Outcome-based contracting for human services

Emma Tomkinson
Author contact: emma@emmatomkinson.com

doi: 10.4225/50/57C4E81F2FABD

Abstract

This review examines the evidence of the effect of government-funded outcome-based contracts in public human services. Outcome-based contracts in public human services are defined as those where some proportion of payment is triggered by some measure of change in the lives of clients. There is a lack of evidence comparing outcome-based contracts for public human services with other means of funding. There is also little evidence comparing the effect of payment on the basis of one measure of outcome to another, comparing outcome-based contracts to grants or block-funding models. And there is no evidence of the effect on outcomes of changing outcome-based payment structures as contracts progress. The evidence that does exist suggests that, given sufficient flexibility to do so, providers of services will deliver on the outcome metrics their contracts pay for. Outcome-based contracts developed so far have, however, struggled to create incentives to achieve the desired outcomes. The findings indicate that while outcome-based contracts deliver the measures of outcome for which they pay, these measures do not always reflect the intention of the contract designers, or desirable outcomes for the end-client. Measures of outcome that were not related to payment did not improve and sometimes worsened. Some outcome payments created incentives for service providers contrary to the achievement of desired outcomes. For example, employment services contracts that were meant to increase tailoring and flexibility had the opposite effect. Some contract conditions or environments constrained providers’ ability to affect outcomes. The challenge for government is to define payment metrics that represent the outcomes they seek and that encourage behaviour from service-delivery organisations consistent with these outcomes.

Governments around the world are increasingly outsourcing the services they fund. For example, it is estimated that up to 80 percent of US human services are now being delivered under government contracts with external providers, many of them not-for-profits (Collins-Comargo et al. 2011; Koning and Heinrich 2013; Martin 2005). Payments for these services increasingly include a component that is dependent on some measure of outcome (Collins-Comargo et al. 2011; Lu 2014; Martin 2005). Outcomes can be defined as changes made to people’s lives, or according to Martin (2005, 66) “results, impacts and accomplishment”. Outcome-based contracting for human services in most Western democracies has the longest history in employment services. It was introduced into US employment services in 1982, into Australian employment services in 1998, and via several models in New Zealand in the late 1990s and early 2000s (Kibblewhite and Ussher 2002). The UK introduced outcome-based contracting much later in employment services, in 2011, following the introduction of activity-based payments in the UK National Health Service in 2005.

The increased use of outcome-based contracting sits within the wider trend towards performance-based contracting, which includes contracts where payments are tied to performance, and also those where performance simply affects contract renewals and extensions (Martin 2005). This review focuses on contracts where payments are triggered by some measure of the desired outcomes. The enthusiasm of governments to adopt outcome-based contracting has outpaced the production of evidence on its effect (Collins-Camargo et al. 2011; Lu 2014). There has been a reluctance to capture and build on lessons from previous regimes, with new Governments presiding over rounds.
of consultation and re-launch of contracts, rather than conducting a program of incremental, informed improvements (Abetz 2015).

Outcome-based contracting of public services can also be contextualised as part of a wider reform context referred to as New Public Management, part of which is the increased outsourcing of publicly-funded services (Nicholls and Tomkinson 2013). In addition to the quantity of services outsourced or privatised, new service delivery environments have aimed to:

1. tailor services to the needs of the individual;
2. allow more flexibility as to how services are delivered;
3. increase accountability for outcomes (Considine and Lewis 2010).

In New Zealand, the reforms of the late 1980s and early 1990s shifted the focus of the budget from inputs to outputs, with outputs linked to outcomes through budget papers. Accountability for public spending was a key driver behind these reforms (Kibblewhite and Ussher 2002).

This paper begins with a definition of the terms and scope of the evidence review. It then sets out the aims and the search process used to identify the relevant literature. The methodological issues are then outlined before the findings are explained. This explanation touches firstly on the objectives of human services. Secondly, the effect of outcome-based contracts on people receiving services; governments commissioning services; and organisations delivering services are examined in turn. These are followed by a more technical discussion of the incentives and barriers created by different payment metrics. Finally, suggestions for future research are made and conclusions reached.

The aims of the review

This review examined the evidence of the effect of government-funded outcome-based contracts in public human services on clients and their communities. This began by looking at the basis upon which outcome-based payments are made in different contracts and the incentives that these payments created. The focus of the review was on empirical evidence, studies that contained some kind of evidence of effect, rather than discussion of the general policy area. The evidence of effect of outcome-based contracts was examined according to three stakeholder groups: government commissioners of services; service delivery organisations and their staff; and people receiving services. This includes literature on the unintended consequences created by outcome payments.

The review aimed to determine the range and findings of evidence related to the use of outcome-based contracting for public human services. It hoped to uncover whether outcome-based contracting is more effective at achieving results for clients and what the unintended consequences of these contracts are. It sought to uncover the benefits and challenges of outcome-based contracting for people who work to deliver services. And finally, it aimed to suggest areas where future research on outcome-based contracting should be generated.
Definition and scope

For the purpose of this paper, the term ‘outcome-based contracting’ is used to refer to a contract where at least some payments are dependent on some measure of change in a person’s life. While not analogous with all the following terms, it includes contracts referred to as:

- results-based contracting
- performance-based acquisition
- results-based funding
- pay for success
- pay for performance
- performance-based contracting (PBC)
- payment by results
- performance-based service contracting
- payment by outcomes

Social impact bonds are briefly discussed in this review. A social impact bond is the specific case where the upfront service delivery costs of an outcome-based contract are met by a third-party investor. The investor’s eventual repayment is dependent on the outcomes achieved.

