Payment-by-Outcome in Welfare to Work

Lauren M. Cumming
This case study series was an input into the report *Payment by Outcome: A Commissioner’s Toolkit*. The case studies do not provide an explanation of how to design a payment-by-outcome system in the sectors studied. Rather, through the case studies, the authors sought to understand the challenges involved in using payment-by-outcome and the tools that have been employed to cope with them.
About the 2020 Public Services Trust

The 2020 Public Services Trust is a registered charity (No.1124095), based at the RSA. It is not aligned with any political party and operates with independence and impartiality. The Trust exists to stimulate deeper understanding of the challenges facing public services in the medium term. Through research, inquiry and discourse, it aims to develop rigorous and practical solutions, capable of sustaining support across all political parties.

In December 2008, the Trust launched a major new Commission on 2020 Public Services, chaired by Sir Andrew Foster, to recommend the characteristics of a new public services settlement appropriate for the future needs and aspirations of citizens, and the best practical arrangements for its implementation.

For more information on the Trust and its Commission, please visit www.2020pst.org.

The views expressed in this report are those of the author and do not represent the opinion of the Trust or the Commission.
Contents

1 Introduction 5

2 Payment-by-Outcome in Welfare to Work in Three Countries 6
   Outcomes 7
   Population 7
   Measures 7
   Incentives 8
   Procurement 8
   Managing gaming 9

3 Why Payment-by-Outcome? 10

4 Approaches to Service Delivery 12
   ‘Work first’ 12
   ‘Human capital development’ 13
   A mixed approach 14
   Effectiveness 14
   Implications for payment-by-outcome 15

5 Selecting Primary Outcomes 17
   Rationales for work activation programmes 17
   Programme funding 17
   Multiple departmental goals 19

6 Segmenting the Population 20
   Setting prices 20
   Measuring impact 21
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>Selecting Measures</td>
<td>23</td>
</tr>
<tr>
<td>8</td>
<td>Designing Incentive Systems</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Financial incentives</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td>Other incentives</td>
<td>30</td>
</tr>
<tr>
<td>9</td>
<td>The Procurement Process</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Pathways to Work</td>
<td>33</td>
</tr>
<tr>
<td></td>
<td>Working Nation and the Job Network</td>
<td>35</td>
</tr>
<tr>
<td>10</td>
<td>Managing the System Over Time</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Data and information</td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>Monitoring</td>
<td>39</td>
</tr>
<tr>
<td>11</td>
<td>Conclusion</td>
<td>43</td>
</tr>
<tr>
<td>Appendix</td>
<td>The United States</td>
<td>44</td>
</tr>
<tr>
<td></td>
<td>Australia</td>
<td>50</td>
</tr>
<tr>
<td></td>
<td>The United Kingdom</td>
<td>55</td>
</tr>
<tr>
<td>Endnotes</td>
<td></td>
<td>63</td>
</tr>
</tbody>
</table>
Introduction

A spell of unemployment is a difficult time for most people. After losing their jobs, people often lose their confidence and motivation as well, which can make finding new employment difficult. Welfare to work programmes help jobseekers find work through facilitating job matching, improving basic job search skills such as CV writing and interviewing, helping jobseekers overcome non-vocational barriers to work and, perhaps most importantly, increasing jobseekers’ self-confidence and belief they can succeed. Such programmes play an important role in promoting social inclusion and reducing welfare dependency, which cost the British taxpayer £10.3 billion in 2009/10.¹

Years of experience of performance management in welfare to work have generated many lessons about the tools required to ensure payment-by-outcome delivers the best quality service and outcomes for jobseekers. One example of using measurement to improve welfare services dates to 1948, when the department head of a US state’s public employment agency instituted new records to monitor employment advisors’ performance.² Since this early use of reputational incentives, the use of financial incentives to improve performance has become increasingly commonplace. Beginning in the US in 1982, countries around the world, including Australia, the UK, Denmark, the Netherlands, France and Germany, have implemented payment-by-outcome in welfare to work services. Focusing on the US, Australia, and the UK, where the most research has been conducted, this case study seeks to understand the functioning of different programmes in order to derive general lessons about the tools required to ensure payment-by-outcome works.
Payment-by-outcome was first implemented on a large scale in welfare to work in 1982, when the Job Training Partnership Act (JTPA) was rolled out in the US. In 2000, the JTPA was replaced by the Workforce Investment Act (WIA). Alongside these federally-funded schemes, some states also have outcome-based welfare to work programmes.

Australia began to experiment with payment for outcomes in its Working Nation programme in 1994. Working Nation was replaced by the Job Network in 1996, and in 2009 the Job Network evolved into Job Services Australia.

The UK’s first payment-by-outcome programme, Employment Zones, was rolled out in 2000. Subsequent programmes, including Pathways to Work and the Flexible New Deal, have built on the model of another welfare to work programme, the New Deal, which operated at the same time as Employment Zones. These programmes saw the government transfer increasing levels of risk for the achievement of outcomes to providers. A brief comparison of some of the programmes that operated in each of the three countries follows. More detailed descriptions of the programmes can be found in the appendix.

Although all of these programmes paid for outcomes, the differences among them are striking. The schemes had varying policy objectives, divided clients among programmes or streams of programmes in different ways, measured outcomes differently, rewarded different kinds of outputs and outcomes, were procured differently and have managed problems such as gaming using a variety of tools.
Outcomes
In the American JTPA programme, the primary outcome specified in the legislation was to maximise the “return on investment in human capital”, where human capital was defined as a “set of skills and knowledge” acquired through education and training. The population targeted for intervention suggests poverty reduction was another aim. In Australia and the UK, the focus has been more explicitly on ensuring benefits claimants find employment, promoting social inclusion and reducing the cost of unemployment to government.

Population
The client groups targeted and the ways in which outsourcing was used to serve different groups of jobseekers varied quite dramatically. Certain American programmes have been aimed at individuals living below the poverty line who do not have to be unemployed to qualify to receive services. In Australia, jobseekers are classified according to their level of disadvantage, and this determines their eligibility for services. Those categorised as closest to the labour market are eligible only for ‘job matching’ services, while the most disadvantaged qualify for ‘intensive assistance’. The private and voluntary sectors play a role in providing services for all clients. In the UK, eligibility for programmes has historically been determined by length of unemployment and the type of benefit claimed. In general, those unemployed for less than twelve months qualify for services from Jobcentre Plus, the public employment service, while the long-term unemployed are served by private and voluntary sector providers specialising in personalised case management, with a separate programme for those with disabilities. The Coalition Government is in the process of commissioning its Work Programme, which will be for all jobseekers regardless of benefit type. Referrals to this programme will be more flexible than under FND, with certain jobseekers (for example those classified as ‘facing significant disadvantage’) having access after only three months of unemployment.

Measures
The measures selected to evaluate performance reflect the variation in primary outcomes. Under America’s JTPA, clients’ average wage on entering employment was one of four measures used to determine an agency’s eligibility for an outcome award. The US is the only one of the countries studied here that places emphasis on increasing jobseekers’ earnings, used as a proxy for increased human capital.
Australian and UK programmes have focused on measures of sustainable employment, namely the number of jobseekers achieving 13 and 26 weeks in work. In the UK, longer measurement periods will feature as part of the Government’s new Work Programme.

**Incentives**

Incentive systems also differ. Under JTPA, the states had discretion over the ways in which eligible agencies were rewarded. In most states, agencies simply had to meet a certain standard of outcome achievement to receive an award. However, some states gave the entire bonus to the best-performing agency while some divided it among all agencies performing above a certain threshold. Other states varied the award according to performance relative to the standards, so that agencies that far surpassed their targets received more than those that simply met their standards.8 The Employment Zones programme in the UK was quite different. Providers received some funding upfront, a proportion of which they were entitled to retain if jobseekers found work within a certain period of time. Providers also received standard payments for jobseekers placed in employment and those who retained work for 13 weeks. Subsequent UK programmes such as Pathways to Work and the Flexible New Deal have similar incentive systems to Australia’s Job Networks, with providers receiving service fees of varying levels for accepting jobseekers onto their programmes and then earning outcome payments when jobseekers remain in work for 13 and 26 weeks. A key difference between these two countries is that Australian programmes have also tended to measure and reward education and training outcomes, while in the UK the incentives have steered providers to focus much more heavily on employment outcomes.9

**Procurement**

Programmes were procured differently. In Australia, commissioners initially fixed the price for intensive assistance services so that providers had to compete based on the quality of the service they were offering, stimulating innovation and avoiding a winner’s curse situation. In the UK, the Pathways to Work programme was procured through a very competitive, price-focused competition. Pathways Providers have experienced extreme financial pressure which has forced them to deliver a minimal service. All providers have failed to meet their targets for job outcomes.
Managing gaming

Finally, gaming has been managed in different ways. Gaming occurs when providers increase their measured performance in ways other than by increasing their performance in relation to the primary outcome sought. In Australia, reputational incentives to deliver high quality services are strong, as providers are rated using published ‘star ratings’ and contracts of providers that perform above a certain level have been automatically renewed. Employment Zones in the UK used a prize to discourage providers from parking those furthest from the labour market. Providers were encouraged to achieve as many job outcomes as possible since the level of performance required to win the prize was not disclosed.

This brief analysis demonstrates that payment-by-outcome can be applied to welfare to work programmes in many different ways. Each programme has a specific context which affects its design and level of success. This case study seeks to help policymakers and commissioners better understand the tools they can use to overcome the challenges involved in implementing a payment-by-outcome system in welfare to work.
Why Payment-by-Outcome?

Payment-by-outcome is not the only method that can be used to encourage service providers to achieve commissioners’ goals. In the past, employment services were largely delivered by public agencies which had mandates to provide certain services to eligible jobseekers, including dispensing benefit payments and providing basic job search facilities. Governments have also contracted services using highly prescriptive contracts which pay for inputs. Why are so many governments now moving to a payment-by-outcome approach?

Payment-by-outcome is related to but distinct from payment-by-results and invest-to-save. Payment-by-results is an advanced form of performance management where providers are paid based on the delivery of outputs or outcomes rather than inputs or processes. There is often an assumption that performance incentives will be very strong and providers will carry substantial financial risk, although these are not essential features of the model.

Payment-by-outcome differs from payment-by-results in its focus on the ultimate policy or programme objective. Payment-by-outcome involves clearly specifying the primary outcome of the service and, where this cannot be measured directly, developing a measurement system using intermediate outcomes or high-level outputs as surrogates.

Invest-to-save requires providers to finance the initial investment necessary for service improvement, with payment being made out of the savings that are directly attributable to that investment. In this way, it creates very strong incentives for providers to deliver outcomes. While invest-to-save can be added to a payment-by-outcome model in order to obtain an alternative source of funding service improvement, the incentives, and correspondingly the fear of failure, may be so strong in such a system as to prevent providers from delivering significant innovation in service delivery.
There are several advantages of using payment-by-outcome in welfare to work. Most obviously, it encourages providers to deliver better outcomes through the use of powerful financial incentives. Sometimes, the financial incentives serve to clarify the goals of policy. For example, payment-by-outcome has cemented in the ‘work first’ approach commissioners have been advocating for years, but often unsuccessfully. It is now abundantly clear to providers that their primary aim is to get jobseekers into work, and not simply to ensure they complete training programmes. Despite the fact that Wisconsin had been promoting an emphasis on employment for some time, when payment-by-outcome was implemented in the state the director of one of the providers said, “I had thought we were supposed to be getting recipients general equivalency degrees [the equivalent of a high school diploma]”. The true purpose of the programmes was clarified by the incentive structure of the payment-by-outcome system.

The model also gives providers the space and creates stronger incentives to find innovative ways of achieving outcomes. This includes improving the overall quality and efficiency of service delivery as well as increasing the level of personalisation for each individual client. Thus an exploration of ‘what works and for whom’ can occur more easily than under a contract that is very prescriptive about the inputs providers should use. This is important in welfare to work where each client will have a different mix of barriers to work which may need to be addressed in different ways.

Furthermore, payment-by-outcome may lead to the integration of departmental funding streams, since more than one department is likely to have a stake in a high-level outcome. There is some evidence of this occurring in the UK in the Pathways to Work programme. Providers of this programme, which serves clients who have been assessed as having some form of disability, are required to include a Condition Management Programme in their range of interventions. Where the public sector delivers Pathways to Work, these Condition Management Programmes are joint ventures between the Department of Health and the Department for Work and Pensions, as the benefits of patients better managing their health conditions accrue to both departments.

