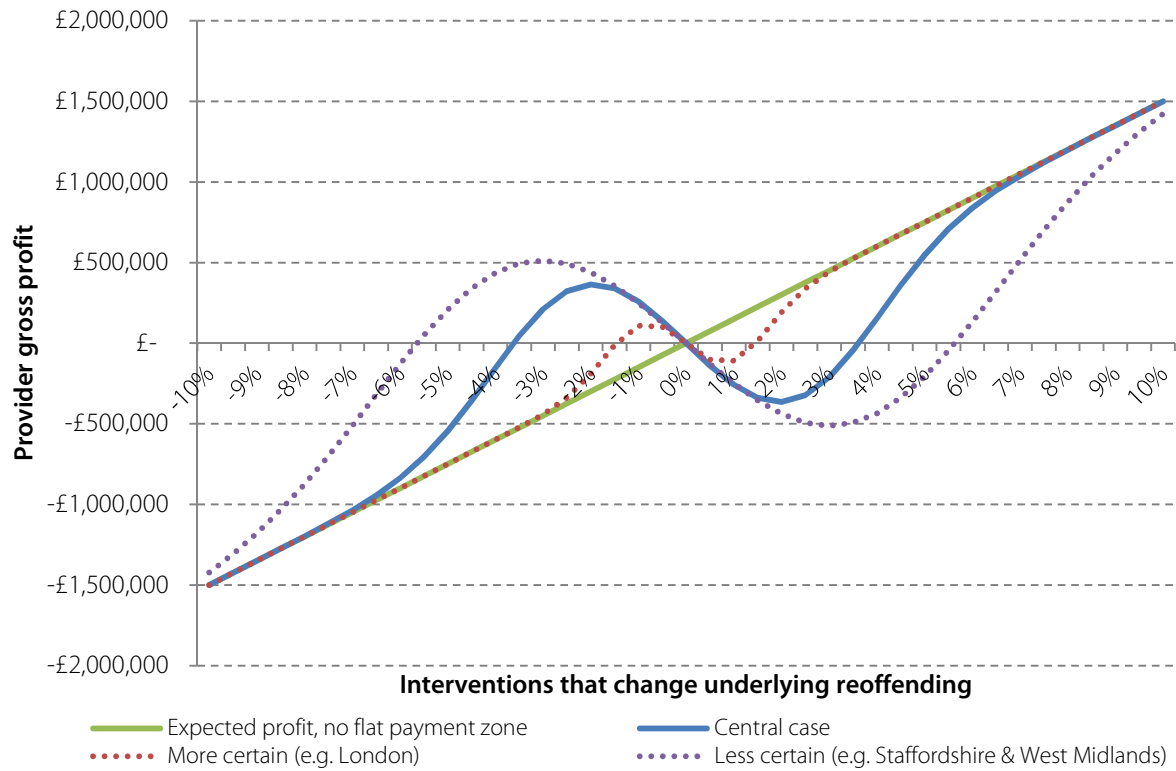
The statistical uncertainty zone is likely to be bigger in areas with smaller cohorts, such as the CPAs of Essex and South Yorkshire, and smaller in the largest CPAs, such as London, which will have a cohort five times bigger. This means that the problem identified above could be less severe in London, but more severe in the smaller CPAs, the majority of which will have caseloads less than one third the size of London's.

The smaller CPAs were proposed in the Department's response to the consultation process, in the wake of concerns that the initial proposal for 16 bigger CPAs meant "too few contract package areas for smaller providers to be able to participate easily".⁷ Ironically, providers bidding for the smaller CPAs could face much bigger risks that they will not be rewarded for improving underlying performance. This is simply because the statistical significance zone should be larger in these areas. Of course, this also means that, in these zones, providers will also be less likely to be penalised for poor performance.

Chart 1.2, below, illustrates how the disincentive for providers to invest in cutting reoffending changes according to the degree of statistical uncertainty one might expect in smaller and larger CPAs. The central case remains as in Chart 1.1, with the zone of statistical uncertainty three percentage points either side of the baseline. If the zone were to narrow to plus/minus 1.5 percentage points in the larger CPAs, the expected gross profit curve would look like the red dotted line on the chart. If the zone were to widen to plus/minus five percentage points in the smaller CPAs, the expected gross profit curve would look like the dotted purple line on the chart. Historical data from the probation trusts suggests that this (entirely illustrative) degree of variation between CPA would not be unusual. The perverse incentives in smaller CPAs could therefore be huge.