The review is limited to government-funded contracts that deliver direct services to people to help improve their lives. It does not include the provision of infrastructure such as roads, or the provision of services in relation to regulatory systems like licensing of drivers. For the purpose of this paper, direct services for people are referred to as ‘public human services’ and include services that are referred to as:

- welfare services
- social services
- social welfare services
- services to vulnerable people

The term ‘outcome’ is used to refer broadly to the changes the service makes for clients of the services. The word ‘clients’ is used to refer to people receiving or using services.

Public services: Outcomes or activities?

It is apparent in the literature that many outcome-based contracts included fees for services, or service requirements, as well as outcome payments (Brucker and Stewart 2011; Department of Employment and Workplace Relations 2003, 2006; Department of Education, Employment and Workplace Relations 2009, 2012; Department of Employment 2015). This reflects the mix of requirements each contract places upon providers. For example, employment services contractors in Australia partly deliver support to get a job, and partly manage the requirements of receiving unemployment benefits from the Australian Government. This means that providers are financially rewarded for both the jobs they help people access (outcome payments) and the administration of the ‘mutual obligation requirement’ (service fees) (Department of
Employment 2015). Sometimes these fee-for-service components work against the achievement of outcomes services (Thomas 2007; Considine et al. 2011; Leon 2014).

Proponents of outcome-based contracting often describe it as a way to optimise public spending and ensure better outcomes for clients (Ng et al. 2009, Outcomes Based Healthcare and Capsticks 2014). There is evidence, however, that outcome-based contracting is more applicable to some circumstances than others. For example, Leonard and Graff Zivin (2005) showed that, given the choice, patients preferred outcome-focused treatments and payments for illnesses that require ongoing partnership and management, but fee-for-service treatments for health needs that required only one interaction with the service. It is important to recognise that there are publicly funded human services that should help clients achieve outcomes, and services that are much more immediate or administrative. Therefore, we should be careful not to demand all publicly-funded services pursue outcomes.

**Search process**

Five search strategies were employed to locate studies meeting our inclusion criteria for this review. These strategies were:

1. A keyword search of research databases relevant to social policy and public policy.
2. Examination of the bibliographies of the most relevant research papers and reviews.
3. Searches for literature citing relevant research papers and reviews, or citing key government departments as an author.
4. A search of the publication pages of key government departments.
5. Consultation with academics studying performance-based contracts.

Only published and professional reports were included. This review only includes studies published in English. Appendix 1 details the specific search elements used to find the studies included in this review.

In addition a tweet was sent out to several hundred people involved in social services, data promotion and social research. It read: “Know of any empirical evidence re effect of using #performancebased/#outcomesbased contracts for social/human services? Tweet me a link?” No responses were received.

Only peer-reviewed studies containing empirical evidence were included in the review, apart from the report from the Office of Federal Procurement Policy (1998), which was included because it was itself a review of a number of contracts.

The search strategy uncovered 19 studies showing empirical evidence of the effects of outcome-based contracting. Publication dates ranged from 1998 to 2014 and most are from the United States. These studies are listed in Appendix 2.

**Overall methodological issues**

Overall, while most studies looked at the effect of outcome-based contracts on clients, the type of effect on those clients was highly varied. This meant that studies were not easily compared in terms of what they tell us about outcomes-based contracting in general. For example, some studies focused on time to achieve outcomes, others focused on the number or quality of outcomes achieved (Appendix 2). Some of the evidence compared outcomes from services delivered under outcome-based contracts to outcomes for people who received no services, rather than for people who received
services under other types of contracts. This makes it very hard to separate conclusions about the effect of the contract from those about the effect of the service.

Most of the studies were non-experimental, which limited the findings to data or surveys collected after the service itself. There were three quasi-experimental studies (Lu 2014; McBeath and Meezan 2009; Washington et al. 2009), two of which looked at the impact of outcome-based contracting on clients, one in child welfare and one in employment. Both found improvement in performance on measures related to payment, but worse performance or no improvement on measures of desirable outcomes not related to payments.

Some studies were able to obtain administrative data to compare the same population or service before and after changes were made (Brucker and Stewart 2011; Office of Federal Procurement Policy 1998; Faith et al. 2010). Three compared survey results from just after the program was implemented to ten years later (Considine et al. 2014; Considine et al. 2011; Considine and Lewis 2010), and one simply asked interviewees to reflect back on changes (Ng and Nudurupati 2010). Other comparisons pursued in the study designs included similar services with different payment bases (Gordon and Heinrich 2004; Koning and Heinrich 2013; McBeeth and Meezan 2009; Verbeeten 2008), similar public services (Washington et al. 2009), and propensity score matching (Lu 2014). Five qualitative studies did not use any comparison groups, but looked at relationships between responses or common themes in responses (Collins-Camargo et al. 2011; Dias and Maynard-Moody 2007; Gates et al. 2005; Levy et al. 2012; Meezan and McBeath 2011).

The focus of the 12 studies related to employment was split evenly between effect on jobseekers and service delivery staff. Even within studies of the effect of outcome-based contracts on jobseekers there is wide variation of outcomes studied. These include job placement rate (Considine et al. 2014; Gates et al. 2005; Koning and Heinrich 2013; Lu 2014), time from registering in the program to gaining employment (Gates et al. 2005; Lu 2014), job retention/duration (Gates et al. 2005; Koning and Heinrich 2013), choice of clients served (Koning and Heinrich 2013), working hours (Lu 2014), and wages (Gordon and Heinrich 2004; Lu 2014).