Finally, payment-by-outcome increases the accountability of providers to taxpayers for the effective use of funds, as providers are only paid for achieving concrete results. In the welfare sector, where taxpayer buy-in is crucial to maintaining the redistribution settlement, heightened accountability is especially important.
Approaches to Service Delivery

It is important for commissioners to have a broad understanding of the linkages between inputs and outcomes, as the decisions commissioners make about the design of the payment-by-outcome system, particularly in relation to the selection of primary outcomes and measures of progress towards them, may encourage or prevent the use of a particular approach to service delivery and affect the scope for innovation by providers. One of the principal aims of payment-by-outcome is to enable flexibility and innovation in service delivery, so it may seem unnecessary for commissioners to understand how providers might deliver services. However, because of the system design decisions they will need to make, it is unlikely that commissioners can be entirely agnostic with regards to approaches to service delivery. In welfare to work, there are two competing theories about how the long-term unemployed can most effectively be helped into sustainable employment, known as the ‘work first’ and ‘human capital development’ approaches.

‘Work first’
Proponents of the ‘work first’ approach (also known as the ‘labour market/force attachment’ approach) argue that being in work promotes good work habits and skills which can in turn lead to better paying jobs. Moreover, it is a cost-effective way of acquiring skills (compared with education programmes). Advocates argue that rather than welfare recipients remaining on benefits or participating in further education or training, the best thing for them is to accept the first available employment opportunity, and then work their way into their preferred sector or position. Work first programmes tend to emphasise short-term job search
assistance and their success is measured by how many jobseekers are placed in work (rather than how many complete training or education or achieve other outcomes).

The UK and Australia have adopted the work first approach, albeit with a focus on sustainable employment which elevates the importance of ensuring there is a good match between the jobseeker and the employment opportunity. The UK’s approach was evident in the incentive structure of Employment Zones. Providers that found employment for jobseekers quickly were able to retain some of their benefit entitlements, but providers also received outcome payments for clients who retained work for 13 weeks as an incentive to ensure a good match between job and jobseeker. While the Australian government has not formally described the Job Network as having a work first approach, this is the broad thrust of the measures and is evident in the incentive structure which is heavily weighted towards rewarding employment outcomes. In the US some programmes are highly work-oriented while others are more ‘human capital development’ focused.

‘Human capital development’
The alternative is the ‘human capital development’ or ‘life first’ approach, which emphasises the importance of finding work at a reasonable wage and with opportunities for salary progression and ongoing employment. The human capital development approach recognises that assistance to deal with non-vocational barriers to work may be essential for some jobseekers. These latter may be especially important for those that have been unemployed for a long time and may not be ready for employment in the short to medium term. Welfare to work schemes that focus on human capital development tend to emphasise longer-term skill-building activities, especially basic education, before entry into the labour market. Those advocating this approach argue that it leads to better and more stable jobs, increased income and a reduced likelihood of participants reclaiming benefits.

Australia has used the human capital development approach in its Personal Support Programme for highly disadvantaged jobseekers with complex, non-vocational needs. The Personal Support Programme enables providers to be rewarded for making progress with participants, as payments are made for non-employment outcomes. Clients completing the full two-year programme are assumed to have achieved a social outcome such as increased self-confidence and attract an outcome payment worth AUD$825. Providers also receive payment
for clients who achieve economic outcomes which include transition to employment assistance (Job Network or disability employment assistance services), employment, study, vocational education and training, or apprenticeships and traineeships.\textsuperscript{20}

The human capital development approach finds many advocates in the US where a job paying the minimum wage may result in families still living below the poverty line. This underlines the importance of the context in which such an approach operates. Where the minimum wage will allow a person to live above the poverty line, then employment may be the most important means to lift people out of poverty. Where this is not the case, work first approaches may result in a decline in the standard of living of welfare recipients.\textsuperscript{21} It also speaks to the fundamental reasoning behind welfare to work programmes. Where the primary reason for moving people into work is to generate savings, then the conditions of employment are less important. Where employment programmes are used to combat poverty, net earnings and other benefits are clearly very significant factors in achieving that goal.

A mixed approach
A third approach combines elements of the work first and human capital development approaches. In a mixed programme operating in Portland, Oregon in the US, caseworkers assessed participants’ skills and needs and assigned them to the activity most appropriate to their situation. This programme “offered education or training classes to a substantial minority of its enrollees and encouraged everyone to hold out for a job that paid more than the minimum wage and offered a good chance of stable employment.”\textsuperscript{22}

Effectiveness
The National Evaluation of Welfare-to-Work Strategies (NEWWS) was conducted in the US to answer the question: “What programme strategies work best, and for whom?” The study examined the long-term (five year) effects of 11 different welfare to work programmes on single-parent welfare recipients and their children.\textsuperscript{23}

In three sites which operated both work first and human capital development approaches, side-by-side comparisons of the two approaches showed that differences in their five-year impacts on employment, earnings, months on welfare, and welfare payments were not statistically significant. That is, it was not possible to ascertain that the differences in impacts did not occur by chance. Where there
were differences, these tended to be for certain sub-groups or outcomes. The human capital development approach did not, as hoped, increase the quality or duration of employment, and programmes based on that approach were more expensive to operate than those based on the work first approach. Moreover, human capital development did not appear to be more effective than work first for the most disadvantaged jobseekers. The evaluation shows that in fact, “though neither the employment-focused approach nor the education-focused approach was clearly more effective for the most disadvantaged recipients, the employment-focused programs had slightly larger earnings impacts for this subgroup than did the education-focused programs.”

The success of the Portland programme, which adopted a mixed approach, suggests key features of very effective programmes include a focus on employment, the use of both job search and short-term education or training, and an emphasis on waiting for a good job (the Portland programme advised clients to wait for a job paying at least 25% above the minimum wage and offering a good chance for stable employment). As the state had a relatively high minimum wage, it would appear that waiting for a better paid job does not only matter where the minimum wage is very low. One might also speculate that one of the reasons for the programme’s success was that activities were more personalised to individual circumstances and needs, and this may be as important as its emphasis on quality employment. To the author’s knowledge, this aspect of the programme has not been studied. It should be noted that relative to the other NEWWS programmes, the one in Portland worked with a less disadvantaged welfare caseload, and since there was only one programme of its kind, the results should be treated with some caution.

Implications for payment-by-outcome

There would appear to be an inherent tension between pursuing a work first policy for those jobseekers for whom immediate employment is the best option, and encouraging providers to work with those for whom entry into employment will take some time and other services that can address personal (rather than vocational) barriers to work are required. It may be that these two objectives cannot be incentivised under the same scheme.

If all jobseekers are served by the same programme, and that programme measures only employment outcomes, the risk of providers parking disadvantaged jobseekers who have complex, non-vocational barriers to work may be high.
Therefore commissioners may need to create multiple programmes, each incentivising different objectives reflecting their target clientele. For jobseekers closest to the labour market, measures of entry into the workforce and retention of employment are appropriate. However, for more disadvantaged jobseekers, for whom employment outcomes are unlikely in the short to medium term, a separate programme that incentivises providers to improve education and skills outcomes may well be necessary.

Measuring only employment outcomes also narrows the scope of interventions providers are likely to offer. Providers may feel obliged to adopt a work first approach which may not successfully address the needs of very disadvantaged clients. Where education and skills outcomes are rewarded, providers are much more likely to invest in developing the human capital of their clients. Therefore it may be appropriate to offer one or more programmes and refer jobseekers to the one which is likely to better serve their needs.

It may also be that the work first theory has dominated in recent years precisely because it fits well with a payment-by-outcome approach. Employment outcomes are relatively easy to measure; commissioners and providers feel comfortable contracting on the basis of job placements and 13- and 26-week employment outcomes. Human capital development, on the other hand, is much more difficult to measure objectively, so providers and commissioners may feel less comfortable contracting on that basis. While evidence suggests that the work first approach is best for most jobseekers, when selecting primary outcomes and measures, commissioners should ensure that all jobseekers can be well served under the system they create.
Selecting Primary Outcomes

Before a payment-by-outcome system can be implemented, commissioners must understand the outcome or outcomes sought from a particular programme, and, if there are multiple outcomes, which one is the most important to policymakers. This is essential to creating a coherent and functional incentive structure. In order to prioritise outcomes, commissioners need to understand the objectives of the department or departments funding the programme.

Rationales for work activation programmes
The two main rationales governments give for funding welfare to work programmes are to reduce poverty and promote social inclusion, and to save money. Where there is a decent minimum wage, obtaining employment is often the quickest and most sustainable way for people to raise themselves out of poverty. In some cases, welfare programmes have a specific aim to reduce child poverty through helping parents, and in particular single mothers, find employment. Social inclusion can also be achieved through welfare to work programmes. Unemployed people often feel excluded from society and can suffer from poor mental and even physical health as a result. Ensuring that those who want to work are able to find suitable employment is an important aspect of combating social exclusion. By reducing the number of people on benefits and increasing tax revenue, governments also hope to save money.

Programme funding
To help them understand which of the possible programme objectives is the most important, commissioners will need to identify the department or departments that will fund the programme. Although there is some variation across countries with
regard to the departments responsible for benefits payments and employment programmes, in general responsibility for both tends to be vested in one department.

In the UK, the Department for Work and Pensions (DWP) was a merger of the Department of Social Security (DSS) and the Department for Education and Employment (DEE). In the past, the former was responsible for benefits payments while the latter was in charge of employment and training programmes. This departmental division meant that benefits savings created by the DEE by helping people into work accrued to the DSS. In 2001, the departments were merged as part of government plans to help people who wanted to work find and keep a job, so that the savings from the programmes would accrue to the department delivering those savings.31

While one department tends to be in charge of funding employment programmes, the level of government vested with these responsibilities differs from country to country. In the UK and Australia, a national government department discharges these duties. In the US, the federal Department of Labor coordinates overall policy, but the states are responsible for implementation. They have some discretion over how they interpret policy and therefore each state has a slightly different way of delivering federal directives.

Although programme funding tends to come from one department, other departments may have a stake in the outcomes being pursued. For example, in the UK and perhaps elsewhere, the Treasury may be an implicit stakeholder, often influencing the emphasis placed on cost savings. The Department of Health may also have a stake in certain welfare to work schemes and, as noted above, has acted as a co-funder of part of the Pathways programme. However, the primary focus has still been on the DWP’s primary goal, increasing employment.

This may change. Given the fiscal situation in the UK, it is to be expected that Treasury will exert more pressure on departments to extract cost savings. Indeed, this is probably why the attachment fees for the Work Programme (paid when providers accept jobseekers onto their programmes) are quite small and will only be paid for the first three years of the contract.32 Thus, while DWP remains the funder and primary decision-maker with regards to the outcome selected, Treasury may exercise substantial influence over the payment structure used to incentivise the outcome.
Multiple departmental goals

“Many have attributed the difficulty [of identifying performance measures] to the complex or ill-defined nature of goals in the public sector.”33

Allocating responsibility to one department does not preclude the possibility of programmes having multiple objectives since departments generally have several reasons for funding a programme. Sometimes the goals seem to form a coherent set, such as welfare to work programmes which aim both to reduce benefits dependency in order to promote poverty reduction and social inclusion, and to reduce public spending on welfare. By achieving the former, the latter is likely to follow.

Other times goals are more obviously contradictory. For example, the criminal justice system has at least two aims: to punish and to rehabilitate criminals. Judges, who have the responsibility for sentencing, are often required to choose between using incarceration, which may achieve the punishment objective but may not contribute to the rehabilitation process, and using community sentences, which are more likely to achieve the rehabilitation goal.

It seems obvious that when goals are contradictory, departments will need to prioritise their goals before creating an incentive structure. However, what is less intuitive is that even where goals appear coherent, an incentive structure may create tensions among them. This is because it is very difficult to create an incentive system where all goals are treated equally; often, one goal is (intentionally or accidentally) elevated above the others. This will be discussed further in section 7. For now, it is sufficient to note that honesty about the relative importance of various goals is essential to creating a functional incentive structure.
Segmenting the Population

One of the primary challenges commissioners face in designing a payment-by-outcome system lies in ensuring the system incentivises providers to meet the needs of the entire unemployed population. Jobseekers face different types of barriers to work and some are further from the labour market than others. Although the primary outcome for all jobseekers may be finding sustainable employment, the measures used to assess progress may depend on how far a group is from the labour market and providers may need to be paid more to help highly disadvantaged jobseekers who will cost more to get into work. In addition, differences in the profiles of clients providers serve may make it difficult to compare their performance, so commissioners may find it necessary to segment the population into groups in order to make payment-by-outcome feasible.

Setting prices

To reduce incentives to cream and park, commissioners may pay providers more for securing outcomes for highly disadvantaged jobseekers. This demands a reliable and trusted means of classification. Under Job Services Australia, the Job Seeker Classification Instrument is used to categorise clients into three different groups, each attracting a different level of payment. Providers have the option of petitioning to reclassify jobseekers they feel have been miscategorised. The instrument is only accurate enough to classify jobseekers into broad categories, so that the groups are still quite heterogeneous. While imperfect, it is likely this segmentation does reduce the incidence of parking.

Commissioners may also choose to establish the proportion of jobseekers that fall into each price category and increase the level of payment to providers as more clients are served. For example, commissioners might not pay anything for
the first 10% of employment outcomes, on the assumption that these jobseekers would have found work without intervention, £600 per outcome for the next 15% of jobseekers, £1,000 for the following 15% and so on. A target accelerator model such as this would eliminate the risk of misclassifying individual jobseekers, but it would still rely on commissioners correctly identifying the proportion of jobseekers in each category and the approximate cost of getting them into work.

In its Work Programme Prospectus released in November 2010, the government stated its preference for using classification rather than a target accelerator, since it feared the latter would encourage providers to help the ‘low hanging fruit’ first. This concern appears to be linked to an aversion to paying for deadweight (jobseekers who would have found work without external intervention). The concept of deadweight can be disputed, however. Some jobseekers may find work by themselves eventually, but with assistance would do so more quickly, generating some savings. Therefore, without endorsing the accelerator funding model, it can be argued that there are strong public benefits to assisting those closest to the labour market first, as long as the government does not pay an unacceptably high price.

Measuring impact
Because the population of jobseekers is diverse, it is likely that providers will serve different proportions of easy- and difficult-to-help clients. This makes comparing the performance of different providers, a common way of distinguishing service impact from external variables, very difficult.

Measuring raw outputs – the numbers of jobseekers retaining employment for 13 and 26 weeks – is relatively simple. However, in order to measure the impact of providers on those outcomes, it is necessary to control for external factors that may have impacted on outcomes. This can be achieved in various ways. In general, commissioners of welfare to work programmes have not been able to compare programme results with those of a control group. An alternative way of attributing impact lies in using a statistically derived comparator, based on the results that commissioners would expect providers to achieve taking into account jobseekers’ characteristics and labour market conditions. However, this relies on the statistical model being robust.

One common way of attributing impact is therefore to compare the results of two or more providers operating in the same conditions, referred to as yardstick competition. Providers operate in the same labour market and serve similar
populations of jobseekers, achieved through referring similar proportions of easy- and difficult-to-help jobseekers. For example, in the UK, Employment Zone outcomes could be compared to those of the New Deal because both programmes faced similar circumstances in terms of the clients they served and labour market conditions in which they operated.\textsuperscript{38} Yardstick competition reveals information about the impact of external variables since if two or more providers are underperforming in relation to expectations, there is a \textit{prima facie} case that difficult economic conditions have impacted on outcome achievement. Alternatively a large discrepancy between the results of different providers is an indication of underperformance by one.

Collusion, often raised as a problem with this model, does not seem to have been an issue in the UK, where several Employment Zones and Flexible New Deal districts operated with yardstick competition. This is possibly due to the shadow of the future and the importance providers place on protecting their brands.

Another challenge lies in classifying jobseekers accurately so that providers serve similar client mixes. The tool used to classify jobseekers must be robust, taking into account all the variables that may impact on an individual’s ability to obtain and retain employment. Alternatively, providers can serve different types of jobseekers and commissioners can adjust for this statistically before making comparisons. However, statistically controlling for client mix still requires agreement on the characteristics that make jobseekers more difficult to help. Australia makes no attempt to ensure providers serve similar mixes of clients, but rather uses information about the jobseekers served over a certain period to make statistical adjustments to the number of outcomes achieved. This forms the basis for ‘star ratings’, the principal method of comparing providers.\textsuperscript{39}
Selecting Measures

“Few organizations’ stated objectives are sufficiently verifiable to be used directly in incentive contracts and reward systems.”

Selecting good measures of progress toward primary outcomes demands more than simply tailoring the measures to the different segments of the population. First, a good measure should reflect, as closely as possible, the primary outcome sought. In some cases, commissioners may need to incentivise more than one intermediate outcome to ensure that all aspects of the primary outcome are achieved. If a programme has multiple primary outcome goals then the likelihood of requiring several intermediate outcomes to ensure their achievement increases. As further explained in the next section, financial incentives change the focus of providers’ efforts and failing to incentivise a broad enough range of intermediate outcomes could jeopardise the achievement of primary outcomes. Clearly, having too many measures is not desirable either, since it increases the costs of verification and can restrict the freedom of providers to innovate and be flexible. Choosing the correct range and number of measures is therefore a balancing act.

Second, intermediate outcomes should be simple and clearly defined so that they can be easily and cheaply measured and verified. Intermediate outcomes which are expensive to monitor can drive up cost, since providers will build the cost of demonstrating achievement into their bids. Measures should be independently observable to ensure they will be credible to all parties.

Many different measures have been used in welfare to work over the years. The earliest performance measurement framework in welfare to work about which the author has information dates to 1948. A US state employment agency began to measure its employees based principally on the number of jobseekers interviewed
by employees, the number of clients referred to jobs and the number of jobseekers hired. A study of this measurement system explains, “The first index… impelled them to work faster… The next two indices… curbed the tendency to attain speed by neglecting the very objective of employment interviewing, namely, finding jobs for clients.” Even this early attempt at performance management recognised the danger of measuring only one aspect of performance.

Measures have changed over time through learning and changing priorities. Since this early example of performance management, commissioners appear to have recognised that simply measuring the number of clients hired would have the same effect on employees and would reduce measurement costs. The first JTPA measures thus included a measure of the number of clients in jobs on the day their participation in the programme ended.

However, commissioners then began to worry about the problem of churn; there were many jobseekers who found employment, lost it, and reclaimed benefits. Thus the focus shifted to developing measures of sustainable employment, hence programmes currently operating in the UK and Australia measure 13- and 26-week employment outcomes. These intermediate outcomes were chosen because sustainable employment cannot be measured quickly enough to provide good management information. Commissioners cannot forego having information on providers’ performance for very long periods, as this would reduce providers’ accountability and could result in political embarrassment if outcomes were discovered to have been poor. The 13- and 26-week employment measures are a compromise between measuring employment for a number of years to ensure it is truly sustainable and gathering performance information in a timely manner.

All of the measures discussed here are simple, clearly defined, easily measurable and verifiable. However, some studies raise questions about whether the 13- and 26-week employment outcomes reflect closely enough the primary sustainable employment outcome. Studies of the long-run effectiveness of programmes reveal a wide range of results. It is not clear that a person who retains work for 26 weeks will be better off five years later than a person who did not find work while on the programme or who only retained work for 13 weeks. For example, JTPA’s measures of employment rates and earnings levels 13 weeks following discharge from the programme were weakly and in some cases negatively correlated with longer-term employment and earnings impacts estimated for participants using experimental data. This is not necessarily an argument for extending the timeframe over
which providers are measured, which would delay the collection of performance information and likely increase cost. Moreover, the 13- and 26-week measures remain useful as measures of cost savings to government. These may be ‘good enough’ measures for government even if they are not strongly correlated with long-term outcomes.
Designing Incentive Systems

Financial incentives
There is ample evidence from welfare to work programmes in several countries that private sector organisations paid for delivering certain results increase their focus on the activities which will deliver those results, sometimes to the detriment of other activities which are still important to the overall missions of the programmes. In some ways, this does not come as a surprise. In general, it is expected that organisations driven by the profit motive will respond to financial incentives.

However, government agencies are also motivated by financial incentives. A welfare to work programme in Wisconsin was encountering opposition by employees that worked for the counties’ welfare departments. In order to stem this, the counties that wanted to compete for the new programme contracts could earn a right of first refusal if they “met certain performance requirements in the twelve months prior to contract selection.” As a result, by the time contracts were due to be selected, “all but about half a dozen counties had succeeded in meeting the requirements”. As one report notes, “The right of first selection process showed that in certain circumstances government entities are motivated by financial incentive structures as well as private vendors are”.46

There is also evidence that voluntary organisations respond to financial incentives. In Australia, the Salvation Army, a Job Network provider, was obliged to repay AUD$9.4 million following an audit which revealed irregularities in some of its employees’ job outcome claims.47

As this latter example shows, precisely because all types of organisations respond to financial rewards and penalties, commissioners must carefully structure incentive systems to ensure the primary outcome is achieved. Well designed incentives will encourage providers to achieve the objectives of policymakers,
motivate providers to focus the correct amount of resources on each intermediate outcome (where multiple outcomes are incentivised), and reduce the incidence of gaming.

**Scale of risk**

As discussed in section 4, policymakers and commissioners must clarify what they would like to achieve by using payment-by-outcome. This will help commissioners determine the optimal level of risk to transfer to providers. Outcome payments can form a very large proportion of total payment (revenues are at risk), a low proportion (less than the anticipated profit margin is at risk) or somewhere in between.

The level at which risk is set will in part determine the level of innovation in service delivery commissioners can expect. This is because where providers could lose substantial sums of money, they may become risk-averse and restrict themselves to using well-known and tested models of service delivery. Pathways to Work may be one example of where this has occurred. Where a large proportion of revenue is guaranteed, providers may not be incentivised to deliver outcomes. If increasing the level of innovation is an objective, a balance might be struck in which a sum equivalent to, for example, twice the profit margin is at risk, to ensure providers are sufficiently motivated to deliver outcomes but are also willing to take the risk of innovating.

For commissioners wishing to obtain alternative sources of financing for programmes, an invest-to-save model in which no payments are made until savings from the achievement of outcomes are delivered will be preferable. Such a model may not deliver much innovation; given the considerable risks providers would be required to carry, they may opt for more familiar delivery models which are easier to price.

**Establishing priorities**

Where commissioners have multiple (non-contradictory) objectives, they will need to decide whether these are equally important or one has priority. In cases where all goals are of equal importance, commissioners will need to create a ‘neutral’ incentive structure. In order to achieve this, the amount of incentive payment must be ‘relatively equal’; all providers must be able to make an equal return from each outcome achieved, over and above the effort required and transaction costs associated with claiming outcome payments. If payments are the same,
then providers may focus on the task that requires the least effort or is the easiest to measure. The principal reason why ‘neutral’ incentive structures are difficult to design is that providers have more detailed knowledge of the costs of various elements of service delivery and the costs associated with demonstrating they have achieved outcomes, so it is highly likely that commissioners will inadvertently incentivise one outcome more than another.

Of course, it is entirely possible that commissioners will want to incentivise one goal over another. In public services, this can be very difficult, since agencies tend to acquire contextual goals as a result of ongoing negotiation with stakeholders. Overtly placing more value on one of these goals can be politically controversial. However, when designing incentive structures, it is important to be honest about what matters most. The risk of attempting to appear ‘goal neutral’ is that the incentive structure that is eventually implemented will not encourage the achievement of the most valued outcome.

There are a number of ways in which incentive structures can be designed to prioritise goals. First, payment can be weighted to reflect importance. Again, commissioners must take into account the costs of achieving the outcome and claiming payment. Second, commissioners can build in an element of contingency. In some cases this will occur naturally, because one outcome cannot be achieved without achieving another. For example, in many welfare to work programmes, providers can earn payments for clients who achieve 13 and 26 weeks in work. The 26-week outcome payment is inherently contingent on providers achieving the 13-week payment. In Australia, the importance of the 13-week outcome is additionally reinforced by attracting a larger payment than the 26-week outcome. However, contingency can also be artificially created through the design of the incentive structure. This has occurred in the health sector in the US, where providers of long-term condition management programmes only became eligible for payments based on quality once they had achieved a certain threshold of cost savings.48

**Limiting gaming**

Problematically, outcome measures are only able to capture some dimensions of the primary outcome, which means that providers may have an incentive to increase their measured performance rather than their actual performance in relation to the ultimate outcome. For example, 13- and 26-week employment measures do not capture the totality of what commissioners want providers to achieve. They
may not, for example, specify which kinds of jobseekers commissioners want providers to prioritise or give a detailed indication of the quality of employment providers must help jobseekers find. These omissions create opportunities for providers to game, for example by focusing on jobseekers with few barriers to work (creaming), providing minimal services to very disadvantaged jobseekers (parking) or by encouraging jobseekers to accept jobs with little opportunity for advancement (shading service quality). Commissioners have experimented with scaled payments, retrospective adjustment and prizes to limit the perverse incentives created by paying for intermediate outcomes.

**Scaled payments**

Increasing the payment providers receive for helping highly disadvantaged jobseekers into work should decrease the incidence of parking. Under Australia’s Job Network, jobseekers were categorised according to their distance from the labour market so that providers earned more for achieving outcomes for disadvantaged jobseekers. Since 2006, experts in the UK have been advocating a form of ‘target accelerator’ in which commissioners pay progressively more as providers return more jobseekers to work, but such a system has yet to be implemented.

Most systems that group jobseekers in order to assign them different tariffs create thresholds, which may create further gaming opportunities. If commissioners understood perfectly all the qualities that made jobseekers more difficult to help and could assess these with perfect accuracy, providers would be unable to contest how particular jobseekers had been classified. However, since commissioners recognise that their classification tools are imperfect, they often allow providers to dispute jobseekers’ classifications. This creates the possibility of providers gaming the system by inappropriately requesting that jobseekers be re-classified into groups with higher tariffs. Moreover, even within relatively homogeneous groups, those jobseekers that cost less to help than others are likely to receive disproportionate attention from providers. Those on higher tariffs who are relatively easy to help may even be prioritised over jobseekers closer to the labour market on lower tariffs, depending on the scale of the difference in outcome payments between the two groups.

While it is clear that designing a continuous payment system where each individual jobseeker attracts a different level of payment is not feasible, great care should be taken in establishing the thresholds for scaled payments so that, as much as possible, perverse incentives are minimised.
Retrospective adjustment

Retrospective adjustment is a way of taking into account the impact of labour market conditions as well as jobseeker characteristics. Rather than estimating in advance the proportions of different types of jobseekers providers will serve and the likely labour market conditions in which they will operate, retrospective adjustment is based on actual clients served and conditions experienced. As the adjustment must be made at the end of a certain period, commissioners could either pay providers an expected rate and then request a refund or arrange additional payment to providers, or wait until the end of a measurement period to make payments. To the author’s knowledge, no payment-by-outcome systems are currently using retrospective adjustment, although programmes such as the JTPA have adjusted based on predictions of expected performance.

Prizes

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 target were not disclosed.\(^{52}\) This ensured that providers had an incentive to find employment for as many jobseekers as possible, rather than simply finding work for those jobseekers whom it was profitable to place based on the amount of the outcome payments. The Work Programme also uses prizes, called ‘incentive payments’, but these may not influence provider behaviour in quite the same way since the level of performance required to receive a prize has already been disclosed.\(^{53}\)

Other incentives

Reputational, regulatory and governance incentives may also influence provider conduct. Where substantial risk is transferred to providers, financial incentives will tend to be an important determinant of behaviour, but even in these cases there may be a role for other types of incentives.

Reputation

Reputational incentives work where firms see long-term commercial value in being recognised as high-achieving organisations and are therefore motivated to meet commissioners’ expectations. ‘Star ratings’, for example, are a very important
reputational incentive where providers believe the probability of future business is high and the ratings have an impact on winning that future business. In Australia, such ratings are a particularly important part of the incentive structure. They have been used since 1998, and providers know that performing well could qualify them for an automatic contract renewal or extension. Since June 2008, the UK’s Department for Work and Pensions has also awarded star ratings to its contracted service providers. The ratings “support the Department in making informed decisions about contract letting, extension and management.”

Where a good reputation is unlikely to influence decisions about the allocation of future business, providers may be more inclined to maximise their short-term financial positions by engaging in gaming behaviour. Moreover, where providers carry a high level of financial risk, concern about achieving outcome payments may well outweigh reputational concerns. For example, providers of Pathways to Work programmes have been forced to reduce the provision of services to a minimum, causing harm to their reputations, because of the problems they face in financing these programmes.

**Market governance**

Regulation and regulators provide a legal incentive for providers to comply. For example, in the UK, all providers of employment services must comply with the Employment Agencies Act of 1973, which includes provisions relating to the proper conduct of employment agencies, how they may be inspected and what constitutes fraud.

Often, as markets grow, the number of contracts proliferates and services become more standardised, what began as specific contractual obligations can become enshrined in contractual governance instruments. Indeed, the Australian Employment Services Code of Practice may be an example of the codification of best practice. All organisations contracted to deliver government-funded employment services have agreed to observe the Code, which “sets out the principles and standards that underpin the delivery of employment services and other services to increase employment outcomes and participation in economic activities in Australia especially for disadvantaged client groups.” The Code’s explicit reference to disadvantaged jobseekers seems to be another attempt by the Department to reduce the incidence of cream-skimming. In the UK, the Department for Work and Pensions Supplier Charter sets out the expectations for all organisations supplying
the DWP, while the Code of Conduct governs the relationships between prime and subcontractors and the Merlin Standard helps ensure adherence to the Code.\textsuperscript{56}

**Institutional governance**

Studies of contracting often treat the provider as one actor and the customer as the other. However, the internal dynamics of providers and the way they are governed can impact on the quality of the service delivered and the way employees respond to incentives in the contract, so commissioners may wish to be aware of different providers’ governance mechanisms.
The way in which services are procured can have a large and lasting impact on the quality of services and the outcomes they achieve. There are some differences between ‘traditional’ procurement and payment-by-outcome which makes the procurement process for the latter more complicated. Since payment-by-outcome involves the specification of an intermediate outcome or in some cases one or more outputs, potential providers may propose quite different service models with varying costs. In this context, commissioners required to assess the quality of the bids would have to use their discretion, drawing on available evidence and past experience, rather than comparing bids to a standard quality specification. Moreover, commissioners would have to exercise judgment in comparing the prices of different bids, since these may refer to very different types of services. Price-based competitions may not reveal the best value for money. This begs the question: what kind of procurement process is best suited to payment-by-outcome?

A scan of the methods used to commission for outcomes reveals substantial variation. Commissioners have made different choices about the extent of dialogue between commissioner and providers, the weight given to price and quality, and the value placed on past performance. Some of the decisions have unquestionably been factors in subsequent poor service provision, while the impact of others is less obvious. No process has been without flaws, but important lessons can be learned by examining why some have worked better than others.

Pathways to Work
The procurement process for Pathways to Work in the UK is widely recognised as having contributed to a winner’s curse. First, providers believed they understood the claimant group for which the programme was being procured, but there were
in fact substantial differences between the assumptions made by providers and the characteristics of jobseekers they actually served. Providers submitted bids on the assumption that a similar service would be required to that supplied under the voluntary New Deal for Disabled People. However, jobseekers on the mandatory Pathways programme faced complex barriers to work and were far less job-ready than providers had anticipated. Thus the cost to providers was higher than expected. The problems associated with the misunderstanding of the population to be served were “exacerbated by the highly competitive bidding process”.

Pathways bids were assessed based on the numbers of outcomes providers pledged to achieve. Since the value of the contract was fixed, this led to a competition based on the unit price per outcome. As providers promised to deliver more outcomes, they became more likely to win a contract but also lowered their unit price per outcome.

Partly because they knew so little about the population and partly because the process was so competitive, many providers were overly optimistic which resulted in low unit prices. Subsequently, they were unable to achieve the numbers of outcomes to which they had committed in their contracts, and thus faced severe cash flow problems. In response, providers had to pare down their services and performance expectations were revised. Although the recession and other issues surrounding the implementation of the new Employment and Support Allowance and the Pathways programme compounded these problems, it is widely recognised “that the vast majority of Pathways targets were never achievable.”

Several lessons can be drawn from this example. One is that early assumptions about the nature of new programmes can be unreliable. New services may be similar to existing services in many ways, but it is possible that there will be unforeseeable differences. In this situation a procurement process that puts a heavy emphasis on price is probably not appropriate. Rather, a process that encourages providers to be innovative about their service offerings and treats price as a secondary consideration may deliver better outcomes. As providers and commissioners develop a better understanding of the service offering and its cost, then more emphasis can be placed on price. This leads to a lesson regarding the dynamic nature of payment-by outcome. Learning comes from experience, and tools to capture learning should be integrated into the procurement process. Commissioners should make changes to the system as necessary based on the lessons learned.
Working Nation and the Job Network

Australia is an interesting case because the way in which welfare to work services have been procured has changed substantially over the years. For the Working Nation programme (1994), the Department of Employment and Workplace Relations (now the Department of Education, Employment and Workplace Relations), which commissioned the programme, opted to run a competition based on quality since the market was new and providers inexperienced. By fixing the price, rather than allowing it to emerge through the procurement process, the commissioner assumed the risk of either having set it too low and service quality suffering or of having set it too high and providers making large profits. When the next government scrapped the programme, it argued Working Nation had been expensive and ineffective at obtaining permanent employment for jobseekers. It is possible that the high programme cost was partly a result of the government setting the fixed price too high and the considerable profits made by some providers seem to confirm this.

The first Job Network contract tender process was tightly controlled. Three-person panels in each local region considered bids against the tender specification. These panels were not permitted to exchange information, so each region functioned, in effect, as a separate mini-tender. The true reason for adopting this process is unknown, but one might speculate that the Department wanted to encourage some innovation on the commissioning side. By allowing the panels to use slightly different methods to select winning bids, the Department could be fairly certain that the provider market would be diverse.

Providers could compete on price for ‘job matching’ and ‘job search training’, which were fairly standard services. Bids were assessed for quality and then ranked by price. The Department was not required to accept the lowest bid for any service, but could instead trade off aspects of quality and price. However, prices for ‘intensive assistance’ were fixed and providers competed solely on quality for this particular service. This was due to “concerns that the bidders would initially lack the expertise to cost the new service.”

Despite what were probably sensible decisions about the balance between quality and price, there were several weaknesses in the first Job Network tender process. Panel members were not experienced service providers, so the qualitative judgments about the services required by the process of weighing quality and price had to be defended with reference to the quality of the bid. Although previous provider performance could be taken into account, “informal forms of political
“steering” occurred in which panel members received indirect messages not to give too much weight to past performance. The final result was that many agencies that had performed well under Working Nation were not offered contracts, and many new agencies were offered contracts despite having “no track record, no facilities and little obvious claim on expertise.”

Another potential weakness was that negotiation with bidders was not permitted. This resulted in some bidders being offered business at lower rates than what they had submitted in their tenders. However, because they could not negotiate with the Department, they simply had to either accept or decline. If they declined, there was little likelihood they would receive a second offer. There does not seem to be any evidence that this resulted in providers offering poor quality services; however, this could have created a situation in which providers that did not believe they could deliver the service for the offered price accepted anyway in order to secure their place in the market.

In the second round of Job Network contracts, some competition on price for intensive assistance services was permitted. Outcome payments for 26 weeks in work or two semesters in education were fixed, but the service fee and payment for a 13-week job outcome were set by competitive tender. The Department set different minimum prices for each category of jobseeker, but providers could bid above the minimum price if they thought the conditions in which they operated justified it. The bids were assessed 75% based on quality and 25% based on price.

It would seem that in this contract round, a slightly higher level of expertise on the part of both providers and commissioners was assumed. However, commissioners were not sufficiently confident that certain elements of service quality would not suffer if price competition were allowed, so some of the outcome payments were fixed or a floor price put in place. This system was clearly “very much dependent upon the skill of the [Department] to set a reasonable administrative [fixed] price.”

Thus there was some price competition, heavily controlled by the Department, but in practice the floor price set the price, as bidders bid down to the floor: “the difference between the average and minimum price for upfront payments for level B clients in [intensive assistance] was less than 7 per cent, compared with nearly 90 per cent for [job search training] (where no floor price was set).” Moreover, those providers that did bid above the floor and won contracts were not necessarily the best at achieving better outcomes because the Department had to try to evaluate this based on the service proposition in the bid, and they were not always right.
Therefore, the Australian Productivity Commission concluded that simply setting the price administratively would have given a similar result.\textsuperscript{70}

The reason given by the Department for its decision to set a floor price was “to ease the transition to a fully competitive market for Intensive Assistance, and as a safeguard to protect service quality and reduce the risk of market failure”.\textsuperscript{71} However, in the event, competition on price for the third Job Network contract was abandoned because of concerns about gaming. Instead, 60\% of business was reserved for existing providers supplying a certain standard of service, as identified by providers’ levels of ‘star rating’. The other 40\% was put out to open, fixed-price competition.\textsuperscript{72} This change was likely the result of a report by the Productivity Commission, published in 2002, which exposed the extent of parking of the hardest-to-help jobseekers. Fixing the price was the Department’s attempt to soften the incentive to park by ensuring a sufficiently high level of compensation for helping the most disadvantaged.

In its evaluation of the procurement processes for the first two Job Network contracts, the Productivity Commission argued in favour of fixed pricing because (1) competition based on price was likely to lead to a lower quality service than what the Department desired and (2) there was little evidence that competition with a floor price succeeded in awarding contracts with higher prices to those providers that could best achieve higher marginal net outcomes.
10
Managing the System Over Time

Creating an effective and efficient payment-by-outcome system is an iterative process, not an event. Commissioners may wish to pilot new approaches before rolling them out more widely. Even when programmes have reached scale, commissioners will need to continue to gather evidence, conduct reviews and be prepared to make major changes to improve outcomes.

Data and information
Commissioners will need to collect data in order to manage payment-by-outcome systems well. In welfare to work, data on jobseekers and labour market conditions is especially important. Such data ensures commissioners can adjust figures to compare the performance of various providers. Since 1998, Australian jobseekers have been classified according to the Job Seeker Classification Instrument (JSCI). The JSCI gathers information about jobseekers’ characteristics and histories, providing commissioners with valuable information about regional differences in the populations served and how the flow of jobseekers has changed over time. In 2008, a review of the JSCI was conducted amid concerns about its ability to reveal certain characteristics of jobseekers that might make them more disadvantaged than they first appeared. For example, providers raised concerns that employment histories were incomplete or did not go back far enough to give an accurate picture of relative labour market disadvantage. The new JSCI will be able to examine jobseekers’ receipt of a variety of different kinds of benefits for the previous ten years, in order better to assess their accumulated time in unemployment.\textsuperscript{73}
It is equally important that commissioners are able to identify best practice in employment services. Austria has a particularly advanced ‘Datawarehouse’ “which enables detailed longitudinal analysis of the experiences of different customer groups in the labour market and how this is influenced by [the public employment service’s] interventions.” The Datawarehouse integrates data from a number of government databases, including the public employment service and the social security system which tracks transitions in and out of employment. Data date back to the mid-1990s and can be analysed by individual or by group.

Information about the cost of achieving outcomes for different categories of jobseeker is also crucial to ensuring commissioners obtain value for money from employment programmes. Open book accounting, possibly with profit sharing agreements, would help ensure value for money.

**Monitoring**

Since fewer processes are specified in the contract and providers have much more discretion with regards to the service they deliver, monitoring becomes especially significant in payment-by-outcome. The role of monitoring must then be to verify that the services that are being delivered not only achieve the outcomes specified but also fulfil other requirements to do with the quality of service, equality of treatment of different service users, and other difficult-to-measure elements of service delivery. Monitoring involves costs; therefore, commissioners must decide what level and kind of monitoring is optimal.

**Contract compliance**

A relatively simple form of monitoring lies in scrutinising providers’ achievement of outcomes against what is stipulated in the contract. Since providers must prove that they have achieved outcomes in order to receive payment, the monitor must compare the numbers of outcomes achieved with the contractual obligation. Where providers are failing to meet their outcome targets, commissioners must take action to remedy the situation. In general, poor performers are given a period of time to improve their attainment levels. If they are still performing below par, then commissioners may need to replace them. If the failing provider is a public body, management may need to be replaced.

Contracts may also specify processes which providers must perform. For example, some welfare to work contracts require providers to develop, sign and
implement ‘Return to Work’ plans for jobseekers. This is relatively straightforward to monitor, since inspectors can examine a proportion of each provider’s files to ensure that this has been done. Where there are omissions in the sample inspected, a more thorough inspection may be required. Other processes that do not have an immediately visible or tangible result, such as job advisors explaining jobseekers’ obligations and the consequences of non-compliance, will need to be monitored through site visits.

Programmes monitor contractual compliance in various ways. In New York, one contract manager from the Human Resources Administration (HRA) oversees eight providers. This individual has a very close relationship with the providers so when problems arise, they are identified early. Moreover, providers are likely to be open with the HRA contract manager about emerging problems, and the contract manager may even help providers resolve issues. This ‘early warning system’ has been highly effective and compensates for the fact that poor performance may otherwise go unnoticed until an outcome target is missed. However, there are concerns that this type of close relationship with incumbent providers stifles competition since the HRA contract manager effectively acts as a support that is not available to potential new entrants. New York also facilitates monthly meetings between commissioners and providers, which provide an opportunity for both parties to discuss difficulties and review performance statistics.76

Under Pathways to Work in the UK, ‘contract managers’ were responsible for monitoring providers’ outcome achievements against the profiles agreed in their contracts.77 In addition, they monitored compliance with other contractual obligations such as delivering five work-focused interviews to each client. ‘Third party provision managers’, on the other hand, were responsible for monitoring the quality of service provision throughout clients’ journeys. The separation of these two responsibilities led to confusion about roles and requirements for access to data, which decreased the overall effectiveness of the monitoring system.78

Under the Flexible New Deal in the UK, prime providers played a major role in monitoring the performance of their supply chains, including compliance with contracts and less tangible elements of service quality, as they remained responsible for jobseekers’ experiences throughout the programme. Prime contractors managed this monitoring responsibility in different ways. Those that mostly delivered services in-house tended to have informal arrangements with partners, while those with a large number of sub-contractors or other partners made more formal arrangements
including staff dedicated to managing the supply chain. While the prime contractor model is attractive because it devolves the monitoring role from commissioner to prime provider, and the Department for Work and Pensions has stated that “it will not become involved in what are commercial negotiations between prime and sub-contractors”, some have called for a greater role for the Department in monitoring Work Programme sub-contractors in order to identify and disseminate best practice.

**Quality**
Commissioning services is often more difficult than procuring goods since there are many aspects related to the quality of service delivery, such as the level of courtesy staff should demonstrate to clients, that are difficult to specify in a contract. There are several ways to monitor service quality.

**Voice**
The UK government plans to enable jobseekers to voice complaints about services under the new Work Programme. As part of their bids, potential providers will be required to “provide a summary... of the minimum service they will offer to all customer groups.” Based on this, Key Performance Indicators will be developed and articulated to jobseekers so that they will be able to judge services and complain if there is a problem. Providers failing to deliver as promised could face repercussions for breaching their contracts. A potential problem with the use of the voice mechanism to ensure minimum service standards is that those jobseekers who are likely to be parked are also likely to be unmotivated and may therefore be less inclined to complain about not receiving the level of service to which they are entitled.

**Choice and exit**
In choice-based markets, customers can change providers if they are unhappy with the quality of the service. In some public service markets individuals are able to exercise choice in this way, giving commissioners information about service quality. However, as with voice, choice may not raise quality where those jobseekers most likely to be parked are not those most likely to exercise their right to exit. Moreover, in many markets for public services citizens are not free to choose. Even in Australia’s welfare to work market where clients have the right to choose a provider initially,
once they have accepted to work with a provider, they are not entitled to switch providers unless major problems emerge in the relationship. This characteristic of many markets for public services increases the importance of monitoring.

**Inspections**

Inspections, whether announced or unannounced, are a common element of most monitoring systems. Inspectors observe how providers deliver a service and produce a report which is publicly available, so there is a reputational cost to providers who perform poorly. This is particularly so where performance on such inspections is likely to affect the commissioner’s decisions about the award of subsequent contracts.

**Surveys**

Surveying customer satisfaction is another way of monitoring service quality. The advantage is that feedback comes directly from those receiving the services, rather than from an independent third party experiencing the service ‘second-hand’. However, surveys can be expensive to design, administer and interpret, and there may be an element of self-selection in the individuals choosing to take the time to participate, so the results may not be valid for the entire population. Moreover, overall perceptions of service quality may be affected by a number of different service elements, so it is best to combine survey findings with third party inspection to obtain an overall picture of the quality of provision.

Staff surveys are another option. Employees may be able to give insights into how they feel about the services they deliver, the parts of their work they enjoy and what frustrates them. This can enrich the picture of service quality, but must be combined with other measures of quality since employees may give biased responses to questions that relate to their own skills and behaviours.

**‘Star ratings’**

‘Star ratings’ are used to monitor the quality of service provision in the Australian and British welfare to work markets. They are an aggregated performance measure that helps determine subsequent contract award and in some cases automatic contract renewal. In Australia, the ratings take into account client and labour market characteristics using regression analysis, and are weighted in favour of outcomes for the most disadvantaged jobseekers.83
Conclusion

Many governments around the world have implemented welfare to work programmes which pay for outcomes, which indicates the perceived value of the approach. Payment-by-outcome is indeed a powerful way to align incentives, promote innovation, increase accountability, and ensure that service provision is flexible and personalised. Equally, however, despite almost 30 years of cumulative experience, programme design continues to evolve, illustrating the complexity of implementing payment-by-outcome in employment services. Commissioners should not expect to design a perfect system the first time. It is very difficult to anticipate providers’ responses and a certain amount of gaming is therefore inevitable. Commissioners should seek to learn from such opportunistic behaviour and adapt the system over time so that it delivers the desired outcomes.
Appendix

The United States
There is a long history of performance management in welfare to work programmes in the US. Because responsibility for the administration of welfare to work programmes is divided between the federal government and the states, there has also been a diversity of different systems in the country, many of which have been well-studied. Comparisons of the programmes provide a good starting point for an analysis of payment-by-outcome.

Federal programmes
The Job Training Partnership Act
The 1982 Job Training Partnership Act (JTPA) “was the first large-scale, federally funded program to mandate the use of performance standards in state and local programs.”\(^8^4\) The federal Department of Labor (DoL) funded the programme and determined its broad outlines, while the states were responsible for administration. The programme was operated by about 620 agencies (usually called ‘providers’ in the UK) operating in geographical monopolies.\(^8^5\) Agencies were allocated budgets based on the number of unemployed and economically-disadvantaged individuals living in their geographical areas.\(^8^6\)

The DoL stated that the programme was to be targeted at “those who can benefit from, and are most in need of, [job training] opportunities”.\(^8^7\) The DoL specified that at least 90% of the enrollees be economically disadvantaged; that recipients of Aid to Families with Dependent Children (later replaced by Temporary Assistance to Needy Families, see below) and high school dropouts receive an equitable share of JTPA services; and that 40% of all JTPA service dollars be expended for services to youth.\(^8^8\) Within these guidelines, agencies had discretion over which clients they served. Those eligible to participate included “stably employed working poor persons and persons who were out of the labor force.” In practice, only about 1% to 5% of the eligible population was served each year.\(^8^9\)
The primary outcome specified in the legislation was to maximise the “return on investment in human capital”, where human capital was defined as a “set of skills and knowledge.” Originally, four measures were used to assess agencies delivering the programme: the percentage of clients entering employment, the average wage at placement, cost per employment outcome and the percentage of welfare recipients entering employment. These measures were compared against a standard of expected performance for each agency, based on regression analyses of the health of the local labour market and the characteristics of the jobseekers referred. This comparison helped to capture the impact of the programme on the employment and wage results. In 1990, three of the measures were changed so they were taken 13 weeks after a participant had been terminated from a programme. The cost standard was eliminated because JTPA officials were concerned that it encouraged ‘quick fix’ job placement activities instead of investment in skills or training that would make a greater difference to participants in the long run.

States had substantial discretion over the administration of the programme. They could choose to add performance standards beyond the four specified by the DoL, in order to further state objectives. They could also make adjustments to agencies’ standards beyond the regression analysis of the DoL. Finally, states were in charge of the funds to award to agencies. Award money amounted to about 7% of a state’s JTPA budget. Meeting the standards was a condition for receiving an award, but states could dispense bonuses to agencies in any way they chose. In most states, agencies simply had to meet a certain standard of outcome achievement to receive an award. However, there was some variation in how funds were dispersed in different states. Some states gave the entire bonus to the best-performing agency while some divided it among all agencies performing above a certain threshold. Other states varied the award according to performance relative to the standards, so that agencies that far surpassed their targets received more than those that simply met their standards. The bonuses could not be used to reward program managers or case workers but could only augment the budget of the agency.

The Workforce Investment Act
In 2000, the JTPA was replaced by the Workforce Investment Act (WIA) and a number of important changes were made. First, while the decentralised nature of the system was maintained, competition for market share was introduced with jobseekers being given a choice of service provider.
Second, the incentive system was amended. There are 17 performance measures under the WIA. The ‘core’ measures for adult jobseekers are:

- the percentage of clients entering employment;
- the percentage of clients retaining employment for six months;
- the earnings change from six months prior to entry into the programme to six months after exit; and
- the percentage of clients who have entered unsubsidised employment who obtain qualifications.

There are also customer satisfaction measures for both employers and participants. The earnings change from before to after the training is intended to capture the impact of the programme on earnings. However, before and after earnings measures can be unreliable as average enrollees’ earnings will dip just before they enter training, so a rise in earnings may not be a true indicator of programme impact.

Third, the standard an agency must meet to be eligible for an award is no longer derived by applying a regression formula based on the characteristics of jobseekers served and labour market conditions. Rather, states negotiate standards with the Department of Labor. This move to negotiated standards was primarily aimed at reducing the data collection burden on states. However, it resulted in low enrolment rates for the programme, “due to a reluctance in local areas to officially register people in WIA because of concerns about their ability to meet performance goals, especially the ‘earnings gain’ measure.”

Finally, states are no longer allocated fixed sums to be awarded as incentives to agencies. Instead, the amount given to the states depends on the aggregated performance of their agencies. Only those states whose results exceed the standards in all areas of measured performance are eligible for awards.

Temporary Assistance to Needy Families

Another programme funded by the federal government ran in parallel to the WIA. In 1996 the United States Congress passed a law called the Personal Responsibility and Work Opportunity Reconciliation Act (PRWORA). This law established the Temporary Assistance to Needy Families (TANF) as a block grant to states. TANF
legislation established a ceiling on the federal government’s contribution to welfare costs and “placed a time limit on eligibility for federal income assistance, setting a cumulative lifetime limit of 60 months.”\textsuperscript{104} Under this legislation, states were required to demonstrate that certain percentages of welfare recipients were working or participating in work-related activities for a specified number of hours per week in order to be eligible for federal government reimbursement of welfare expenditure.\textsuperscript{105}

Table 1: TANF Requirements for Percentage of Participants in a Job or a Welfare-to-Work Programme and Weekly Work Hours

<table>
<thead>
<tr>
<th>Fiscal year</th>
<th>One-parent family</th>
<th>Two-parent family</th>
</tr>
</thead>
<tbody>
<tr>
<td>1998</td>
<td>20 hrs/week</td>
<td>35 hrs/week</td>
</tr>
<tr>
<td></td>
<td>30% participation</td>
<td>75% participation</td>
</tr>
<tr>
<td>1999</td>
<td>25 hrs/week</td>
<td>35 hrs/week</td>
</tr>
<tr>
<td></td>
<td>35% participation</td>
<td>90% participation</td>
</tr>
<tr>
<td>2000</td>
<td>30 hrs/week</td>
<td>35 hrs/week</td>
</tr>
<tr>
<td></td>
<td>40% participation</td>
<td>75% participation</td>
</tr>
<tr>
<td>2001</td>
<td>30 hrs/week</td>
<td>35 hrs/week</td>
</tr>
<tr>
<td></td>
<td>45% participation</td>
<td>75% participation</td>
</tr>
<tr>
<td>2002</td>
<td>30 hrs/week</td>
<td>35 hrs/week</td>
</tr>
<tr>
<td></td>
<td>50% participation</td>
<td>75% participation</td>
</tr>
</tbody>
</table>


States were also eligible for bonus payments if they reduced the number of families receiving assistance. Notably, however, the bonus was contingent on the absolute reduction in families on welfare, which could be achieved by means other than lifting poor families out of poverty.\textsuperscript{106} In fact, “more individuals participating in contracted programmes lost welfare benefits, despite lack of work, than left welfare because they found work.”\textsuperscript{107} There was also a modest financial incentive for states to contract out some welfare to work programmes to faith-based organisations.\textsuperscript{108} These new requirements were accompanied by a block grant of $3 billion to subnational units of government to enable them to expand their services.\textsuperscript{109}
State programmes

In addition to federal programmes, there are a number of state-funded welfare to work schemes that use payment-by-outcome as a tool to incentivise the achievement of objectives. The following is not an exhaustive discussion of all state-funded programmes, it is simply a description of some of the most well-studied.110

Wisconsin

‘Wisconsin Works’ (W2), a new welfare system based on the ‘work first’ principle, was established in 1995.111 The programme was deliberately designed to reduce welfare dependency and re-link the receipt of income to work. Under W2, individuals are required to enter one of four tiers of employment. Wherever possible, they are encouraged to take up unsubsidised employment. Where individuals lack skills or experience and this prevents them from entering unsubsidised employment, they may take up a ‘trial job’, which is subsidised in order to offset the cost to the employer of training the new employee. Jobseekers who cannot find private employment can enter the third tier, ‘community service jobs’, which involves working for a non-profit or government organisation in exchange for benefit payments rather than wages. To simulate real wage work, the benefit payment to each individual is reduced by the amount of the hourly minimum wage for each hour of work the recipient misses. Finally, for individuals unable to perform independent work, perhaps due to disability, transitional jobs are available which involve vocational training or work activity consistent with personal abilities.112 As individuals move up the tiers of the employment ladder, their net after-tax income increases, so there is an incentive for jobseekers to accept the highest tier they are capable of obtaining.113

Wisconsin is split into counties, and for the purposes of W2, Milwaukee, the largest city, was split into six regions.114 An open competition was held in which government agencies, not-for-profits, for-profit companies and quasi-governmental organisations competed to provide services in the counties and regions.115 The first round of contracts ran from September 1997 to December 1999.116 Counties calculated the total amount of welfare expenditure, including benefits, employment services and administration, and contracts were worth 80% of this figure after benefits had been subtracted. Agencies that generated savings beyond the contract value could keep the first 7%. Beyond that, savings were split: the state received 45% (in the form of budget savings), 45% was allocated by the agency to their local community for expenditures in general support of low-income families
Payment-by-Outcome in Welfare to Work

(With approval by the state), and 10% reverted to the W2 agency in the form of unrestricted profit. Agencies whose expenditure exceeded the value of the contract were obligated to pay the difference.117

W2 was regarded as a huge success in terms of reducing the welfare caseload and generating cost savings to the state. In 1996-97, the welfare caseload was 55,000. Three years later it had decreased by 80% to 11,000.118 Of those no longer enrolled in W2 at the end of its second year in operation, 76% were working; of this group about 50% were earning between $12,000 and $18,000, about 33% were earning less than $12,000 and about 20% were earning more than $18,000.119 Moreover, the cost savings were higher than even the most optimistic projection.120 Critics have accused W2 of being responsible for the rise in the child poverty rate from 1998, but the poverty rate among participants in the programme fell from about 77% to 67% from 1998 to 1999 while average family income rose from about $12,000 to $15,000 over the same period.121

W2 has had some problems, especially in subsequent contract periods. Two agencies improperly billed W2 for work done outside the programme and were forced to repay the sums with compensation. The state government also made some mistakes in terms of managing the market. Recently, a new agency was created dedicated to managing W2, which may improve the situation.122

Chicago

As seen earlier, under the TANF legislation, each year for five years the federal government required a higher percentage of welfare claimants to be in work and for more hours per week in order for states to be reimbursed for a proportion of their welfare costs. When Chicago contracted out some of its employment services provision, contracts were tied to the federal government’s terms. Providers were paid fees for case management and job preparation, and outcome payments of between $700 and $1300 for individual job placements. At first, these job placements could be at minimum wage, and had to last between 30 and 90 consecutive days with a minimum number of hours worked.123 Job placements were later redefined to include a job retention feature, which required providers to ensure individuals remained in work for 90 consecutive days or 150 non-consecutive days. Jobseekers could fulfil the requirement by working one or multiple jobs. Some contracts further rewarded jobs remunerated at more than the minimum wage or placements of clients designated as ‘hard-to-serve’.124
Providers were required to agree a ‘responsibility and service plan’ with clients, detailing the responsibilities of each party. This should have led to greater accountability of providers to clients, but because the process was very one-sided, more often resulted in clients being forced to accept providers’ terms in order to continue to receive their benefits.\footnote{125}

Fifty-seven percent of those participating in welfare to work programmes in 1998-99 found work. Forty percent retained their jobs for six months and 12% for nine months. Only 8.1% of those who found jobs earned salaries that put them above the poverty line, and one study found that these higher earnings were due to client characteristics rather than any effect of the programmes.\footnote{126}

Australia

In Australia, welfare to work policy and programmes are designed and implemented nationally. Due to the similarities of the UK’s system to that of Australia, the various programmes that have used payment-by-outcome are important examples to study.

Working Nation

Between 1994 and 1996 Australia established a new system of welfare to work provision with the objective of ensuring more unemployed people found jobs more quickly, thus generating savings to government. Working Nation’s central initiative was a Job Compact which guaranteed jobseekers who had been unemployed for 18 months or more access to a case management service, leading to employment, training or both.\footnote{127} Alongside this, the delivery of employment services was restructured. The public employment service, Commonwealth Employment Service (CES), remained responsible for job matching services for jobseekers. However, part of the provision of more intensive services was outsourced.\footnote{128}

Employment Assistance Australia (EAA) was created as an agency of the CES. EAA remained responsible for providing case management services to approximately two-thirds of jobseekers.\footnote{129}

At the same time, a network of non-profit and for-profit providers was made responsible for about one-third of jobseekers. An independent regulator, the Employment Services Regulatory Authority (ESRA), was created to license the private providers and allocate workloads among them, and was also responsible for monitoring all providers, including EAA.\footnote{130} The CES was responsible for classifying jobseekers according to the ‘degree of difficulty’ of placing them in work when they reached 12 months of unemployment, so that providers were paid more for those most difficult to help. Jobseekers were obliged to
enter into a ‘case management activity agreement’ with their providers and to undertake all activities prescribed by them.\textsuperscript{131} Providers were required to report any instances of non-compliance and recommend the suspension of social security payments. Providers were paid a fee for each jobseeker they registered and earned additional payments when clients obtained work or began an approved training programme and when they completed 13 weeks of work or training.\textsuperscript{132}

\textbf{Job Network 1}

In 1996, the new government declared Working Nation a failure, claiming it was expensive and ineffective, and established a new system.\textsuperscript{133} Centrelink, a public provider, was made the gateway to employment services, determining jobseekers’ eligibility for services and providing information. Centrelink then referred jobseekers to the network of private and voluntary sector providers, known as the Job Network. Jobseekers were allowed to choose their providers, but in practice most did not exercise choice.\textsuperscript{134} The Job Network developed through three distinct periods and has now been replaced by Job Services Australia, which is a very similar system.

Under Job Network 1 (1998-99), providers performed one or more of ‘job matching’, ‘skills training’ and ‘intensive assistance’ functions. They had flexibility over which services they provided and how they provided them. Providers were paid service fees by the Department of Employment and Workplace Relations when they took on jobseekers. In addition, providers could earn outcome payments for ensuring jobseekers completed education or training courses, entered work, or sustained work for 13 or 26 weeks. Providers could also earn payments for jobseekers placed in jobs they listed on the National Vacancy Data Base. Under this new system, although providers could ask jobseekers to sign activity agreements, they were no longer able to ‘breach’ (sanction) clients, as this was made the responsibility of Centrelink.\textsuperscript{135} Jobseekers were assessed using the Job Seeker Classification Instrument (JSCI) as Level 1, 2 or 3 in terms of barriers to work, and based on this, referred to one of the three major services. Jobseekers eligible for intensive assistance were further categorised into three groups. Each group attracted a different level of payment, so that providers could earn more by placing harder to help jobseekers.\textsuperscript{136}

\textbf{Job Network 2}

A variety of changes aimed at resolving the problems revealed in Job Network 1 were made in the second round of contracting (1999-2003). The Department was concerned that providers were making large profits from their job matching and
job search training services, but neglecting the more difficult to help jobseekers on intensive assistance. In Job Network 2, therefore, providers were required to agree a ‘preparing for work’ plan for anyone unemployed for 13 weeks, identifying barriers to employment and strategies to overcome them. In addition, the use of the JSCI was changed slightly, with individuals being classified as either Level A or B, reflecting the severity of barriers to work.

**Job Network 3**

Further changes were made under Job Network 3 (2003-2006, extended to 2009). The distinction between the three levels of service was abolished in favour of a continuum of services. All providers were therefore obliged to provide ‘intensive support customised assistance’ which began after 12 months unemployment or immediately for those jobseekers classified as highly disadvantaged. Providers could earn higher payments for jobseekers who had been unemployed for longer periods. Jobseeker Accounts, funds which could be used only to pay for training for jobseekers, were also introduced during Job Network 3. Jobseekers classified as the hardest to help received the highest amount of funds; however, Jobseeker Accounts were not individual entitlements, so providers could allocate funds across the jobseekers as they saw fit. The funds could not be retained by providers as profit, so there was an incentive to use the funds for training purposes that would increase the likelihood of jobseekers obtaining work. Jobseekers were able to choose a provider but had to remain with their chosen provider throughout a period of unemployment.

In all of the Job Networks, multiple providers were operating in most geographical areas. ‘Star ratings’, relative measures of provider performance taking into account labour market conditions and jobseeker characteristics, were introduced to facilitate this choice.

Evaluations of the Job Network have consistently revealed substantial cost savings. A 2002 evaluation showed that placing a client in a job under JN cost between $5000-6000 compared to $10,000-16,000 under previous arrangements. The JN has also been praised for establishing clearer objectives and incentives for welfare to work schemes. Both employers and jobseekers have given favourable feedback.

More disputed, however, is the performance of the system in terms of achieving sustainable job outcomes, and there are legitimate concerns about deadweight cost (the cost of achieving outcomes for jobseekers who would have found employment without help) and parking of the hardest to help. For example, in Job Network 2, only
15% of those that began the programme found work and retained it for 13 weeks, and approximately 70% of provider income was earned through non-outcome-based commencements fees. Moreover, while the Department claimed Job Network’s intensive assistance programmes had a net 10% positive impact on employment outcomes, the Productivity Commission found that this was “significantly overstated”.

**Job Services Australia**

In 2009, a new model was introduced called ‘Job Services Australia’. This system has some similarities to Job Network 3. Providers have access to Employment Pathway Funds (similar to Jobseeker Accounts) and receive service and job placement fees, as well as outcome payments for 13- and 26-week job placements, with those jobseekers furthest from the labour market attracting higher payments. Job Services Australia, however, includes a more prescriptive service delivery model. Jobseekers are classified into one of four ‘streams’ based on the barriers to work they face. Each stream has ‘service guarantees’, or services that every organisation must provide to all jobseekers in that particular stream. The result is that providers’ flexibility to tailor services is curtailed. Star ratings have been modified but continue to play a role.

**The Personal Support Programme**

The Personal Support Programme (PSP), introduced in 2001, operated separately from the Job Network and continues to operate separately from Job Services Australia. It provides intensive case management over a two year period to jobseekers facing multiple personal (non-vocational) barriers to work. The aim of the PSP is to achieve increased economic and social participation. The “outcomes [PSP providers aim for] are expected to be matched to the abilities, capacities and circumstances of the participants.” Participants may exit into employment, and this is seen as a desirable outcome. However, it is recognised that this is not always possible for PSP clients, so participants have also exited into the Job Network (although providers estimate only about 5% to 10% did so) or Disability Employment Assistance (DEA), moved onto the Disability Support Pension or become ‘inactive’.

PSP participants experience a high level of disadvantage: 50% have been homeless in the five years before joining the programme, 70% have completed Year 11 of secondary school or lower as their highest level of education and 78% suffer from a mental health problem. On average, participants have been unemployed for about two and a half years before beginning the PSP.
Providers are required to engage with clients a minimum of once every four weeks, face-to-face or over the phone. Outreach is required by the PSP contract, so many providers conduct home visits or meet clients in cafés.\textsuperscript{153}

The programme receives comparatively little funding; providers have access to between $0 and $120 per client per year, compared to the $1350 per disadvantaged jobseeker in the Job Network allocated through the Jobseeker’s Account.\textsuperscript{154}

The PSP payment structure combines payments for inputs, outputs and outcomes, as shown in the table.

<table>
<thead>
<tr>
<th>Payment type</th>
<th>Amount</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Administration payment</td>
<td>$660</td>
<td>Commencement payment when participant starts with provider</td>
</tr>
<tr>
<td>Administration payment</td>
<td>$660</td>
<td>Action plan payment on the production of an action plan with the client</td>
</tr>
<tr>
<td>Administration payment</td>
<td>$165</td>
<td>Exit payment when a participant exits and an exit report is submitted</td>
</tr>
<tr>
<td>Timing payments</td>
<td>$660 each</td>
<td>When a participant completes eight and 16 months on the programme</td>
</tr>
<tr>
<td>Outcome payment</td>
<td>$825</td>
<td>Social outcome payment after two years on the programme and submission of an exit report detailing the social outcomes achieved</td>
</tr>
<tr>
<td>Outcome payment</td>
<td>$1100</td>
<td>Durable economic outcome payment at 13 weeks</td>
</tr>
<tr>
<td>Outcome payment</td>
<td>$440</td>
<td>Durable economic outcome payment at 26 weeks</td>
</tr>
<tr>
<td>Other payment</td>
<td>$550</td>
<td>Remote loading payment for providers to work with participants in remote areas ($275 at commencement and $275 with the eight month timing payment)</td>
</tr>
<tr>
<td>Other payment</td>
<td>$220</td>
<td>Interpreter payment paid on commencement (eligibility determined by Centrelink)</td>
</tr>
<tr>
<td>Other payment</td>
<td>$165, $165 and $330</td>
<td>Transient participant payments: a reconnection payment on referral to a new PSP provider; a payment for production of exit report detailing progress to date; a recommencement payment for the new provider</td>
</tr>
</tbody>
</table>

Providers can receive outcome payments for ‘economic outcomes’, which include clients achieving 13 and 26 weeks in work. Those who complete two years on the programme and have not achieved an economic outcome are automatically considered to have achieved a ‘social outcome’, such as increased self-esteem or better interpersonal skills. Providers receive a payment on submission of a report detailing the social outcome achieved for clients completing the 24-month programme.

Providers are also judged according to certain criteria for the Performance Indicator Framework, which helps determine whether providers will be offered further PSP contracts. These criteria are the referral to commencement ratio, the retention rate and the timeliness of exit reports.

The United Kingdom
Welfare to work programmes in the UK are designed and administered centrally by the Department for Work and Pensions (DWP). For many years, jobseekers have accessed different welfare to work schemes depending on the type of benefit they claimed and where they lived. For example, Employment Zones (EZ) ran at the same time as the various New Deal programmes, and jobseekers with the same characteristics accessed either EZ or a New Deal programme depending on the region in which they lived. Rather than discussing all the schemes in detail, this section analyses three programmes, concluding by describing the Coalition Government’s plans for a new Work Programme.

Employment Zones
Employment Zones (EZ) was “the most radical” experiment of the use of the private sector to promote contestability in employment services. It is also arguably one of the most successful models of payment-by-outcome in this sector. Employment Zones was piloted in five areas from April 1998 to 2000. In 2000, 15 zones with high levels of deprivation were selected to participate in the rolled-out programme. The scheme was originally intended for individuals aged 25 years and over who had been claiming Job Seeker’s Allowance (a type of unemployment benefit which requires jobseekers actively to seek work) for at least 18 of the previous 21 months. It was subsequently extended to certain young jobseekers and some single parents. The zones had varying populations of between 1000 and 7000 eligible jobseekers.
In 2003-2004 when contracts were re-tendered, the number of zones was reduced to 13 as some zones in London were merged.\textsuperscript{162} These new contracts were for five to seven years.\textsuperscript{163} Whereas previously there was a single provider in each zone, the second contracts introduced competition through multiple providers in six of the zones. In these multiple-provider zones, participants were randomly allocated to a provider, allowing evaluation of performance through yardstick competition.\textsuperscript{164}

Jobseekers on the Employment Zones programme passed through three different stages of service provision. Stage 1 was overseen by the public ‘gateway’ provider, Jobcentre Plus. Certain process requirements had to be completed during this phase, including writing a personal, costed ‘action plan’ and completing a ‘better-off in work’ calculation.

Stage 2 was spent entirely with a private or voluntary sector provider, and the only process requirement was that jobseekers have ‘meaningful contact’ with their providers at least every two weeks. Providers were given 21 weeks of benefit payment from the government when jobseekers began Stage 2, which they had to pay onto the jobseekers until they found employment. However, Stage 2 lasted 26 weeks, so if jobseekers remained unemployed for the entire duration of Stage 2, providers had to pay the last five weeks of benefit payment themselves; that is, providers made a loss on those jobseekers. If, however, jobseekers found employment before the end of 21 weeks, providers retained the benefit payments as revenue. There was therefore a strong incentive for providers to support jobseekers into work quickly. In addition, providers received outcome payments for job placement and job retention for 13 weeks. They could also receive bonus payments for placing a certain proportion of their total caseload in employment, but providers did not know what proportion they had to place to earn this bonus, nor were they aware of the amount of the bonus. This further incentivised providers to perform.\textsuperscript{165}

Providers could ‘sanction’ jobseekers who did not comply with their action plans by stopping benefit payments. Providers did not require Jobcentre Plus approval to take such action; however, the appropriateness of sanctions was assessed ex post as jobseekers had to re-apply for benefits through Jobcentre Plus.\textsuperscript{166} There is no evidence that the ability of providers to sanction jobseekers was abused; one large-scale survey of New Deal and EZ participants found that 12% of EZ clients had been sanctioned in the previous year, compared to 18% of New Deal clients in matched comparison areas.\textsuperscript{167}
By the end of June 2003, 80,000 people had participated in a zone and more than one-third had held a job for 13 weeks. A quantitative evaluation of administrative data found that there was “a small but significant programme impact in exits from unemployment during the first year” of zone activity. This effect was not at the expense of other groups of unemployed people; that is, EZ participants did not take jobs away from non-participants. However, by the third year of operation, programme impact had been eroded considerably.

The findings from the quantitative evaluation were confirmed by qualitative research based on the results of large-scale participant surveys. By the end of the first ten months of EZ, 34% of zone participants were or had been in paid employment, compared with 24% in New Deal. This was statistically significant. At the second round of interviews, ten months later, 55% of zone participants and 51% of New Deal participants were in jobs or had been in them, a difference which was significant but not statistically so. However, “it is important to note that over the research period the comparable New Deal programme was itself changed, incorporating some key features from the zone approach”, so the second comparison may not be valid.

In both EZ and New Deal, one significant finding was that jobs lasting 13 weeks often did not last over the longer term. Many jobs were temporary, with small employers and offered low wages. This suggests that 13-week outcomes are not strongly correlated with long-term employment outcomes.

In terms of cost-effectiveness, assuming that the number of jobseekers who would have found work in the absence of the programme, the duration of employment and wage levels were the same in EZs as in comparator New Deal areas, EZs cost significantly more than the New Deal. For clients aged 25 and over, each programme start cost 40% more and each job cost 15% more than the New Deal, while sustained jobs cost about the same. Other measures of cost-effectiveness also show that EZs were more expensive than the New Deal.

Pathways to Work
Pathways to Work (Pathways) built on the New Deal model, and was thus a very different programme from Employment Zones. Pathways was first piloted in October 2003; it was expanded to 14 areas in 2005-6. In these areas, representing about 40% of the country, the programme was administered by Jobcentre Plus. The
remaining districts were contracted out to private and voluntary providers in 2007-8; this was known as Provider-led Pathways (PL Pathways). Contracts were for three years; Phase 1 contracts expired in December 2010 while Phase 2 contracts will terminate in April 2011.

Pathways provision was originally available to all Incapacity Benefit claimants (people who cannot work because they have an illness or disability). This group of approximately 2.63 million people had not previously been targeted for mandatory participation in work activation programmes, as the existing New Deal for Disabled People was entirely voluntary. The introduction in October 2008 of Employment and Support Allowance (ESA), which is replacing Incapacity Benefit, has meant that all new Incapacity Benefit claimants are assessed to determine if they are capable of working. Those who are deemed capable must participate in Pathways. Pathways providers can also attract voluntary clients for whom participation is not yet an obligation.

Clients attend up to six Work Focused Interviews, five of which are delivered by Pathways providers, in the first 12 months of their claim. Failure to attend can result in a reduction in the amount of benefit of 25%. In practice, the rate of sanctions is very low, around 1% for Pathways participants.

The aim of Pathways is to move people into work. PL Pathways providers are assessed based on the number of clients achieving 26 weeks in work and are also evaluated against their contractual targets which specifies the number of people they will return to work. Providers receive a monthly service fee worth a total of 30% of the contract value. The other 70% of payment is outcome-based. Job outcome payments, paid when clients begin employment of at least eight hours per week while on or within six weeks of leaving the programme, amount to 50% of the contract value. Sustained job outcome payments, paid when clients have been in employment for a minimum of 16 hours per week for at least 13 of the previous 26 weeks, are worth 20% of the contract value. Providers have monthly targets for job and sustained job outcomes.

The average payment received per job outcome (as all providers have different unit prices) is £1003. Providers receive the same payment for all jobseekers placed in work, whether they are mandatory or voluntary participants. Originally, there was a cap on the number of claims providers could make for placing voluntary clients. This was subsequently removed, but has not resulted in a significant change in the numbers of claims.
There is evidence that the lack of variation in payment has led to a focus on the easier-to-help (often voluntary) clients, at the expense of other jobseekers. In fact, 75% of prime and second tier contractors who responded to an NAO survey said that the current payment system diverted efforts away from jobseekers who needed extra support. This claim is supported by the fact that 40% of all job claims under PL Pathways are for voluntary clients, compared to 9% in Jobcentre Plus-administered areas. Overall, PL Pathways is helping clients into work at a similar rate to Jobcentre Plus and at similar cost. Jobcentre Plus has performed better for mandatory clients than PL Pathways; 11.1% of mandatory clients served by Jobcentre Plus achieved job starts compared to 9% served by PL Pathways. Contractors have underperformed compared to targets in contracts with the Department for Work and Pensions; the reasons for this are explored in Section 8.

**Flexible New Deal**

In 2009, the government introduced the Flexible New Deal (FND). This programme was designed to support individuals who had been claiming Jobseekers Allowance for 12 months or more into work. Under FND, jobseekers receive support from Jobcentre Plus for the first 12 months of their claim. At the end of 12 months they are referred to private and voluntary providers of FND. FND thus replaced most of the New Deal programmes.

The programme was split into two phases. The first phase involved rolling the programme out in 14 contract areas. In ten of these, two prime contractors deliver services in competition, while in the other four there is a single prime provider. The roll out of Phase 2 has been cancelled due to the Coalition Government’s plans to replace FND provision with the Work Programme (see below).

FND is intended to give providers more freedom to deliver the services they believe would make a difference to individual jobseekers’ employment prospects. However, there are certain requirements; for example, providers are required to develop an ‘action plan’ of mandatory activities for jobseekers including a minimum of four weeks full-time work-related activity. Jobseekers generally receive FND services for up to 12 months, but if at the end of the period they are still unemployed, provision can be extended for six months.

Under FND, prime providers were supposed to be paid 20% of the total contract value through an upfront service fee, 50% for 13-week job outcomes and 30% for 26-week job outcomes. However, due to the economic circumstances, DWP
revised this payment structure so that the service fee would be worth 40% of the total contract value for the first 18 months; the remaining 60% would be for job outcomes. DWP gave providers targets of achieving 13-week job outcomes for 55% of participants and 26-week job outcomes for 50%.

Provision is structured so that prime providers contract directly with DWP, but are able to subcontract various aspects of service delivery to smaller organisations. In practice, some prime providers deliver all FND services in-house and only use subcontractors for certain specialist services. A similar model involves the prime provider focusing on its strengths, delivering one or more elements of delivery and subcontracting the other elements to smaller organisations. Other prime providers deliver end-to-end services and subcontract with other organisations to do the same in other areas. Finally, some prime providers outsource all services, maintaining a small staff to manage the subcontractors. In one outsourced model, jobseekers move to a different subcontractor for each stage of the prime provider’s service model.

Despite the ‘black box’ approach to service provision, which gives prime providers substantial freedom in designing services, most providers deliver a similar set of services, including:

- diagnosing jobseekers’ needs and writing action plans;
- employability courses;
- motivational courses;
- work-related activity;
- specialist provision; and
- in-work support.

Some variation in service provision has been observed in the initial diagnosis phase. Some providers deliver the same core service to all jobseekers, but provide tailored wrap-around services. Others classify jobseekers according to their distance from the labour market, so that groups of jobseekers receive services tailored to their particular needs. Other notable service innovations included employing previous clients as job advisors and using a contract between the jobseeker and the provider as part of the action plan process.

A quantitative analysis of the impact of FND on employment outcomes is planned, but data is not yet available.
Work Programme

The Work Programme is the Coalition Government’s new employment support programme, intended to be operational by summer 2011.196 The Work Programme will replace many of the New Deal and Flexible New Deal (FND) programmes currently in place and as such will serve most jobseekers regardless of benefit type. The aims of the programme are to ensure more jobseekers return to work, more quickly, and sustain their jobs for longer than under previous programmes, with a particular emphasis on achieving these goals for disadvantaged groups. In addition, the programme is intended to “contribute to a decrease in numbers of workless households”.197

Referrals to the programme will be more flexible than under FND; certain jobseekers (for example those classified as ‘facing significant disadvantage’) will have access to this programme after only three months of unemployment and Jobcentre Plus advisors will be able to delay referral if a client is very close to finding work.198 Providers will work with jobseekers for two years.199

The prime contractor model will continue to be used. There will be competition between providers in each contract area, and while referrals will be random and equal at the start of contracts, the better performing provider will be able to increase its market share over the course of the contract.200 Contracts will run for five years, with providers working with clients for up to two years after the end of the contract. An element of flexibility will be built into contracts so that unforeseen changes can be accommodated within them.201

Providers will be paid an ‘attachment fee’ when a jobseeker joins the programme and a ‘job outcome payment’ when a client obtains employment and retains it for 13 or 26 weeks. The amount of the attachment fee will decrease to 75% of its original value in Year 2 and 50% in Year 3, and in Years 4 and 5 will be nil. Maximum job outcome payments will also decrease starting in Year 3 of the contract, for certain client groups. In addition, for every four weeks a client retains work, providers will receive a ‘sustainment outcome payment’, in some cases for up to two years after a client achieved a job outcome. They will earn more for achieving outcomes for jobseekers classified as harder to help. Providers who achieve a certain standard of performance in relation to job outcomes for certain groups of clients will be eligible for incentive payments.202

Providers will be required to specify ‘minimum service standards’ in their tenders, which will be translated into Key Performance Indicators and articulated
to clients so they can hold providers to account. Moreover, the government will apply minimum performance standards expected of providers for certain client groups, with contractual action, including contract termination, a possibility for providers failing to meet those standards.

The government plans to award contracts based on the quality and price of bids, which will be given equal weighting. Costs will be controlled by the establishment of a maximum price for job outcome fees, which providers will be able to bid under but not above.
Endnotes


3. Ibid., 38-39.


15. Ibid., 1.


18. Ibid., 2.


23. Ibid., 1.

24. Ibid., 29, 32.

25. Ibid., 32.


27. Ibid., 36.

28. Ibid.

29. In the case of Australia’s PSP, both parties appear comfortable with the very basic measure of a written submission by the provider at the end of two years describing the improvements in human capital that a particular client has achieved. However, this is not a measure that ‘proves’ in any objective sense an increase in human capital.
30. These can be extrapolated from the DWP’s vision to: contribute towards fair, safe and fulfilling lives, free from poverty for children, people in work and retirement, disabled people and carers; reduce welfare dependency and increase economic competitiveness by helping people to work wherever they can and helping employers to secure the skills and employees they need; and provide greater choice and personalisation and higher quality of service for customers where it is in their interests and those of the taxpayers. Graham Taylor, *An analysis of the productivity of the Department for Work and Pensions 2002/03 to 2008/09* (Leeds: DWP Planning and Performance Management, 2010): 4.


35. Please note these numbers are used for illustrative purposes only.


42. Ibid., 40.


44. These included: Reducing the welfare caseload by 10 to 25% depending on prior results; placing greater numbers of recipients into unsubsidised employment of at least 20 hours per week, increasing over three measurement periods; placing those not working in private employment into the equivalent of community service or other jobs, increasing over three measurement periods; reducing expenditures for cash assistance; credit given to counties which had cooperated with the state and performed well in the past. Turner, *op. cit.*, 41.

45. Ibid., *op. cit.*, 41.

46. Ibid.


49. Most commissioners do specify a certain number of hours per week that clients must be in work in order for providers to be able to claim the outcome payment.


57. For a more detailed explanation of the winner’s curse and the conditions under which it is most likely to occur, please see Alastair Dick, *Blind Dates and the Winner’s Curse* (London: The Serco Institute, forthcoming).


59. Prime Provider Manager, quoted in Hudson et al., *op. cit.*, 24.

61. Ibid., pp. 52.


63. The new government was keen to break from the previous government’s model of ‘case management’, so the informal message panels received was not to give too much weight to past performance where this was based heavily on successful case management.

64. Considine, op. cit., p. 53.

65. Ibid., p. 52.


68. Productivity Commission, op. cit., p. 10.4.

69. Ibid., p. 10.5.

70. Ibid.

71. Quoted in Productivity Commission, op. cit., p. 10.4.


75. Ibid.


77. Hudson et al., op. cit., p. 24.

78. Ibid., pp. 24-25.


82. Ibid.


89. Barnow and Smith, op. cit., p. 31.


93. Ibid., p. 910.


97. Ibid., p. 23.

98. Ibid., p. 45.


100. Courty and Marschke (2003), op. cit., p. 46.

101. Barnow and Smith, op. cit., p. 28.

102. Quoted in Barnow and Smith, op. cit., p. 32.

103. Courty and Marschke (2003), op. cit., p. 46.

104. Brodkin, op. cit., p. 75.

105. Ibid.

106. Ibid., p. 76.

107. Ibid., p. 93.

108. Ibid., p. 81.

109. Ibid., p. 77.
110. Turner asserts that W2 is “almost certainly... the most well-studied and analyzed program of welfare reform in the U.S.” with 53 studies completed or underway in its first three years in operation. Turner, op. cit., 43.
111. Ibid., 35.
112. Ibid., 37.
113. Ibid., 37-38.
114. Ibid., 40.
115. Ibid., 41.
116. Ibid., 40.
117. Ibid., 40-41.
118. Ibid., 43.
119. The 1998 poverty threshold for a family of three was $13,001.
120. In Milwaukee, $319 million had been allocated for all expenditures including cash benefits. Of this amount 84% was spent, leaving $26 million for agency profits, $13 million for reinvestment in the community and $13 million to the state in the form of savings, as provided for in the contract. Turner, op. cit., 46.
121. Ibid., 43-44.
122. Ibid., 46-47.
124. Ibid., 88.
125. Ibid., 93-94.
126. Ibid., 90.
127. Considine, op. cit., 43.
129. Saunders, op. cit., 16; Considine, op. cit., 43.
130. Considine, op. cit., 42.
131. Ibid., 45.
132. Ibid., 46.
133. Saunders, op. cit., 16.
134. Ibid.
135. Breaching is the Australian term for sanctioning, which involves reducing the amount of benefits jobseekers receive if they fail to comply with the obligations government has placed on them. Considine, op. cit., 49.
136. Ibid., 50, 53.
138. Ibid., 23, 24.
139. Considine, op. cit., 65.
141. Considine, op. cit., 64.
142. Morrell and Branosky, op. cit., 39.
144. Saunders, op. cit., 21.
150. Ibid., 9.
151. Ibid., 10.
152. Ibid., 22, 25, 28.
153. Ibid., 40.
154. Ibid., 34.
155. Economic outcomes are: transition to employment assistance (JN or DEA); employment; study; vocational education and training; apprenticeships and traineeships. Perkins, op. cit., 46.
156. Ibid., 44, 47.
162. Ibid., 392.
164. Finn, op. cit., 109-110.
166. Finn, op. cit., 112.
167. Ibid., 111.
168. Ibid., 112.
169. Quoted in Finn, op. cit., 113.
170. Finn, op. cit., 113.
171. Ibid.
172. Ibid.
174. Ibid., 26-27.
176. Ibid., 14.
177. Ibid., 19.
178. Hudson et al., op. cit., 8.
179. Ibid., 8-9.
180. Ibid., 24.
182. Ibid.
183. Ibid.
184. Ibid., 23.
185. Ibid.
186. Ibid., 24.
189. Vegeris et al., op. cit., 10.
191. Ibid., 21.
192. Vegeris et al., op. cit., 59-60.
193. Ibid., 61.
194. Ibid., 64-65.
195. Ibid., 11.
201. Ibid., 6, 9.
203. Ibid., 7.
204. Ibid., 14.
Payment-by-Outcome in Welfare to Work

Lauren M. Cumming