Chart 1.2: Risk-adjusted gross profit and cohort size

⁷ Ministry of Justice, *Transforming rehabilitation: a strategy for reform* (London: HMSO, 2013), 16.



Intervention costs rise for harder cases

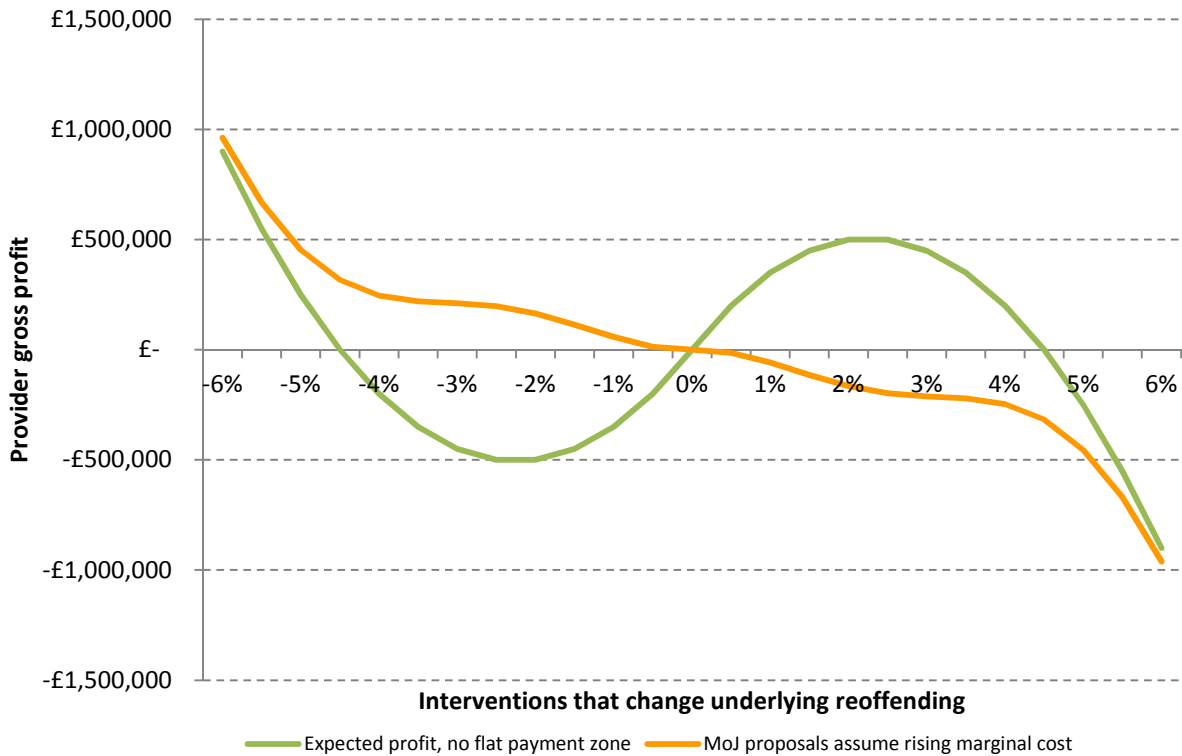
The analysis thus far has assumed that the costs of services to reduce recidivism are uniform for each extra offender who desists. In reality the cost of successful interventions is likely to rise as services begin to target harder cases. This is an intuitively reasonable assumption: while getting some offenders accommodation and a job might help them desist, for harder cases a much wider range of services – substance abuse, training, mental health services etc. - might be required, costing substantially more.

The orange line in Chart 1.3, below, shows what the risk-adjusted gross profit curve would look like for providers under the plans if they think that costs of achieving desistance rise gradually as they move through the cohort. In this model the marginal cost of desistance is assumed to go on rising by £1,000 per offender for each half per cent of the cohort. This means it is assumed that, among the first half percentile (50 of the 10,000 offenders in this CPA), providers have to spend just £1,000 per person to achieve desistance. For the next half percentile, the costs of successful interventions rise to £2,000 per person, and so on. The total cost of reducing underlying reoffending by one percentage point is therefore £150,000 (£50,000 plus £100,000).

The consequences of this 'rising marginal cost of desistance' are dire for provider incentives: once providers hit three percentage points of desistance, the marginal costs of further reductions in reoffending (£6,000 per person) outweigh the payments on offer from the Department (£5,000 per person, by assumption). This model is entirely assumption based and different bidders will make different assumptions about the cost of services. But any plausible assumption of rising marginal costs raises the question of whether there is any point where the risk

adjusted gross profit of cutting reoffending turns positive for providers who try to cut reoffending. As the orange line illustrates, under these assumptions there is no prospect of being rewarded for cutting reoffending.