There was only one study related to child welfare that focused on outcomes for the child, aided by administrative data from Wayne County (Detroit), Michigan (McBeath and Meezan 2009). The lack of studies on outcomes for children may be due to the sensitivity of the data and the difficulty of collating it from thousands of disparate case files. The other four studies on child welfare looked at service delivery organisations and their staff. Two looked at the changes in backend organisational practices such as IT systems and internal communications (Collins-Camargo et al. 2011; Meezan and McBeath 2011), and two looked at factors affecting staff satisfaction and intention to quit (Levy et al. 2012; Washington et al. 2009).

While we know there are several examples of outcome-based contracting in New Zealand (Kibblewhite and Ussher 2002), the literature search did not uncover any empirical studies of them.

The rest of the review examines the effect of outcome-based contracts by specific stakeholder group. The most common stakeholder group studied in the literature was clients – the people receiving services. Service delivery organisations and their staff were the next most frequent focus, and a few studies looked at the effect on government agencies.
Effect on people receiving services

Most of the studies found an improvement in outcomes for recipients of human services delivered through outcome-based contracts. But while the quantitative measures of performance rewarded by payment almost always improved, Verbeeten (2008) found that unmeasured quality of services did not follow suit. One example of this is frontline staff delivering employment services in Australia, who reported feeling that over time they had less flexibility to deliver services tailored to the individual needs of their clients (Considine et al. 2011).

Employment services

Designers of outcome-based contracts in employment services are continuously adjusting their outcome-metrics and payments in an effort to create the desired set of outcomes for jobseekers (Considine et al. 2011; Department of Employment 2015; Leon 2014; Thomas 2007). Most outcome-based employment services contracts have tiers of payment, where people who face more difficulty finding and sustaining work attract higher payments (Department of Employment 2015; Lu, 2014). Despite this, several studies found that the incentives to service the most difficult clients were insufficient: these clients had poorer outcomes, were underserved, or ‘parked’ (Business Council of Australia 2014; Koning and Heinrich 2013; National Audit Office 2015). At the other end of the spectrum, ‘cream skimming’, the practice of favouring easier to serve clients, was also evident (Davidson and Whiteford 2012).

These practices, which were a direct response to incentives in employment services contracts, became a significant political issue in Australia in the late 1990s (Considine et al. 2011). In response, payments for finding work placements for those who have been unemployed for longer were increased (Davidson and Whiteford 2012) and payments for providers to work with those most likely to find jobs were lowered (Department of Employment 2015). More external regulation was also introduced, which Considine et al. (2011) found led to more conservative and standardised practices within service delivery organisations. This contrasts with the intention that outcome-based payments would lead to more service flexibility and tailoring. Employment services in New Zealand were not significantly privatised, although those on jobseeker benefits are still required to meet work obligations in order to receive benefits (Ministry of Social Development 2016).

As discussed earlier, most employment services contracts are structured using a mixture of service fees and outcome payments. Service fees are related to the administration of what in Australia is called the ‘mutual obligation requirement’ (Department of Employment 2015), i.e. the activities that must be fulfilled in order for unemployment benefits to continue to be paid. Outcome payments are attached to milestones related to achieving and then sustaining employment. For unemployed people who are most likely to get a job, the service delivery organisation is paid more through outcome-based payments. Conversely, payment for those who are less employable is more through service fees (Department of Employment 2015; Koning and Heinrich 2013). While this encourages the service of clients who have more difficulty accessing the job market, it lessens the incentive to find them employment, as it becomes more efficient to work for service fees than outcome payments.

The balance between fees and outcome payments is hard to get right and employment contracts are frequently adjusted to correct for unintended consequences (Considine et al. 2011; Leon 2014; Thomas 2007). The UK National Audit Office (2015) found that sometimes the value of service fees and outcomes payments has been
set such that providers will pursue only the service fee. They reported that clients were registered with programs, triggering the payment of an ‘attachment fee’, but payments for achieving outcomes were insufficient for service providers to invest the time and effort necessary to pursue them. Gates et al. (2005) found that weighting payments heavily towards sustained employment milestones (up to nine months employment) encouraged providers to work with people early and liaise with other government support agencies to find jobs that people want to keep. In contrast, Dias and Moody (2006) found that the absence of service fees resulted in insufficient time being spent with clients to address their barriers to employment and thus secure rewarding jobs that would last.

The number of fees in Australian employment service contracts more than quadrupled from 2000 to 2009 (Department of Employment, Workplace Relations and Small Business 1999; Department of Education, Employment and Workplace Relations 2009), creating incentives for service providers to provide services that attracted fees, such as training. The result of this was that services provided training that did not progress a person towards work (Leon 2014; Business Council of Australia 2014). The Business Council of Australia (2014) described the system as one where the set of rewards and penalties were not aligned to employment outcomes.

Three studies showed an improvement in employment outcomes that were directly related to payments (Gates et al. 2005; Koning and Heinrich 2013; Lu 2014), but there was no evidence of improvement in other desirable outcomes that were not payment-related (Koning and Heinrich 2013; Lu 2014). One example of this occurred when outcome-based contracts rewarded placement in a job; they achieved higher placement rates and shorter time to placement, but there was no evidence of these contracts increasing job duration or wages (Dias and Moody 2006; Koning and Heinrich 2013; Lu 2014). Considine et al. (2014) looked at profit-maximising attitudes and there was no relationship between these attitudes and improved outcomes for jobseekers.

*Child welfare*

There was only one study looking at the effect of outcome-based contracting on outcomes for children in care. Like employment services, services for children in care performed better on outcomes related to payment, but showed no improvement or worse outcomes for outcomes not related to payment. Reuniting a child with their biological family is widely considered the optimal outcome, however requires more resources than placing the child with another family or carer. When the outcome payments related only to placement in any family situation, placement rates rose overall, but the rate of placement with biological families dropped (McBeeth and Meezan 2009).

**Effect on service delivery organisations and their staff**

Meezan and McBeath (2011) found that for service delivery organisations and their staff, the transition to outcome-based contracting has been a source of anxiety, but concerns relating to revenue loss and high staff turnover were greater in anticipation of the transition than after it had occurred. In contrast, they found the organisational change required for the transition was underestimated by those yet to go through it. Other studies found the flexibility to respond to outcome payments was far more apparent in some contract settings than others, and strongly affected the role staff felt they played in decision-making (Considine et al. 2011; Nicholls and Tomkinson 2013).
Findings were not consistent across countries. Considine and Lewis (2010) found that with the introduction of outcome-based contracting for employment services, work satisfaction decreased in Australia, increased in the UK, and remained steady in the Netherlands. The daily number of clients seen by staff increased in Australia but decreased in the other two countries. They reported that the workforce in Australia became younger and more female, while in both the UK and the Netherlands the workforce aged and became less female.

Despite being introduced with the intention of allowing providers to tailor services to individual needs (Davidson and Whiteford 2012), Dias and Moody (2006) found that outcome-based employment service providers and their staff reported a decrease in flexibility to tailor services, and other studies described how frontline staff chose options that were explicitly approved by their government agency in order to avoid penalties for making the wrong decisions (Considine et al. 2011; Thomas 2007). Further iterations of the Australian employment services contract sought to reduce the administrative burden, while strengthening incentives to produce outcomes and increase flexibility (Department of Employment 2015).

The reason for this decrease in flexibility was partly an increase in government compliance measures and partly the reaction of service providers. Considine et al. (2014) found that staff reported a reduction in organisational appetite for innovation, in addition to an increasing administrative and compliance burden imposed by government. This manifested in onerous reporting requirements by delivery organisations on their staff, on top of what was required by government. The constantly changing contracting environment led to an increase in compliance-focused administration, furthering the loss of flexibility and innovation. Considine et al. asked about the computer system prescribed by the Australian contracting department, finding ‘close to 70 per cent of Australian employment services staff reported that a computer program directs them in their job, suggesting that flexibility was not a prominent feature of Australian employment services by 2008’ (2011, 7). Providers at the time of these findings were delivering services under contracts that detailed 132 different payments (Department of Employment and Workplace Relations 2006).

In contrast, staff delivering the Peterborough Social Impact Bond (delivered under only one payment metric) reported an increase in flexibility for innovation and the ability to continuously improve how outcomes are pursued: ‘We’re not being dictated to. We can shape the delivery model to meet needs’ (Nicholls and Tomkinson 2013). One of the key constraints for the achievement of employment outcomes is the number of jobs available. While most social outcomes are affected by the economic, policy, legal and cultural context in which they are pursued, the relationship is not as direct as the one between employment outcomes and availability of jobs. The number of unemployed people in Australia correlates more strongly with the number of job vacancies than any changes to employment services contracts.

In relation to child welfare services, McBeath and Meezen (2008) found that contract incentives sometimes negatively influenced employee behaviour. Performance deadlines were found to adversely impact employees’ ability to conduct thorough assessments, provide necessary services, and access community resources.

**Effect on government commissioners of services**

Government commissioners are hindered by a lack of evidence with which to design effective outcome-based contracts (National Audit Office 2015) as well as a lack of evidence as to the effect of these contracts on outcomes once implemented (Gates 2004;
Hedderman 2013; Lu 2014; National Audit Office 2015). Despite the challenges, studies found the benefits of outcome-based contracting for governments included lower prices (Office of Federal Procurement Policy 1998; Sturgess et al. 2007) and improved internal practices (Collins-Camargo et al. 2011), and no discernable increase in administrative costs (Koning and Heinrich 2013).

Government agencies designing outcome-based contracts for the first time have had difficulty developing metrics on which to base their payments. This includes establishing a baseline for outcome improvement, and the size of the potential improvement. The National Audit Office (2015) found that the Department of Work and Pensions overestimated the proportion of participants that would have found jobs without the support of the program, and also failed to model performance in different economic environments, resulting in every provider failing to meet minimum targets and being placed on a performance improvement plan. They also designed the payment regime before a system of validating it; it was estimated that from 2011 to 2014 they paid £11 million in invalid payments.

There were some benefits to public servants’ practices as they moved to managing outcome-based contracts with private providers of child welfare services. Collins-Camargo et al. (2011) observed positive changes in four areas: collaboration; communication; clearly defined and agreed performance; and integrated systems. While these practices would be considered desirable for any system of contracting, they became necessary for outcome-based contracting and were implemented in response to the demands of delivering the new contracts.

Koning and Heinrich (2013) did not find a difference between the administrative costs to government of full or partial outcome-based payments (i.e. whether 100 per cent or 50 per cent of payments were dependent on performance). A US review of 26 contracts from fifteen agencies that had transitioned to include performance-based payments showed that this transition had led to an overall 15 per cent decrease in price and 18 per cent increase in customer (government agency) satisfaction (Office of Federal Procurement Policy 1998). This is consistent with evidence relating the use of performance-based contracting to reduction of contract cost/size (Sturgess et al. 2007). The Office of Federal Procurement Policy (1998) also found an increase in the time spent to develop contracts for technical and professional services, alongside a 93 per cent decrease in audits overall. This investment of time in contract development, rather than contract management, has been mirrored in the development of outcome-based contracts for social impact bonds (KPMG 2014).

**The basis upon which outcome-based payments are made**

Many of the issues discussed above are a direct result of the incentives created by outcome-based payments. Outcome-based contracts create incentives linked to the way the payment-related outcome is defined and measured. There may be several different ways to measure one outcome, for example reducing reoffending. Different measures will contain different incentives. We can compare outcome-based payments for prisoners exiting Peterborough and Doncaster prisons to examine how this plays out in practice. The outcome measures are (Ministry of Justice, 2014):

For Peterborough prison the outcome measure is the frequency of re-conviction events (based on offences committed within 12 months of release from prison and convicted at court within those 12 months or a further 6 month period). This is often referred to as a frequency measure.
For Doncaster prison the outcome measure is the proportion of offenders who commit one or more offences in the 12 months following release from prison and are convicted at court in those 12 months or in a further 6 months. This is often referred to as a binary measure.

Nicholls and Tomkinson (2013) found that the Peterborough prison measure encouraged service providers to work with prisoners over the entire 12 months post-release, regardless of whether they were reconvicted in that time or not. The potential financial reward to work with prisoners who are expected to reoffend more frequently over that period is greater than the potential financial reward from working with prisoners who might reoffend once or twice. In contrast, Hichens and Pearce reported that the Doncaster prison measure encouraged the service provider to cease working with a prisoner if they were reconvicted within the 12 month post-release period, as this meant no financial reward would be gained from the prisoner: ‘Some delivery staff reported frustration that support is withdrawn, undermining the interventions previously undertaken’ (2014, 1).

In some cases outcome-based contracts are introduced to correct for the perverse incentives created by previous contracts. Faith et al. (2010) studied Philadelphia’s mental health residential services, where payments had been based on the availability of beds, which resulted in very low occupancy rates. Despite there being unmet demand for these beds, the financial incentive for providers to keep them empty had been strong. Performance-based contracting was introduced to correct this. Provider payment became based on actual occupancy, with incentive payments for high levels of occupancy and sanctions for low levels. Outcome-based payments were then added to improve patient outcomes.

**Difficulties, limitations and barriers**

The incentives created by outcome-based payments aren’t always pursued, and depend on the wider contracting environment, relationships between stakeholders, and limitations placed on the ability of service providers to respond.

The New Zealand Department of Corrections uses public performance reporting, related to a series of financial penalties and rewards, to manage the contract for the Mt Eden Corrections Facility (MECF). MECF’s new buildings opened on 30 March 2011 (Department of Corrections 2010). The quarterly reporting requirements of the Mt Eden Correctional Facility contract, awarded to Serco by New Zealand’s Department of Corrections in 2011, include 37 performance measures, 14 key performance indicators (KPIs), and a rehabilitation score. Most of the performance measures are inputs, activities or outputs (although a few, such as unnatural death, would be considered outcomes). The KPIs are mostly outcome measures, such as improved wellbeing, and a rehabilitation score (Department of Corrections 2011). Despite being the only prison ranked ‘Exceptional’ in the 12 months to March 2015 (Department of Corrections 2015), a number of violent incidents at the Mt Eden Correctional Facility led to the government assuming control of the facility in July 2015 (Gulliver 2015). This is an interesting contrast between performance reflected by a fairly extensive range of measures, public perception, and government action.

A significant barrier to the incentives created by outcome-based payments is the freedom by which they can be pursued. Outcome-based contracts represent a significant cultural shift for people who have been involved in human service delivery over time. If the contract tightly prescribes the service that must be delivered, the ability of the provider to adapt services and their resourcing in the pursuit of outcomes is curtailed. Tomkinson (2015) explored the example of the Rikers Island Social Impact Bond, contracted by New York City government. The service was defined by a rigid service
model that had proven effective on other populations of juveniles in detention. Payments from the New York City government would occur, however, from the reduction in reoffending by young men exiting Rikers Island. Although the service providers recognised that achieving a reduction in reoffending would require services outside the scope of the program that had been funded, they were restricted from delivering these services by their contracts. In this example, the primary purpose of the program as it was defined in the contracts was to test a particular service model, while the payment related to a different purpose – that of reducing the reoffending of the young men incarcerated on Rikers Island.

The specific case of social impact bonds

Social impact bonds have two distinct features: (1) an outcome-based contract between a service commissioner (usually government); and (2) private investment to cover working capital until outcome-based payments start to flow. Some of these outcome-based contracts have involved stepped payments, where a greater improvement in some measure of outcome reaps higher payments. These contracts have all, however, been implemented in service gaps. This means their effect is compared to no service, rather than to services delivered under traditional contracts or grants; their effect compared to another type of contract not being measured would be very difficult to establish. An example of the evidence as to the effect of these contracts is that in comparison to a matched group of ex-offenders who received no additional services, the ex-offenders eligible to participate in the Peterborough Social Impact Bond were reconvicted 8.4 per cent less (Ministry of Justice 2014).

Two social benefit bonds, known elsewhere as social impact bonds, were implemented in New South Wales (NSW) in 2013. They are based on outcome-based contracts between the NSW Government and not-for-profit service providers, where a proportion of payments is dependent on outcomes. Some funding is provided by external investors, whose returns are also dependent on outcomes. Payments are based on the proportion of children in care returned to their families or remaining with their families (NSW Treasury 2014). Outcome-based contracting was a new experience for the NSW Department of Family and Children’s Services, and also for the two not-for-profit providers, UnitingCare and The Benevolent Society (Tomkinson 2012). The difficulty of defining a suitable measure of outcome was evident in the amending of the UnitingCare contracts less than two years into the program, to recognise that some children restored to their families will return to state care (Tomkinson 2015). Once again, the effect of these contracts is measured in comparison to families that did not receive services at all, so cannot be conclusively attributed to the payment basis.

Some social impact bonds, for example those contracted by the Department of Work and Pensions, do not measure effect size at all. They make outcome payments without comparison to an estimate of what might have happened for clients in the absence of the service.

Further research

Further research should produce an evidence base that can be called upon by policymakers to inform their decision-making. It will also need the cooperation of policymakers, as quality research will most effectively be produced if it is designed along with new contracts and services. The contribution of more robust data in the form of
randomised controlled trials could be delivered in the process of rolling out future policy or contracting changes (Haynes et al. 2013).

A robust evidence base would include studies that focus on the effect of changing the basis of payments for outcome-based contracts. This would compare contracts that have different payment bases for a similar service. Comparison could be pre- and post- a particular change, or across areas as outcome-based contracts are rolled out across a jurisdiction. As well as studies comparing different payment bases, studies could compare the outcomes of entirely different payment models, for example, block funding or grants.

In studies where the comparison is a service delivered by the public sector, the literature has not separated the effects of outcome-based contracting from the effects of privatisation and market-based approaches. The literature is interesting where the evidence contradicts beliefs about market-based approaches – for example, where outcomes-based contracting led to greater collaboration. Further research might compare services delivered by the public sector with those delivered under both outcome-based contracts and more traditional contracts. Real value will be produced if studies are able to isolate the effects of different contracting variables.

The comparability of studies would be greatly increased if the variation between them were reduced. The replication of studies across different contracts and services would be highly valuable. The inconsistency of findings across countries could be further investigated, in an attempt to isolate the variables that are causing differences. The identification of specific practices in different countries, and their effects, would help policy-makers to deliberately introduce features of contracts that have desirable outcomes.

As the number of studies increases, one approach (for example) could be to categorise welfare/human services in terms of how they meet user needs: primarily for maintenance (e.g. aged care), for long-term development (e.g. increasing employability of long-term unemployed people), or for short-term development (e.g. getting a job, passing an exam), as well as being framed to regulate users. A key issue for policy and research would then be how the requirements or challenges for successful outcome-based contracting differ in each of these cases.

A further area for exploration is comparing outcomes between outcomes-based contracting and relational contracting (where a relationship of trust forms the basis of decisions) or alliance contracting (where unanimous decisions are made by all parties as the contract progresses).

Longitudinal analyses of outcomes would allow for understanding of the permanence of outcome changes, which would inform the understanding of cost-effectiveness of different interventions, intervention points, and outcome-based contracts (Gordon and Heinrich 2014).

Finally, empirical evidence on New Zealand examples is ripe for future studies. These would be strengthened if they were constructed in parallel with similar studies in Australia and other countries.

**Conclusion**

The main finding of this review is that there is a severe lack of evidence for policy-makers to draw on when constructing outcome-based contracts. The studies that have been done are not comparable, and largely fail to isolate contracting variables, so it is difficult to identify cause and effect. Despite the scarcity of studies containing empirical evidence, several conclusions can be drawn.
The literature reflects that outcome-based contracts have a longer history and are more widely used in some human services than others; 12 of the 19 studies reviewed relate to employment. There is, however, no evidence that employment is any better or worse suited to outcome-based contracting than any other service area.

The studies discussed suggest that outcome-based contracting can help achieve specific, observable changes, but is less effective for non-contracted outcomes and those that are difficult to define. Also, there is evidence that a strong focus on certain outcomes or outputs may work against achieving other (sometimes more important) outcomes.

Outcome-based contracts were shown to deliver improvements in the metrics by which they were paid, but these metrics proved difficult to align with desired outcomes. The challenge for government has now become to define payment metrics that represent the outcomes they seek, and that encourage behaviour from service-delivery organisations that is consistent with these outcomes. Despite a raft of hopeful promises written into contract revisions and announcements, outcome-based contracts for employment services have not resulted in increased flexibility and innovation, the tailoring of services to the individual, or improved outcomes for participants.

The evidence suggests that the rhetoric of outcome-based contracting does not always reflect its reality. While this may be true for many other aspects of policy-making and implementation, in outcome-based contracting there is a tendency towards a simplistic representation of human services outcomes that exacerbates the divide. As problems are uncovered and corrected, the contracts themselves become complex and unwieldy. These problems may reflect a lack of capacity in the public sector to build and define outcome-based contracts for service, as well as a lack of capacity within delivery organisations to negotiate and deliver them. Both of these areas of capacity need to be taken into account by public sector managers charged with the implementation of outcome-based contracts.

As well as a lack of capacity for designing and implementing outcome-based contracts, there appears to be an inadequate investment of time and skill spent rigorously thinking through the intended and unintended consequences of new contracts. This may be partly due to the additional time the outcome-based contracting process requires, and the pressure that puts on policy-makers. Policy-makers are not always open to the input of those outside government, or may not be willing or able to adequately consider it. In employment services, for example, many of the problems arising were predicted by external organisations, but dismissed during the contracting process.

The studies raise the question of whether it is beneficial to include more than one type of payment in any one contract, or whether perverse incentives might be better managed by separating them into different contracts. This includes contracts that involve more than one outcome-related payment, as well as contracts that mix outcome-based payments and service fees. This separation may solve the problem, or it may just push conflicts out to a system level. The evidence also showed that while there is a trend towards producing life-changing outcomes from public human services, sometimes people seek a simple one-off service. The use of outcome-based payments and service fees should reflect the service that is being sought.

Finally, service providers need to have the freedom to respond to the incentives created by outcome-based contracts. The achievement of outcomes can be severely constrained by the contracting environment or by prescriptive conditions in the contract that work against the intended outcomes.
References


Appendix 1: Search terms

Research databases:
- Web of Science
- Google Scholar

The keywords searched for in the above databases include:
- Outcome* and contract* in title
- Performance and based and contract and Australia
- ‘Payment by results’ and ‘welfare’
- Job* service* Australia* in title
- Outcome based contract*
- Performance based contract*
- DEEWR
- Evaluation Job Network Australia [Google Scholar]
- Evaluation Job Services Australia [Google Scholar]
- Outcome-based contracting
- Payment by results
- Outcome based contract* NZ [Google Scholar]
- McBeath and Meezan in author
- Considine in author
- All of the above generic search terms and New Zealand

The search of the above databases also looked for literature citing:
## Appendix 2

### Table 1: Summary of evidence by outcome and policy domain

<table>
<thead>
<tr>
<th>Study</th>
<th>Study focus</th>
<th>Effect on</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gates et al. 2005</td>
<td>Provider hours</td>
<td>• Getting a job</td>
<td>Outcomes funding was associated with successful employment for people with psychiatric disabilities. The performance-based funding structure may encourage providers to work more intensively with consumers early in the placement process, find jobs that consumers want to keep, and coordinate more effectively with the state rehabilitation agency.</td>
</tr>
<tr>
<td></td>
<td>Harder to serve</td>
<td>• Time to employment</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Job retention</td>
<td></td>
</tr>
<tr>
<td>Koning and Heinrich 2013</td>
<td>Full vs partial vs no performance payments</td>
<td>• Job placement rate</td>
<td>Although job placement rates for some workers increase with performance payments, job duration was unaffected. In addition, job placement rates for those with poorer employment prospects did not increase under high-powered incentive contracts.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Administrative cost</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Parking - underserving or ignoring clients</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Choice of client group</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Job duration</td>
<td></td>
</tr>
<tr>
<td>Considine et al. 2014</td>
<td>profit-seeking behaviour by frontline staff</td>
<td>Rate of jobseekers getting jobs</td>
<td>The increase in profit-seeking behaviours in Australia's not-for-profit employment services has not led to improved employment rates for clients.</td>
</tr>
<tr>
<td>Lu 2014</td>
<td>OBC</td>
<td>• Time to placement in job</td>
<td>The performance-based contract performed much better on measured performance, with a higher chance of employment and a shorter time-to-placement. But the differences in unmeasured performance between the two models in terms of working hours and wages were trivial.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Likelihood of placement in job</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Working hours</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Wages</td>
<td></td>
</tr>
<tr>
<td>Gordon and Heinrich 2004</td>
<td>Comparison of four models of analysis: the linear selection on observables, first difference, difference-in-differences, and multilevel models.</td>
<td>Apparent wage differences of mothers on different programs</td>
<td>Multilevel modelling captures both the jump (approximately $1,800, $1,000–$1,200 higher than comparison group) and subsequent decline in earnings’ for one program's mothers (Just over $100 each quarter), but the other econometric models do not.</td>
</tr>
</tbody>
</table>
### Employment: Effect on service delivery organisations and their staff

<table>
<thead>
<tr>
<th>Researcher(s)</th>
<th>Reform</th>
<th>Metric</th>
<th>Practices</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dias and Maynard-Moody 2007</td>
<td>Payment metric</td>
<td>Frontline service delivery</td>
<td>The payment for job placement resulted in management pushing a 'work-first' approach, so frontline workers felt that they couldn't properly help clients to get good jobs or address employment barriers. Frontline workers and jobseekers felt the service provided exacerbated the problems faced by jobseekers and produced poor outcomes.</td>
<td></td>
</tr>
<tr>
<td>Considine et al. 2011</td>
<td>Privatisation and OBC</td>
<td>Flexibility, Innovation</td>
<td>Privatisation of service delivery has resulted in less flexibility and an increase in standardised practices i.e. less innovation/tailoring services to the jobseeker.</td>
<td></td>
</tr>
<tr>
<td>Considine and Lewis 2010</td>
<td>Country and time (1998 to 2008), classified by governance types</td>
<td>Frontline service delivery practices: (Female staff (%), Staff under 34 (%), Union members, Caseload/FTE, Number of jobseekers seen per day, Time spend with jobseekers, Work satisfaction)</td>
<td>Employment services staff were happier in Neth and UK than in Aust. The proportion of female and younger staff has increased significantly in Australia over the first ten years of outcome-based contracting, contrary to the UK and Netherlands. At the same time, the number of jobseekers seen per day increased in Australia but decreased in the other two countries.</td>
<td></td>
</tr>
</tbody>
</table>

### Child welfare: Effect on children

<table>
<thead>
<tr>
<th>Researcher(s)</th>
<th>Reform</th>
<th>Metric</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>McBeath and Meezan 2009</td>
<td>Reform towards outcome-based contracts</td>
<td>Reunification rates, Permanent placement rates</td>
<td>Pilot children are 18% less likely to be reunified with birth parents, but 24% more likely to achieve permanent placement by the end of the study. Reunification required more effort than other permanency options. Payment was for placement (allocation to a home and carer(s)), with no extra for reunification.</td>
</tr>
</tbody>
</table>

### Child welfare: Effect on service delivery organisations and their staff

<table>
<thead>
<tr>
<th>Researcher(s)</th>
<th>Reform</th>
<th>Metric</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Collins-Camargo et al. 2011</td>
<td>Reform towards outcome-based contracts</td>
<td>Service delivery organisational practices</td>
<td>In an environment of privatisation and performance-based contracts, service delivery organisations focus more on interorganisational partnerships; shared decision making; new/increased internal communication; developing goals, objectives and measures of performance; integrated research, data and case management systems.</td>
</tr>
<tr>
<td>Meezan and McBeath 2011</td>
<td>Reform towards outcome-based contracts</td>
<td>Service delivery organisational practices</td>
<td>When moving to outcome-based contracts, not-for-profits altered (or expected to alter) staffing patterns and staff assignments, staff training, and the ways in which budgeting and financial management were handled. In some agencies, the pilot (outcomes) contracting environment necessitated (or was expected to necessitate) greater interdepartmental coordination of shared tasks between service delivery, finance, and IT systems.</td>
</tr>
<tr>
<td>Study</td>
<td>Policy Area</td>
<td>Outcomes/Variables</td>
<td>Findings</td>
</tr>
<tr>
<td>------------------------------</td>
<td>--------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------</td>
<td>--------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Washington et al. 2009</td>
<td>Reform towards outcome-based contracts</td>
<td>Satisfaction with extrinsic rewards, Rate of salary raises, Expectation of continues employment, Satisfaction with benefits, Rate of promotions</td>
<td>For private sector under outcome-based contracts: Work Environment, Satisfaction with Extrinsic Rewards and rate of salary raises higher, no statistically significant difference for expectation of continued employment, or satisfaction with benefits, or rate of promotions.</td>
</tr>
<tr>
<td>Levy et al. 2012</td>
<td>Job satisfaction, organisational commitment, and conflict between work and family</td>
<td>Intention to quit of service delivery staff</td>
<td>Intention to quit was related to job satisfaction and work-family conflict, but not organisational commitment, family-work conflict, years of experience or hours per week.</td>
</tr>
</tbody>
</table>

**Other policy domains: Effect on clients**

<table>
<thead>
<tr>
<th>Study</th>
<th>Policy Area</th>
<th>Outcomes/Variables</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leonard and Graff Zivin 2005</td>
<td>Health: illness type</td>
<td>Choice of healer - between pay by outcome and fee-for-service</td>
<td>When the illness requires large amounts of effort on the part of both patients and practitioners, patients are more likely to choose a traditional healer (paying by outcomes). Patients prefer mission facilities (fee-for-service) when either the patient or the practitioner needs to exert effort to cure the illness, but not both.</td>
</tr>
<tr>
<td>Verbeeten 2008</td>
<td>Public sector internal</td>
<td>QUANTPERF: quantity or amount of work produced; attainment of unit production or service goals; efficiency of unit operations; QUALPERF: quality or accuracy of work produced; number of innovations or new ideas by the unit; reputation of ‘work excellence’; morale of unit personnel</td>
<td>Quantitative performance was related to incentives, but qualitative performance was not. There were relationships found between quantitative performance and incentives; and clear and measurable goals and performance. Relationships between other variables (inc. quality of service and incentives) were not found.</td>
</tr>
<tr>
<td>Brucker and Stewart 2011</td>
<td>Substance abuse: Introduction of performance incentives</td>
<td>Time to assessment, Time to treatment, Participation, Retention, Completion</td>
<td>The introduction of performance incentives had no effect on timeliness of access to assessments or treatment, length of stay, completion of program. Note that less than 7% of total budget was performance-based, and only 18.83% of it was claimed.</td>
</tr>
</tbody>
</table>
### Other policy domains: Effect on governments

<table>
<thead>
<tr>
<th>Source</th>
<th>Domain</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>Faith et al. 2010</td>
<td>Mental health residential services: Introduction of payment rewards and sanctions tied to bed occupancy rates</td>
<td>Occupancy rates of beds: Tying payments to occupancy rates of beds in mental health residences met objective of increasing occupancy</td>
</tr>
<tr>
<td>Office of Federal Procurement Policy 1998</td>
<td>Across US Govt: Introducing PBC for non-human services only</td>
<td>Price: Satisfaction of government agency decreased, satisfaction increased, both improved more when new providers were chosen, audits decreased and leadtime increased</td>
</tr>
<tr>
<td>Ng and Nudurupati 2010</td>
<td>Defence: Dealing with the risks and challenges of outcome-based contracts</td>
<td>Service delivery organisational (backend) practices: Found 11 factors that mitigate risks/challenges of OBC - Align expectations; Team work; Share information; Share materials; Access resources; Clear roles; Complementary skills; Empowerment; Behaviours and attitudes; Customer’s control; Delivery org's control</td>
</tr>
</tbody>
</table>