Chart 1.3: Risk adjusted gross profit assuming rising marginal costs of desistance



Clearly, the possibility of rising marginal cost of interventions is not a result of the MoJ’s pricing proposals. If the costs of turning around the most prolific offenders are very high, that will pose a problem for any PbR scheme. But the statistical uncertainty zone in the Straw Man proposal discourages providers from making any progress under these assumptions. By contrast, eliminating the statistical uncertainty zone from the payment schedule yields the green line in Chart 1.3: unambiguously positive profits for providers who make marginal progress in cutting reoffending.

To avoid these problems, we propose that the payment curve be simplified to pay (or penalise) on a per-person basis for any observed reduction (or increase) in reoffending below the baseline. This means dropping the proposal for a flat payment in the statistical significance zone.

2. Is the financial incentive likely to be sufficient?

The Ministry of Justice spends almost £1bn per year on delivering sentences in the community. In contracting out probation services the department will be looking for large savings, given its tough departmental budget settlement to 2016. The bulk of the remaining budget will be devoted to the Fee For Service payment that is intended to cover the costs providers will face in delivering mandated activities associated with the offender's sentence. The Government has yet to give any indication of how much cash will be devoted to paying providers according to the reoffending results they achieve. But on a rough estimate, we might expect that between £100m and £200m will be available.

This implies that there will be limited scope for funding interventions. With providers managing around 236,000 starts each year, even if £200m of payments can be structured in such a way as to attract the same amount of spending on services, expenditure would be well under £1,000 per head. If only £100m is available, then the best level of resources per head that might be expected is around £500. By contrast, in the Peterborough pilot, £5m is being spent on services to 3,000 offenders, amounting to just under £1,700 per person. Under the national scheme, it seems likely that providers will therefore have between a quarter and a half of the resources per person available at Peterborough.

In the Peterborough pilot, performance is measured on the frequency of reoffending behaviour, rather than the 12 month reconviction rate proposed as the central measure for the proposed programme. Nevertheless, as many people have argued, there is a strong connection between reduced frequency and desistance. This allows us to get a sense of what kind of reduction in one-year reconviction rates might be possible with the resources available at Peterborough.

In the period from October 2008 to June 2010, prior to the start of the Peterborough pilot, one-year reconviction rates from the prison averaged 40.3%, 1.1 percentage points above the national reconviction rate of around 39.2%. Since the start of the pilot, the one-year reconviction rate, at 38.6% has been slightly below the national figure of 40.4%. A proper evaluation has yet to be completed, and these raw figures take no account of any change in the client mix in Peterborough Prison.⁸ Nevertheless, this suggestion of a 2.9 percentage point reduction in reconviction rates gives a sense of the scale of improvement that is possible with the resources likely to be deployed on this programme.

What we can say is that with less than half of the resources per head, there is no reason to think that providers in the proposed scheme will be able to do better than the Peterborough pioneers. Data from the Doncaster pilot is even harder to read and to compare in advance of any formal evaluation, but the early figures do not imply that greater reductions in reoffending are likely.⁹

⁸ The OGRS scores – which predict likelihood of reoffending – for Peterborough offenders appears to have changed by only a little since before the start of the scheme. The Peterborough caseload seems to have become only marginally less likely to re-offend (absent any different interventions), relative to the national cohort, since the start of the programme.

⁹ While the raw data available imply a reduction in reoffending of between 1.7ppts and 5.7ppts, these reductions also coincided with a four point fall in the OGRS scores of the cohort. This change in OGRS scores may explain away any observed change in headline reoffending once a formal evaluation is undertaken (see for example: Mia Debidin (ed.), *A compendium of research and analysis on the Offender Assessment System (OASys), 2006-2009* (London: Ministry of Justice, 2009), 101).

For providers even to have less than half of the resources available at Peterborough assumes that investors will be willing to invest substantial amounts of money in pursuit of outcomes. But the analysis of Section 1, combined with early data from Peterborough and Doncaster, means there is nothing to suggest that the scale of interventions providers could afford would take their outcomes beyond the statistical uncertainty zone (plus or minus three percentage points for a cohort of 7,000 offenders). For this reason it seems unlikely that providers will be willing to invest comparable amounts to that available in outcome payments, further reducing the likely resources in the scheme and the consequent change in reoffending rates.

To maximise the level of investment in services, the flat payment zone should be dropped and outcome payments concentrated over a reasonably narrow range of reconviction rates. This would ensure that the financial incentive to improve outcomes actually bites.

3. The 'cohort' approach to measuring results

In measuring outcomes from the scheme and paying for them, MoJ proposes to use a cohort approach. This means that they will pay for results on the basis of observing three month 'batches' of offenders. Reoffending rates will be measured for each three-month cohort in an effort to close the time lag between intervention spending and the provider receiving an outcome payment.¹⁰

Even taking offenders in three-month cohorts leaves a long time between spending on interventions and payment for the subsequent results. This is because the three-month period is then followed by a 12 month window in which reoffending outcomes are monitored, and a further six month period to allow for any lag between offences and conviction. Finally, a further four months are required to collate and analyse the data before payments are made. The result is a gap of up to 25 months between expenditure and PBR revenue. The longer the lag, the higher the cost of interventions, since providers have to finance the investment for longer (although see below, on the 'PbR Foundation Payment'). Taking 12-month offender cohorts would extend that cashflow gulf to 34 months.

There is a critical trade-off here. Shorter cohort periods obviously make for smaller cohorts of offenders; but lower numbers of offenders widen the statistical uncertainty zone, making it less likely that providers who have reduced reoffending will get paid. This exacerbates the problems outlined in Section 1: the larger the statistical uncertainty the greater the disincentive to reduce reoffending. In other words policymakers face an apparent trade-off between the incentives on providers and the cost of financing interventions.

This will be a particularly large problem in the smaller CPAs, where three-month cohorts could be fewer than 1,500 offenders, and just one sixth of the size of a three-month cohort in the largest CPA. The very wide statistical uncertainty zones associated with three-month cohorts – around five percentage points either side of the baseline according to the Straw Man document – would effectively mean that interim payments offer no incentive to cut reoffending anyway.

¹⁰ Ministry of Justice, "Rehabilitation programme - Payment Mechanism Straw Man", 12.

If the justification for the quarterly assessments is to ease cashflow problems for providers it seems both unhelpful and unnecessary. Unhelpful because small numbers of offenders will make the zone of statistical uncertainty far too large for any providers to (expect to) get paid. Unnecessary because the Department proposes a 'PBR foundation payment' – an up-front payment that reflects an assumed substantial reduction in re-conviction, but which will be clawed back if providers do not achieve reductions in re-conviction. This loan from MoJ effectively removes providers' cashflow problems, raising the question of why the Department thinks it is necessary to have such small cohorts to reduce the payment lag.

To resolve the first of these problems, the MoJ proposes subsequently to adjust the results of three month cohorts once they have results for a full year of offenders. This seems sensible in principle, if complicated in practice, and could reduce the size of the statistical uncertainty zone. Nevertheless, in most CPAs the zone will remain large because of this cohort approach, encompassing most attainable levels of underlying outcome improvement. As a result, the proposals turn incentives to reduce reoffending upside-down for the reasons outlined in Section 1. There is a better approach that could have the effect of minimising this damaging statistical uncertainty.

The department should instead offer to pay providers for observed reductions in reoffending beyond the baseline, on a *cumulative* rather than a cohort basis. Initially this would mean making payments to providers where reoffending appeared to be below expectations and clawing back money where reoffending appeared to be above the expected level, regardless of the degree of statistical certainty. As more offenders flow through the scheme, the number of outcomes would cumulate, and the payments could be adjusted accordingly. Towards the end of the contract period, the number of outcomes would be many times higher than the size of an annual cohort, minimising the statistical uncertainty, and with payments being adjusted all the time to reflect the best available picture of overall provider performance. This cumulative payment model would eliminate a huge amount of statistical uncertainty, resolving the perverse incentives in the proposed plan. It would also ease cashflow problems.

We propose that the Department drop the cohort payment model and implement a cumulative payment model in its place. This would minimise commissioners' concerns about the statistical significance of measured outcomes.

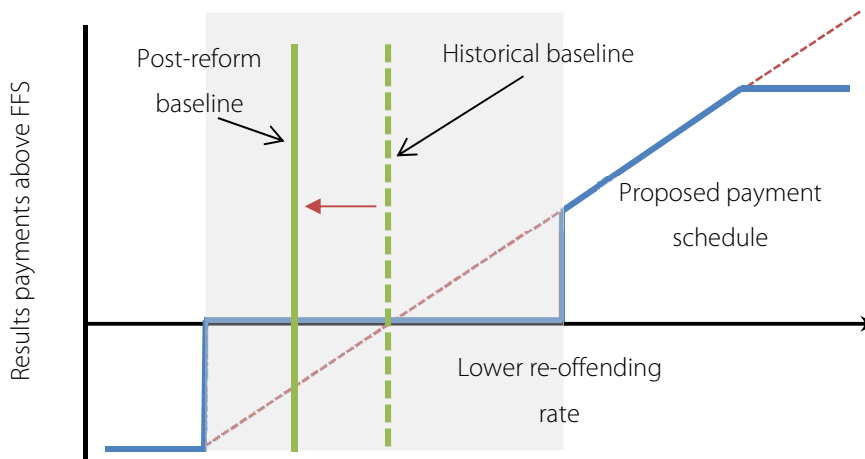
4. Is the baseline stable?

A fourth aspect of the proposed payment mechanism risks exacerbating the problems highlighted above. The 'baseline' re-conviction rate for each CPA will be the expected level of reoffending in each area determined by the most recent pre-reform re-conviction data. This 'non-intervention' re-conviction rate will be adjusted according to changes in the case mix of the cohort, as measured by offenders' OGRS Scores.¹¹

¹¹ Ibid., 9.

The diagram below shows this historical baseline level as the green dashed line, surrounded on each side by the statistical uncertainty zone, within which MoJ is not prepared to pay for measured reductions in reoffending (or to levy penalties for increases in reoffending).

Figure 1.2: The impact of a lower baseline



In the past, predicted reoffending rates for given types of offenders have been reasonably stable. This provides some confidence that the past baseline will remain relevant under the new programme as the level providers can be expected to achieve if they continue providing services as they are at present.

However, the scale and speed of the reform programme creates a significant risk that the past will be no guide to the future. As we have seen in other large scale outsourcing attempts, such as the Work Programme, performance often deteriorates temporarily after a new programme is launched and takes time to recover to its expected performance level. With substantial savings likely to be made from the almost £1bn spent on delivering sentences in the community, a drop in performance for those currently served seems inevitable. There is some scope for improvement in the baseline among the newly supervised sub-12 month group, but given the likely scale of overall savings and the radical nature of the reform, it seems very unlikely that the historical baseline will be a stable or reliable way to measure provider performance.

The result of a downward shift in the baseline would be to further weaken any incentive to invest in cutting reoffending. The ‘target Level’ of performance – the point at which the statistical uncertainty zone ends and payments begin - is now even further away from new baseline because it is set with reference to the old baseline. Investing to cut reoffending will therefore almost inevitably result in lower gross profit for providers than not doing so.

We propose that the MoJ consider running the new scheme entirely on a fee for service basis for the first two years, while the new system beds down, in order to establish a new baseline performance level on which to base the payment by results system.

Conclusion

The payment mechanism proposed by the MoJ is the key tool by which they hope to transform rehabilitation. But as they stand, providers have incentives to cut costs and marginally increase reoffending rather than reduce it. This situation is the unintended consequence to several choices made for political and commercial reasons.

- First the Department does not want to be open to the accusation that it is paying for statistical noise. This is a reasonable aspiration but the proposals result in the scheme paying for failure and penalising results. There are better ways to achieve this goal without wrecking the financial incentives.
- Second the Department has sought to reduce the lag between expenditure and outcome payments for providers, and therefore to reduce costs all round. This is also a reasonable aspiration, but the impact is to increase the statistical uncertainty, further undermining incentives. Further, the cashflow concerns this measure is designed to address are resolved by the Department's 'PbR Foundation Payment' – effectively offering public rather than private finance for services.
- Third, the department is seeking to make huge savings. This means that only a limited amount of cash will be available for the PbR part of the scheme. Combined with the problems of statistical uncertainty it is highly doubtful that any providers will see virtue in spending on additional rehabilitation services given the limited evidence on what is possible with far greater resources.
- Fourth, the Department wants to see no deterioration in reoffending levels. But this will be a huge challenge in the early years of such a large reform. This makes it even less likely that providers will see much to be gained by investing in rehabilitation services.

The unintended consequences of these plans are that this is not a scheme that will pay for results. Indeed it will likely penalise providers for trying to achieve them. These payment proposals, far from improving the productivity of probation in order to cushion the blow of falling spending, risk reducing it. This paper has proposed a number of ways to make the scheme work better and achieve the policy's stated aspirations.

This paper is kindly supported by:

