Provider Incentives in Social Protection and Health

A selection of case studies from OECD countries

CHRISTIAN VAN STOLK
GARDAR BJORNSSON
SIMO GOSHEV

WR-779-WB
August 2010
Prepared for the World Bank
Preface

The Human Development Chief Economist’s Office (the Office) at the World Bank is supporting analytical work on the linkages between governance, service delivery and human development outcomes. The Office commissioned RAND Europe to look more closely at the use of provider incentives in the health and social protection policy areas in Organisation for Economic Cooperation and Development (OECD) countries.

This work consisted of an annotated bibliography (see Appendices D and E) and an overview report contained in these pages. The study looks at financial and some non-financial incentives (education and public release of information) in healthcare and social protection in order to understand their effects and the context in they are used in OECD countries. The study carefully considered evidence and where possible commented on the limitation of the approach chosen in evaluating particular effects. The literature is quite technical and Appendix A provides a glossary to help the reader. Drawing lessons from OECD experiences with provider incentives can prove useful for low and middle income countries which are looking to strengthen accountability and the quality of service delivery.

This report is a discussion paper aimed at the World Bank and client country officials. Its aim is to foster debate on the use of provider incentives in human development in different country contexts.

RAND Europe is an independent not-for-profit policy research organisation that aims to improve policy and decision making in the public interest. RAND Europe’s clients include European governments, institutions, NGOs and firms with a need for independent and multidisciplinary analysis. This report has been peer-reviewed in accordance with RAND’s quality assurance standards. We would like to thank Professor Martin Roland and Dr Ellen Nolte who contributed in quality assurance roles. Dena Ringold, Petronella Vergeer, and Alaka Holla at the World Bank also provided useful feedback on earlier drafts. Deirdre Culley helped edit the document.

For more information about RAND Europe or this document, please contact

Dr. Christian van Stolk
RAND Europe
Westbrook Centre, Milton Road
Cambridge CB4 1YG
United Kingdom
Tel. +44 (1223) 353 329
stolk@rand.org
# Contents

Preface ........................................................................................................................ iii
Table of figures and tables .......................................................................................... vii
Summary ..................................................................................................................... ix

## CHAPTER 1

**Introduction** ..................................................................................... 1

1.1 Background ....................................................................................................... 1

1.2 Objective ........................................................................................................... 2

1.3 Approach ........................................................................................................... 2

1.4 Structure of the report ....................................................................................... 5

## CHAPTER 2

**Provider incentives in social protection** ............................................ 7

2.1 Introduction ...................................................................................................... 7

2.2 Defining P4P .................................................................................................... 8

2.2.1 P4P in employment provision typically allows for competitive provision of job placement services ....................................................... 8

2.2.2 There are differences between P4P schemes ................................. 9

2.3 Intended and unintended effects ............................................................... 10

2.3.1 Incentives trickle down and seem to drive performance ...... 10

2.3.2 There is consensus that efficiency increases, but effectiveness does not ................................................................. 12

2.3.3 P4P may cause perverse incentives to cream and park jobseekers ...... 13

2.3.4 Service delivery is less guided by rules, but services become more generic and goals change ............................................................. 14

2.4 Implementation lessons ................................................................................... 14

2.4.1 The extent of incentivisation can affect market entry ......................... 14

2.4.2 Unincentivised outcome measures may decrease ..................... 15

2.4.3 Information systems can be costly because of difficulties in linking outcomes to incentives and gaming ........................................ 15

2.5 Conclusions ..................................................................................................... 16

2.5.1 P4P tends to improve efficiency but unintended consequences need to be managed ............................................................. 16

2.5.2 There are some lessons from the OECD experience when applying provider incentives in different contexts, including low- and middle-income countries ........................................ 17
CHAPTER 3  

Provider incentives in healthcare .......................................................... 18

3.1 Introduction .......................................................................................... 18

3.2 Most common payment schemes in healthcare ...................................... 19

3.2.1 Fee-for-service, salary, and mixed payment are ways to reimburse physicians .......................................................... 20

3.2.2 Blended schemes are argued to be most appropriate to produce outcomes in an effective way .................................................. 20

3.2.3 Efficiency and quality of care are likely to be costly ........................... 22

3.3 Financial incentives ............................................................................. 22

3.3.1 Bonuses and withhold reward particular provider behaviour ............ 22

3.3.2 The effectiveness of bonuses and withholds is questionable ............. 23

3.3.3 Alignment of incentives across the system is crucial for bonuses and withholds to be effective .............................................. 24

3.4 Non-financial incentives ..................................................................... 24

3.4.1 Non-financial incentives focus on less tangible performance outcomes .......................................................... 24

3.4.2 Active delivery of education and audit with feedback could be effective incentives .......................................................... 25

3.4.3 Multifaceted incentives are important to change to performance of providers .......................................................... 25

3.5 Mixed schemes of incentivisation .......................................................... 26

3.5.1 Pay-for Performance aims to reward providers who meet certain performance expectations with regard to quality of care or efficiency of care .......................................................... 26

3.5.2 Comprehensive P4P schemes have varying effectiveness ................ 26

3.5.3 Complex incentives systems require careful development and substantial resources .......................................................... 28

3.6 Conclusions ......................................................................................... 30

3.6.1 Multi-faceted incentivisation is likely to provide better outcomes in terms of quality and efficiency of care ............................... 30

3.6.2 When considering introducing incentive schemes to different contexts including middle-and low-income countries, it is important to consider that designing incentive schemes in OECD countries requires significant resources .......................................................... 30

REFERENCES .............................................................................................. 33

Reference List ............................................................................................. 35

Appendix A: Glossary .................................................................................. 41

Appendix B: P4P examples in employment provision: Australia and the Netherlands .......................................................... 43

Appendix C: Some examples of studies on provider incentives used in healthcare .......................................................... 47

Appendix D: Annotated bibliography on provider incentives in social protection .......................................................... 51

Appendix E: Annotated bibliography on provider incentives in healthcare .......................................................... 63
Table of figures and tables

Figure 1: indicative keyword searches in social protection ................................................. 3
Figure 2: Standard institutional design of outsourced employment service provision........ 9
Figure 3: Measures to counteract creaming and parking in employment services ............ 14
Figure 4: Comprehensive overview of the Australian system (EC2) ................................. 45

Table 1: Provider incentives covered in current review ................................................. 5
Table 2: Contract types with different degrees of incentive in employment services ....... 10
Table 3: Summary of the types of incentives reviewed in healthcare ............................... 19
Table 4: Overview of cases to illustrate the differences in the use of P4P in employment services .................................................................................... 43
Table 5: Fee structure of the Job Network (EC 2) ........................................................... 44
Summary

Provider incentives aim to improve the efficiency and effectiveness of service delivery
Provider incentives in both healthcare and social policy are used by a number of OECD countries with the aim to improve the efficiency and effectiveness of service delivery. Incentives are used to address particular market failures that arise from operating a complex system consisting of payers, providers, and users of services. In healthcare, these include: moral hazard, where users can have higher than optimal demand for services; asymmetry of information, which may lead to suppliers offering more services than optimal; and double agency, which refers to providers having contractual relationships with payers and users that could imply contradictory priorities. In social protection, incentives are mainly used to offer competition to the previously monopolistic and generic provision of social services by the government.

The review looked at financial, non-financial incentives, and mixed schemes
Provider incentives can comprise a wide range of mechanisms. This report limits the discussion to: financial incentives; non-financial incentives such as public release of information and some education interventions; and mixed schemes. As such, the report does not look at the use of performance management frameworks in improving performance.

Financial incentives can consist of a variety of mechanisms including reimbursement schemes and reward schemes such as bonus payments and withholds associated with performance. Finally, the non-financial incentives category consists of a range of mechanisms aimed at providing information to the provider on how to improve performance. Some are enforceable such as audit with feedback, while others focus on the provision of training and guidelines. The incentive in these cases tends to be reputational or disciplinary (eg not attending a training course or following guidelines may limit ability to practice in case of healthcare workers). Pay for performance (P4P) schemes are often classified as financial incentives schemes but in healthcare tend to be comprehensive multi-faceted schemes that also encompass non-financial incentives.

There are marked differences between the use of incentives in healthcare and social protection. The use of incentives in healthcare is established and well-documented, and studies refer to the use of a full range of incentives. The use of incentives in social protection on the other hand is emergent. The main area where incentives have been used in social policy is employment policy, where P4P schemes are favoured by policymakers. These schemes are often limited in scale or have only recently been introduced.
It is important to reflect on the quality of the evidence
This review fell short of a systematic review of the literature. Sources were mostly identified through expert knowledge and keyword searches. Care was taken to include only relevant evaluations and academic studies on the use of incentives in this review. As such, the review mostly excluded theoretical papers unless they were seen as influential in the two policy areas. There were a number of issues with the papers and evaluations included, which impacts on how the findings can be used and the comparability of findings:

- The number of papers P4P in social protection was limited;
- The report mostly relied on meta-reviews on provider incentives in healthcare limiting what we can say about the quality of evidence and effects in individual cases;
- Some of the evaluations included on employment policy were undertaken by government rather than independent assessors;
- Studies often lack a control group;
- Many studies use different measures for effect, which are not always comparable across papers;
- Studies do not always control for the same systemic or exogenous factors in arriving at effect; and
- There is a general lack of studies looking at patient outcomes associated with the use of incentives in healthcare.

The evidence indicates that the effect of incentives is often more limited than policymakers may expect
The use of provider incentives appears promising in theory but the overall effect of such incentives on long-term improvements in service delivery and human development outcomes appears limited. In P4P programmes in employment services, the evidence indicates that there is some effect on the efficiency of service provision, and less so on effectiveness. In healthcare, a reviewer for this study noted that the effect of individual incentives is often overstated. There is no ‘magic bullet’ and the largest effect is associated with sustained, multiple and multi-faceted interventions.

There are a number of reasons for this:

- The use of incentives may be associated with unintended consequences. An example is the unintended outcome of the use of P4P in employment provision in Australia and the Netherlands. Providers who had previously supplied distinct services started merging and producing more similar services, maximising profits but limiting the chances for innovative service delivery in the system. The findings suggest that employment provision in these countries is converging to the type that the incentives attempted to change. Further unintended outcomes of P4P schemes in employment services are creaming, where providers prefer easy to place jobseekers, and parking, where providers do not provide services to those that are hard to place. In health, there is little evidence of unintended outcomes, mostly because such outcomes are not typically observed or measured in studies.
• **The use of incentives is often contextual**, meaning there are questions around the scaleability of the use of incentives. Incentives are often used in specific locations or within specific systemic contexts. An example is the use of P4P in employment services in deprived areas in the UK. These schemes were often targeted at specific communities that received a lot of national attention, where well resourced providers could be found. In healthcare, positive effects associated with incentives are often linked to specific geographic locations, specific systemic features, and providers with specific characteristics. This is as much a question of design of the scheme as effect, but nonetheless has consequences for what one can say about the use of provider incentives more widely.

• **The effects measured reflect on quality of service or the number of services provided rather than long-term or wider human capital outcomes** related to health or employment. This point reflects on the difficulty and cost of measuring effect. In healthcare, few studies focused on patient outcomes (eg on bonus and withholds), but rather looked at quality of service and the number of services provided. In employment provision, most evaluations did not consider the long-term employment of jobseekers past two years when most programmes stop monitoring. Moreover, most studies focus on measurable outcomes rather than a wider range of outcomes. Work in employment services showed significant unintended outcomes around schooling, training, and job quality when wider outcomes were considered. As such, little is

• **The effect of incentives is often not continuous or sustainable over time.** In the provision of employment services in Australia and the Netherlands, it was noted that the P4P incentives improved efficiency but that quality of service delivery dropped over time. Similarly, the use of P4P in UK primary care suggests that once targets are met the quality of care decreases as well as the continuity of care more generally. Furthermore, studies comment that bundles of non-financial incentives provide an inconsistent effect across cases over time.

**It is clear that some incentives work better than others**
The findings above might appear discouraging. Nonetheless, there is some good evidence that incentives can and may work. In P4P in employment services, incentives appear to trickle down organisations and affect staff performance. Positive effects are noted in several schemes ranging from positive impact on job outcomes, efficiency savings, and the wider spread of performance culture in employment services providers. In healthcare, there is evidence that financial and non-financial incentives can work. In particular, multifaceted incentivisation consisting of a bundle of financial and non-financial incentives is likely to provide better outcomes in terms of quality and efficiency of care. Incentives appear to work best when they are:

• part of a blend of incentives;

• context-specific;

• crowd-in intrinsic motivations;

• sustained over time; and
responsive to clients’ changing needs and providers’ motives.

**Learning from how incentives have been used is important in designing new incentive schemes**

There is no easy answer to how to build an incentive scheme in a complex environment. However, lessons can be drawn from examples in this report. The example of P4P in employment services provision and healthcare shows that implementing a comprehensive incentive scheme requires administrative capacity and resources. In employment services, this refers mostly to the capability of monitoring the implementation of the scheme looking for unintended outcomes and correcting these in the following procurement rounds. In healthcare, the use of multi-faceted schemes appears to be effective.

The studies also point to the importance of setting the incentive at the right level, ensuring payments are in line with outcomes, and aligning incentives across the system over a sustained period of time. In P4P in employment services, there are several lessons for designers of incentive schemes. First, a P4P scheme solely based on outcome measures reduces competition and leads to more generic service provision. Furthermore, cases need to be assigned randomly to prevent creaming, and payments should be different depending on the type of outcome achieved (e.g., more payment for sustained employment of a hard-to-place jobseeker). All outcomes that one wants to achieve need to be incentivised. In terms of aligning incentives, studies on the use of withholds in healthcare show that when designing an incentive scheme it is important to ensure that performance bonuses are complementary and proportionate to salaries.

It is also important to understand the basic prerequisites for implementation. In healthcare, some incentives are unlikely to work without truly competitive markets. The existence of potential competitive providers is also relevant in introducing P4P in employment provision services.

**There are some lessons from the OECD experience when applying provider incentives in different contexts, including low- and middle-income countries**

The main considerations for introducing provider incentives in employment services when moving from the OECD context to middle-income and in particular low-income countries are that: often incentive schemes require the existence or creation of a market, in particular the existence of various (competitive) providers of employment and healthcare services; and schemes require significant administrative capacity and resources in the public sector to run and monitor the scheme. Without the former, it would be difficult to introduce competition in service delivery in employment services. Without the latter, our examples show that adequate monitoring is difficult, which impacts how outcomes related to payment are tracked and unintended outcomes can be addressed. In many cases, those involved in managing incentive schemes have invested heavily in information systems and monitoring and evaluation activities.
1.1 Background

Provider incentives are increasingly used by a number of OECD countries with the aim to improve the efficiency and effectiveness of service delivery. In terms of human development policy, they are most prominently used in healthcare and education, and to a lesser extent in social protection. This report covers the use of incentives in healthcare and social protection. However, the use of incentives in social protection in OECD relates mostly to employment service provision.

Incentives are typically used to address particular market failures that arise from operating a complex system consisting of payers, providers, and service users. In healthcare, these include: moral hazard, where users tend to have a higher than optimal demand for services; asymmetry of information, which may lead to suppliers offering more services than optimal; and double agency, which refers to providers having contractual relationships with payers and users that could imply contradictory priorities. In social protection, provider incentives are mainly used to introduce competition to the previously monopolistic provision of social services by the government.

Provider incentives can comprise a wide range of mechanisms aimed at changing or driving behaviour. This report limits the discussion to: financial incentives; non-financial incentives such as public release of information and education; and mixed schemes. As such, the report does not look at the use of performance management frameworks in improving performance.

Financial incentives can consist of a variety of mechanisms including reimbursement schemes and reward schemes such as bonus payments and withholds associated with performance. The non-financial incentives category consists of a range of mechanisms to provide information to the provider on how to improve performance. Some are enforceable such as audit with feedback, while others focus on the provision of training and guidelines. The incentive in these cases tends to be reputational or disciplinary (e.g. not attending a training course or following guidelines may limit ability to practice in case of healthcare workers). Pay for performance (P4P) schemes are often classified as financial incentives schemes but in healthcare tend to be comprehensive multi-faceted schemes that also encompass non-financial incentives.

Incentives can work on various levels. They can provide a reward for performance, coerce an individual or organisation to perform in a certain way or facilitate an act of altruism. The latter two motivations present the more traditional model of service delivery in the
public sector based on regulation, monitoring and the public service ethos. The former is more closely associated with the reform of service delivery.

Some major questions surround the use of incentives in service delivery, to which this report aims to contribute.

- Do incentives work to improve the effectiveness and efficiency of service delivery?
- Do incentives improve human capital outcomes?
- In which contexts and circumstances do incentive schemes work?
- How can designers of incentive schemes guard against sometimes considerable unintended outcomes?

1.2 **Objective**

The objective of this paper is to synthesise the international experience with provider incentives in the areas of health care and social protection, including employment services. The work will primarily focus on the OECD experience and will aim to categorise different approaches and identify effects from implementation and lessons learnt relevant for developing countries.

1.3 **Approach**

The approach taken in this piece of work is a literature review of a range of provider incentives outlined above in each of the two policy areas.

This review was not systematic or exhaustive and followed a basic four stage approach:

- Firstly, it relied on expert judgment to arrive at a suitable starting point on what incentives and what areas to review.
- Secondly, it identified the relevant literature either using expert knowledge or keyword searches in academic databases.
- Thirdly, it used broad criteria to include or exclude studies.
- Finally, the study used a broad template to capture the key aspects of each study (See Appendices D and E).

In social protection, previous work undertaken at RAND Europe (van Stolk et al., 2006; NAO 2006) indicated a limited use of incentives in social protection and guided the research to look mainly at financial incentives in the provision of employment services (see Table 1), which several countries started using in the 1990s. The most notable examples are Australia, the Netherlands, the United Kingdom and the United States. Thus, the review focused on these countries.

Previous knowledge also helped identify a number of key evaluations to be considered for inclusion in the review. To increase the number of studies for consideration, the reviewers looked at references in the initial literature and also undertook a number of keyword searches (see Figure 1 for different potential permutations). These searches took place
using the RAND Library and GoogleScholar. Finally, the reviewers looked at references in the studies identified through the keyword searches to identify further studies.

These searches identified a range of material ranging from peer reviewed journal articles, reviews of reviews, books, various evaluations by government and non-government bodies as well as a number of discussion papers.

In terms of inclusion, the reviewers selected every empirical study and review of reviews identified. Theoretical papers and most discussion papers were excluded unless they were seminal papers in the field. An example is the work of Bruttel (2004). The reviewers did not: exclude studies on the basis of research design; explicitly score the quality of each study; and keep track of the number of studies excluded. This course of action seemed appropriate given the limited number of studies available and the scope of this report. Rather, the reviewers indicate in the annotated bibliography included in Appendix D the type of study, the type of approach taken, countries included and effects looked at, from which readers can make some inferences on quality. The reviewers put each study in one of the following categories: case study; empirical or applied study; review of case studies; review of empirical studies; review of reviews; and theoretical study (discussion paper). The template of the bibliography captures reference details, year of publication, type of publication, place, type of incentive, effect of incentives, and wider comments relevant to understanding the context of implementation. Readers can find more information on the studies there.

Figure 1: indicative keyword searches in social protection

In healthcare, there is a substantial body of literature on a wide range of provider incentives in healthcare. To identify relevant studies, the team initially consulted RAND Europe experts on provider incentives in healthcare. They provided guidance on the subject, and formulated suggestions on the approach to the issue of incentivisation in healthcare and relevant sources of information. As a result, the team decided to focus on particular provider incentives: financial incentives; certain non-financial incentives; and mixed schemes (see Table 1). The effects of incentivisation at different levels of healthcare provision such as prevention, primary and secondary care (including prescribing) are also captured in the selection. Further studies were identified by contacting those who have
commissioned work on the topic and requesting their assistance in identifying further published and grey literature; and by scanning and selecting references from the lists of references in every study included. The team did not undertake keyword searches.

The team included every review of reviews identified, a number of single empirical studies, a small number of theoretical studies, and papers provided by the client. As before, the reviewers did not: exclude studies on the basis of research design; explicitly score the quality of each study; and keep track of the number of studies excluded. This seemed appropriate given the scope of the study and the reliance in this review on meta-reviews, which often do not contain information on the quality of underlying research. Where possible, the reviewers make some inferences on the quality of the paper/study/meta-review. The annotated bibliography was structured in the same way as the one developed for social protection (see Appendix E). The reviewers used the same categorisation for studies included introduced for social protection, ranging from case study to theoretical study.

The reliance on meta-reviews (of other reviews of case studies) in this report also means that it is sometimes difficult to present a specific effect or give information on specific countries in which that effect is measured. This is particularly true for provider incentives in health.

There are a number of common observations on the studies on provider incentives in social protection and healthcare included in the review that readers should keep in mind when reading this report:

- The number of papers P4P in social protection was limited;
- The report mostly relied on meta-reviews on provider incentives in healthcare limiting what we can say about the quality of evidence and the effect in individual cases;
- Some of the evaluations included on employment policy were undertaken by government rather than independent assessors;
- Studies often lack a control group;
- Many studies use different measures for effect, which are not always comparable across papers;
- Studies do not always control for the same systemic or exogenous factors in arriving at effect; and
- There is a general lack of studies looking at patient outcomes associated with the use of incentives in healthcare.

The specific type of provider incentives covered in the review and presented as general categories in Table 1 will be further elaborated on in Chapters 2-3.
<table>
<thead>
<tr>
<th>Provider incentives covered in current review</th>
<th>Social protection</th>
<th>Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial incentives</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Non-financial incentives (education)</td>
<td>–</td>
<td>X</td>
</tr>
<tr>
<td>Mixed schemes</td>
<td>–</td>
<td>X</td>
</tr>
</tbody>
</table>

1.4 **Structure of the report**

This report has two substantive sections. The first chapter deals with provider incentives in employment service provision and the second chapter focuses on the use of provider incentives in healthcare. The chapters are structured as follows. They explain what each incentive is; give the effects of the incentives; look at wider implementation lessons; draw conclusions for the wider policy area; and pose some questions on applying provider incentives in middle and low income countries. This report uses a technical language. A Glossary is provided in Appendix A to help the reader with key concepts.
CHAPTER 2 Provider incentives in social protection

2.1 Introduction

The use of financial incentives in social welfare service provision is less common than in other sectors such as healthcare (see Chapter 3). Its increasing use is a recent phenomenon (Koning et al., 2010) and has been linked to the adoption of welfare to work policies in social protection in many OECD countries. Such programmes were a response to calls for reform of the delivery of social services in the 1990s. These focused on delivering greater efficiency and effectiveness in the previously monopolistic delivery of services. Aspects of such reforms included more transparency in performance management, a focus on service quality, and wider public sector reform consisting of the use of executive agencies, non-governmental organisations (the third sector) and the private sector. Canada coined this reform package as Alternative service Delivery (OAG, 1997).

One of the earliest examples of a ‘welfare to work’ programme was the Job Training and Partnership Act (JTPA), adopted in 1982 in the US. The wider welfare to work agenda gained momentum with the Personal Responsibility and Work Opportunities Act of 1996, which included the Temporary Assistance for Needy Families (TANF) programme (McConnel et al., 2003). The Act moved social policy towards ‘a welfare to work’ approach, but also provided for the outsourcing of these services. In 2001, this market had grown to $1.5 Billion or 13 per cent of all TANF expenditure (88 per cent non-profit, 12 per cent for profit) (Finn, 2007).

Following the US programme, Australia started its “Job Network” in 1998 (DEWR, 2002) and the Netherlands adopted a welfare to work reform in 2002 (Bruttel, 2004). These consisted of system-wide ‘welfare to work’ overhauls of social services, introducing incentives for employment in service delivery. In 2000, The UK adopted a smaller scale trial project in the form of “Employment Zones”, a ‘welfare to work’ programme in specific deprived areas (Griffiths et al., 2007).

These reforms were recently cited as inspiration for a coming reform of UK foreign aid. However, there is little empirical evidence on the effectiveness of provider incentives in

1 Department of Employment and Workplace Relations.
welfare to work schemes that take into account both intended and unintended consequences (Koning et al., 2010).

The focus of this chapter is therefore on incentivisation associated with a shift in social protection towards a work first approach. The chapter provides an overview of pay for performance (P4P) and a brief summary of key design features and implementation choices in Australia, the Netherlands, the United Kingdom, and the United States. The chapter then proceeds to outline the main effects of the different schemes, commenting on contextual and implementation issues where appropriate and possible. The chapter concludes with a discussion of lessons learnt.

2.2 Defining P4P

2.2.1 P4P in employment provision typically allows for competitive provision of job placement services

The actors involved in employment services are the government, service providers and the clients, in this case jobseekers. To deal with unemployment, governments can use a work to welfare programme to decrease frictional unemployment through making job-matching more efficient, and to raise the productive capacity of the unemployed. In the provision of employment services, governments are principals and service providers, agents. Providers can be government agencies or private and third sector operators.

There are two models for service provision. The traditional model involves government budgeting on the basis of costs incurred by government agencies that are involved in service provision and reimbursing costs of other private and third sector providers. In this model, there are no explicit links to outcomes (e.g. placement into employment), though service contracts might include standards of service delivery and stipulations on specific outputs (e.g. number of cases dealt with). The motivations or incentives conducive to good outcomes are intrinsic based on trust or public service ethos (I am a certified educated case manager, and my work ethos is to provide job outcomes), or coercion linked to stipulations on service provision (I might be sanctioned or lose my job, if cases are mismanaged or if procedure is not followed). These incentives rely on a public service ethos that aligns with government goals, and the power of coercion to force service providers to achieve certain outputs or standards.

The second model is a more explicit P4P model and introduces financial incentives and competition into service delivery. P4P schemes usually have two distinct client paths: one in which government operates one stop shops for screening and assigning the jobseekers as well as processing and paying benefits; the other in which private providers compete in spheres of case management and placement services. In most schemes, the government agency screens jobseekers and puts them in groups. A number of these groups are allocated (either randomly or on the basis of identified risks [disability, length of unemployment, alcoholism, etc.]) to private or third sector providers who then try to place these jobseekers in employment. These providers are typically awarded contracts under competitive tendering. These contracts award payments (mostly a function of the amount of unemployment benefit saved by a successful job placement) on the basis of certain outcomes (mostly based on a successful job placement). These typically vary between schemes. Many jobseekers still receive job advice and placement services in the public
government agency, which also retains responsibility for processing and paying unemployment benefits. The P4P scheme in all cases does not apply to the government agency, which often still provides similar employment services to private and third sector operators. The exceptions are the Netherlands and Australia. The institutional set-up is illustrated in Figure 2.

### Figure 2: Standard institutional design of outsourced employment service provision

2.2.2 **There are differences between P4P schemes**

P4P is in itself is an umbrella term for a set of practices concerned with budgeting, contracting or paying for outcomes/performance rather than costs/time. There are several common models for contracting, each focusing to a different degree on outcomes/performance. The most common models are outlined in Table 2, which differentiates between types of financial incentive, the allocation of the risk of service delivery, operational challenges, and barriers to market entry.
Table 2: Contract types with different degrees of incentive in employment services

<table>
<thead>
<tr>
<th>Type of contract</th>
<th>Financial Incentive</th>
<th>Risk</th>
<th>Operational challenges</th>
<th>Barriers to market entry</th>
</tr>
</thead>
<tbody>
<tr>
<td>Straight P4P aka “NCNP” (No Cure No Pay)</td>
<td>High</td>
<td>On provider (requires high risk-premium)</td>
<td>Sophisticated information systems and control mechanisms required</td>
<td>Can be high due to cash flow problems (outcomes occur far after investment)</td>
</tr>
<tr>
<td>Hybrid versions aka “NCLP” (No Cure Less Pay)</td>
<td>Moderate</td>
<td>Shared in different degrees (requires medium risk-premium)</td>
<td>Sophisticated information systems and control mechanisms required</td>
<td>Can be moderate due to cash flow problems</td>
</tr>
<tr>
<td>Cost reimbursement</td>
<td>Low</td>
<td>On principal (no risk-premium)</td>
<td>Low</td>
<td>Low</td>
</tr>
<tr>
<td>Fixed Price</td>
<td>Low</td>
<td>On principal (no risk-premium)</td>
<td>Low</td>
<td>Low</td>
</tr>
</tbody>
</table>

Source: Adapted from McConnel 2003

The No Cure No Pay (NCNP) is a straight P4P scheme, whereby the provider receives payment only when placing the jobseeker into employment over a period of time. The straight here refers to the absence of any cost reimbursement outside of the performance-related payment. The advantage of this system is that it places the risk in achieving the outcome squarely on the provider who demands a risk premium in return. As payment is directly linked to outcome, payment requires sophisticated verification and information systems. A No Cure Less Pay (NCLP) system gives the provider a certain amount to cover costs and a subsequent amount linked to performance. Similar to NCNP, it requires a verification and information system. A cost reimbursement scheme is principally concerned with covering the costs of the providers and has only small payments linked to performance. A fixed price scene is similar but contracts a certain number of services. Some contracts can have a P4P component.

Among these, hybrid systems are by far the largest group with straight P4P making up a minority. Cost reimbursement and fixed fee schemes are used in a number of cases, sometimes next to hybrid and straight P4P schemes. Below, the review focuses mostly on P4P and hybrid schemes. Appendix A provides further detail on the schemes in Australia, a hybrid system, and the Netherlands, a straight P4P system.

2.3 **Intended and unintended effects**

This section outlines the main findings of government and independent evaluations of provider incentives in employment service provision. There is little evidence that takes into account both intended and unintended effects of the introduction of P4P incentives (Koning et al., 2010).

2.3.1 **Incentives trickle down and seem to drive performance**

There is evidence of a P4P incentives trickling down, affecting organisations staff, teams of staff and often even clients.2 Organisations involved in P4P schemes almost always offer

---

performance bonuses to staff. However, they also foster a culture of performance. They often have weekly discussions on targets and performance goals, a strong performance focus among case workers (McConnel et al., 2003; DEWR, 2006; Hirst et al., 2006) and even practices such as ringing bells and gathering around each other’s desk to give positive feedback when a job outcome is achieved (DEWR, 2006). McConnel et al. (2003) note that there is awareness of performance in providers contracted under cost-reimbursement and fixed price schemes, but that the focus was rather on re-winning contracts rather than outcomes.

Hirst et al. (2006) carried out a qualitative study of the UK Employment Zones (EZs) and observed a pattern of implementation where performance incentives were set for individual workers in the first stage, but were changed to team level when unintended consequences such as hoarding behaviour occurred.\(^3\) They also noted that providers start incentivising process measures (eg correctly filing customer details on the system) and not only pure outcome measures (eg customer satisfaction) to avoid unintended consequences. Interviewees claimed that incentivisation had removed “comfort zones” and rigidities, reinvigorating the provision of services (Hirst et al., 2006).

Burges et al. (2004) and DEWR\(^4\) (2006) study the relationship of individual and team incentives in EZs and the Australian Job Network respectively. Burges et al. found that pay incentive (up to 7.5 per cent of wage) was effective in teams rather than individuals, raising the number of successful job placements by as much as 10 per cent. They further note that incentives followed the theoretical arguments of free-riding leading to smaller effects in large teams (where peer monitoring is difficult). Peer pressure strengthens the effects in small teams. Their findings also correspond with the theoretical prediction that strong incentives steer effort towards the most readily measurable outcomes. Even though quality of service was incentivised, there was no significant change in quality of service delivery. Burges et al. (2004) conclude that team a team-based incentive is more efficient than individual performance awards such as a wage increase.

DEWR found that in the case of Australia, goal focus itself as opposed to financial incentive drove performance. High staff awareness of goals and performance, discussions in weekly meetings, graphs and charts on walls, ringing bells and positive feedback when outcomes are achieved are found to be the most significant factors driving performance (DEWR, 2006).

Two conclusions can be drawn from the aforementioned studies:

- pay for performance incentives trickle down, often first to individuals and then teams;
- team incentives appear more effective than individual incentives; in small teams the goal focus of the organisation can be equally or more efficient than financial incentives.

\(^3\) Hoarding behaviour in this case was for example hoarding job vacancies for one’s own clients instead of sharing them

\(^4\) Department of Education and Workplace Relations, Australia
It is important in this context to point to the limitations of the data. It concerns a very specific context and is limited to two rigorous studies. There is much theoretical and empirical literature on the difference between intrinsic motivation (public service ethos, pride in one’s job etc.) and financially induced motivation, pointing towards the danger of financial incentives crowding out intrinsic motivations.

2.3.2 **There is consensus that efficiency increases, but effectiveness does not**

Numerous studies investigate the effectiveness and efficiency of P4P schemes in employment services. There are a number of weaknesses associated with these studies. Evaluations often:

- are performed by government department and therefore not necessarily independent;\(^5\)
- lack control group data; and
- take place when widespread systemic reforms are still underway.

The evidence presents a relatively consistent picture of performance incentives. They appear to increase the efficiency of existing services, though not effectiveness.

In the UK, EZs have been limited to a pilot programme, which has been evaluated in several studies. Hales et al. (2003) carried out a large group/control group study of EZs with relevant government counterparts and find an eight to ten per cent increase in job outcomes in the incentivised services. They noted that participants in these schemes however got jobs with lower qualification profiles and underlined the increased flexibility of EZs was the primary mechanism for change. Hasluck et al. (2003) carried out another large statistical investigation and confirm these findings.

Bruttel (2005b) and Griffiths et al. (2007) elaborate on these findings noting that increased outcomes mirror, and are therefore driven by, incentive structures. Griffiths et al. (2007), however, note that this increase in efficiency came at a higher price than for the control group, and that ‘value for money’ was not significantly increased.

The Australian experience of provider incentives also provides some insight. There is relatively widespread consensus that the Job Network has increased efficiency in terms of reduced cost, but that effectiveness in terms of jobseekers placed in employment has remained unchanged (Bruttel, 2004; Dockery, 2001; DEWR, 2002 and 2006; OECD, 2001; Productivity Commission, 2002; Struyven et al., 2005). However, it must be noted that the evidence provided is mainly to be found in government studies. These are based on crude outcome measures, which entail difficulty in measuring the specific effect of incentives.

There has been a noted lack of evidence of effectiveness and efficiency in the Netherlands (Bruttel, 2004; Struyven et al., 2005). However, Koenig et al. (2010) have recently presented evidence from a unique dataset on both the intended and unintended outcomes of the Dutch system. They find a 3 per cent increase in job placements, which would indicate a positive outcome from a pure effectiveness point of view. They do however find

\(^5\) This is especially the case in Australia, and until recently in the Netherlands.
that jobs placed by providers with strong financial incentives almost entirely consisted of temporary jobs. Provider incentives are less effective for clients with higher barriers to entry into work where parking and creaming is observed (Koenig et al., 2010).

These conclusions paint a relatively homogenous picture where the introductions of provider incentives have increased the efficiency of services, but have generally had little impact on effectiveness, with the important exception of the UK EZs. It is important in this context to underline the significance of unintended consequences, which will be discussed below.

2.3.3 **P4P may cause perverse incentives to cream and park jobseekers**

One of the most salient theoretical predictions is that P4P can cause unintended consequences when inputs are differentiated but incentives do not perfectly follow that differentiation. The argument is simple. If providers can choose to select low-cost/high-outcome clients in (“cream”), they will achieve better outcomes in terms of placement of jobseekers into work and under a P4P scheme profits will be maximised. Once clients are in, there is little incentive to provide support for the hard to place who can be “parked”.

There are two main problems with this outcome:

- the principle of impartiality/equity is compromised if certain citizens cannot get access to services, and;
- the outcome might not be optimal because for any given jobseeker which is “parked” and remains out of a job, externalities might far outweigh efficiency savings.

Governments in the cases studied have used two primary mechanisms to avoid the creaming and parking problem:

- mechanisms for price discrimination, applying different outcome payments depending on employment prospects of clients (to combat parking); and
- random assignment of client and an element of client choice (to combat creaming).

Mechanisms for price discrimination include the Job Seeker Classification Instrument used by Centrelink in Australia to place jobseekers on a continuum of employment prospects and increasing outcome payments. *Kansmeter/Kwint* in the Netherlands is another example where only the harder-to-place clients are passed to providers and prices are differentiated according to difficulty in placement (Bruttel, 2004).

Despite these measures, evidence indicates that parking is a problem. In Australia, parking behaviour has been noted by analysing administrative, statistical, interview and customer satisfaction data (OECD, 2001; Productivity Commission, 2002; Considine, 2003; DEWR, 2002 and 2008; Bruttel 2004).

A recent study on the Netherlands has concluded that efficiency increases are noted amongst the easier to place jobseekers. Amongst these, there is no evidence of parking behaviour. When harder to place jobseekers are studied, however, the effect of provider incentives decreases as the occurrence of parking behaviour increases. The authors suggest that parking might in fact be driving the decrease in efficiency.
There are tools available to policymakers to counteract parking, these include information and monitoring systems coupled with control mechanisms for minimum services requirements (see Figure 3). An example of minimising parking is the Wisconsin Works programme where standardised customer satisfaction services were implemented as part of the performance measurement in contractual arrangements, giving a clear indications when users did not feel they were getting adequate service (Finn, 2007).

2.3.4 Service delivery is less guided by rules, but services become more generic and goals change

In the Netherlands, the UK and Australia, key informants all perceived a far more flexible working routine in employment service provision (Bruttel, 2004). However, employment services did not become specialised and heterogeneous. Rather services become more similar as providers previously focusing on specific client groups (alcoholics; hard to place; disabled) started offering more generic services and selecting more similar clients.

This phenomenon has been observed in Australia, the Netherlands, the UK EZs and Chicago. Niche providers and NGOs have progressively converged towards the standard for-profit provision model, displaying increases in caseloads, creaming behaviour, and convergence of service delivery strategies. This has occurred to the extent that within a few years, the statistical differences between non-profit and private provision were erased (Considine, 2003).

The evidence suggests that outcome payments have driven generic, one size fits all, solutions, and created a convergence between service providers. Although there is no solid explanatory evidence, a plausible hypothesis is that random assignment of clients fosters a generic approach targeted towards the median, thus maximising outcomes in the aggregate.

2.4 Implementation lessons

2.4.1 The extent of incentivisation can affect market entry

Burgess and Ratto (2003) note a risk of trade-off that causes providers to charge higher rates to compensate for increasing risk, as payments are increasingly shifted towards

---

6 For Australia, see Bruttel, (2004); Considine, (2003); Struyven et al., (2005), DEWR, (2008) and Abello (2002). For the UK and the Netherlands, see Bruttel (2004), for Chicago see Brodkin (2005).
outcomes. This is attributed to potential exogenous factors affecting outcomes (financial crises or local corporate bankruptcy), for which contractors would seek insurance.

The literature supports this theoretical point. For example, the Dutch case shows that shifting towards a larger pay for performance component was met with price increases by providers (Bruttel 2005a).

A further study supporting this conclusion was carried out by McConnel et al. (2003) in the US. They compared the market structure of states with differing degrees of risk, shifting towards providers. They found that a high degree of incentivisation through straight P4P systems resulted in competition between few large providers, because market entry required significant liquidity to invest in the period between commencement of employment services and outcome payments (McConnel et al., 2003). This was also noted in the Australian case (Productivity Commission, 2002).

The evidence suggests that although high outcome payments do indeed incentivise, they also result in significant barriers to market entry. This trade-off needs to be carefully considered in any implementation.

2.4.2 Unincentivised outcome measures may decrease

Burgess and Ratto (2003) predict that hard to measure quality indicators and outcomes that are not incentivised drop. In studied cases, there was a decrease in schooling and training for jobseekers (Hales et al., 2003; Bruttel, 2004), as well as low job stability, short-time placements, regularity of work hours etc. (Brodkin, 2005; Koning, 2010). The Australian government specifically commented on the lack of training and schooling as causing negative externalities in society through lower productivity and sub-optimal outcomes (DEWR, 2008).

Bradegaard et al. (2007) use a qualitative research method and argue that there is a shift in incentivised service providers from a “human capital” approach, prioritising education and training, towards a “work first” approach driven by incentives. It is also noted that this shift also seems to occur at NGO level, previously associated with the “human capital” approach.

In summary, four trade-off challenges were posed in this context:

- schooling and training for jobseekers decreased;
- jobs had lower qualification profiles, wages and conditions;
- job outcomes were short-term, mirroring incentive structures; and
- there was a cultural shift in provider organisations from a “human capital” to a “work first” approach.

2.4.3 Information systems can be costly because of difficulties in linking outcomes to incentives and gaming

Information systems causing often unintended costs are a reoccurring theme in existing research. Several case studies (Heinrich et al, 2007; and Finn, 2007 on Wisconsin; McConnel et al., 2003 on seven US state programmes; OECD 2001 on Australia) point to higher cost and more difficulty in introducing information systems than initially envisaged. There are two main reasons why this might be the fact:
it is difficult to link outcome measures to the specific use of incentives, and;
• information systems need to be expanded to counter gaming by providers.

In terms of linking incentives to outcome measures, P4P implies having adequate measures of performance that allow causal linkage with incentives. There is evidence indicating that common outcome measures for employment services tend to have a number of possible explanations (see DEWR 2002, 2006 and Heinrich et al. 2007). This means that part of the observed outcomes can be explained by exogenous factors outside the control of service providers. A simple example of this is a new factory opening in a village with high long term unemployment that could trigger high outcome payments for local employment service providers, or the opposite.

This has two outcomes: 1) it can cause suboptimal incentives and inadequate provision of employment services; and 2) information systems have to be sophisticated and have high costs in order to achieve the stated aims.

The first effect is noted by the OECD in the case of Australia. It is noted that outcomes for intensive assistance services would have occurred without any intervention at all. Simultaneously, outcome payments do not reflect true societal value. The consequence of this is that the marginal gain for service providers in actually providing value added rather than just hoping for exogenous effects is too low – making incentives much more inefficient than they actually are (OECD 2001).

The key approach would be to minimise all outcome payments relating to random or exogenous effects. This enables fewer but also much larger outcome payments, making the provision of value added by providers profitable, and thus “fixing” the problem. The OECD suggests a competitive bidding mechanism over one to three months, where providers can bid on their current rate (OECD 2001). An alternative could be the expansion of Australia’s current mechanism of star ratings, which uses an econometric model to take a host of regional and local variables into account when awarding ratings. Indeed, the Productivity Commission notes that the star rating system seems to provide a large part of provider incentives though it is primarily geared towards awarding renewed contracts rather than payments (2002).

2.5 Conclusions

2.5.1 P4P tends to improve efficiency but unintended consequences need to be managed

This chapter looks at the use of performance incentives as mechanisms in social welfare provision in the OECD. It has focused on the implementation of such incentives in the US, UK, the Netherlands and Australia.

There is substantial evidence that the implementation of pay for performance entails a “trickling down” of incentives to individual case workers or teams. This financial incentivisation seems to drive efficiency when implemented at a small team/small office level, where free-riding is minimal and peer pressure operates.

On a macro level, there is consensus that P4P has increased efficiency of service provision (three to ten percent) but not of service effectiveness. This increase however has come with several unintended consequences:
• The primary unintended consequence has been extensive parking by providers, where hard-to-place jobseekers are given none or minimal service, which causes equity concerns in government policy.

• The introduction of P4P also leads to increasing homogenisation of service provision towards a single generic model of provision.

2.5.2 There are some lessons from the OECD experience when applying provider incentives in different contexts, including low- and middle-income countries

This review does not aim to be a how-to-guide for applying provider incentives in social protection to different contexts. However, the review highlights some critical lessons in the design of incentive schemes to assist policy-making:

• Provider incentives need to reflect the underlying system of service provision. Provider incentives typically target a market failure. Understanding this market failure is critical to design.

• The level of incentivisation, or degree of output payment, is important to market design. There are significant questions about how much risk is transferred to the party being incentivised and the price set for an outcome. The transfer of risk relates to whether the scheme subsidises input costs, is a straight pay for performance scheme, or does both. A price set too high or too low could result in sub-optimal service delivery.

• Finally, outcomes that are not part of the performance regime seem to decrease in incentivised P4P systems. It is therefore important to incentivise all outcomes of interest (including qualitative measures).

• P4P schemes need significant monitoring and evaluation in order to track outcomes. Information systems are expensive and complex to run. Failure to monitor outcomes means those running the scheme: have little information on the success of the scheme; find it more difficult to link payment to outcomes; are less accountable; find it hard to mitigate unintended consequences; and cannot easily learn lessons on how to improve the programme.

• Contracting and procurement skills are not necessarily found in abundance in public sector bodies. Experience shows that a high degree of specialised knowledge in contractual design issues is needed.

• Schemes need to be responsive to clients’ changing needs and providers’ changing motivations. Contracting cycles cannot be too long to allow service provision to move to the median, but cannot be too short as to entail a risk not worth considering for a potential service provider.

The main considerations for introducing provider incentives in employment services when moving from the OECD context to middle-income and in particular low-income countries are: the existence of a market, in particular the existence of various competitive providers of employment services; and the significant administrative capacity and resources required in the public sector to run and monitor the scheme.
3.1 Introduction

This chapter considers a range of incentive mechanisms, which aim to increase the efficiency of healthcare provision and improve the quality of care for patients. However, only few of these mechanisms target efficiency (driving down the cost of service provision) and quality of care explicitly simultaneously. Indeed, the majority of provider incentives aim directly at only one of these aspects while leaving the other either unattended, or assume a relationship between the two.

Incentives need to be seen in the context of healthcare systems. The provision of healthcare in OECD countries involves three main players: payers, providers and patients. Generally, providers are reimbursed by payers for the services they render to patients. Payers receive funds through various mechanisms including general taxation, social security, social health insurance and individual (private) health insurance. The organisation of healthcare financing and delivery usually separates providers from payers of care. This is considered a necessary condition for the independence of medical treatment from its cost and a guarantee that individual patients, holders (and sometimes non-holders) of health insurance, would be able to afford the care they need (McGuire, 2000).

In theory, there are three key challenges that arise as a consequence of this setting:

- Moral hazard: Where patients are not liable for the full amount of costs associated with treatment, they may demand more care than is clinically necessary. This may lead to higher utilisation of services, higher cost to payers, and may ultimately lead to waste.

- Supplier-induced demand: Information asymmetry may alter patient preferences and providers might deliver more services than are medically necessary. Such behaviour is driven mainly by the self-interest of providers (for higher income/revenue) and is likely to result in higher service utilisation, higher immediate costs to payers and ultimately waste.

- Double agency: Payers have an interest in efficient provision but as they do not provide health services directly there is little control over the quantity of services supplied. As such, payers expect providers to act on their behalf in a principal agent relationship and ensure that patients receive appropriate quantity of care at a minimum cost. However, providers are also agents of the patients who expect
them to act on their behalf and their best interest, and therefore provide best quality care.

Similar to social protection, there are different payment schemes in healthcare. Payment schemes by themselves provide incentives to providers. For instance, in social protection P4P became a payment system used alongside other types of payment, such as cost reimbursement and fixed fees. In healthcare, financial and non-financial incentives are often used to improve existing payment systems or interact with them. Comprehensive P4P schemes can be a complement to these systems and often an approach used to address some externalities of payment schemes such as double agency and supplier-induced demand. Therefore an understanding of the use of incentives requires an understanding of the underlying payment schemes and the inherent incentives they offer. This adds a degree of complexity to the use of incentives in healthcare, perhaps absent in social protection. This chapter starts by outlining the most common payment schemes in healthcare.

The incentives reviewed in this chapter could be divided into three broad categories: financial, non-financial, and mixed. A noted difference with social protection is that P4P in healthcare is seen as a mixed incentivisation scheme rather than a financial incentive. Where possible, this review offers detail of the size and timing of incentives, the setting of implementation (eg primary and secondary care, though with a focus largely on primary care7) as well as whether the incentive mechanisms are targeted at individual physicians, groups of physicians or whole organisations (eg GP practices, hospitals). The chapter also reflects on the unit of analysis used in the evaluations and studies to assess effectiveness. Table 3 provides a summary of the specific incentive schemes reviewed in this chapter. For each incentive, the review gives a brief description of key features, discusses effects and setting of implementation, and outlines key lessons learnt and implications for future applications. The chapter concludes with a summary of the findings.

Table 3: Summary of the types of incentives reviewed in healthcare

<table>
<thead>
<tr>
<th>Type of incentive</th>
<th>Specific incentives</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial</td>
<td>Bonus, stepped bonus and withhold</td>
</tr>
<tr>
<td>Non-financial</td>
<td>Education and guidelines, feedback and public reporting, accreditation and recertification, performance audit and reminders</td>
</tr>
<tr>
<td>Mixed</td>
<td>Comprehensive P4P schemes</td>
</tr>
</tbody>
</table>

3.2 Most common payment schemes in healthcare

Fee-for-service, capitation, salary, and mixed payment schemes are common methods for reimbursing physician labour. While they all achieve their primary objective of rewarding labour in healthcare, they give rise to substantially different incentives to providers and therefore are likely to entail differing consequences for payers and patients.

---

7 Payment systems in secondary/hospital care (such as budgets, per-diem and diagnosis-related groups) and the incentives associated with them have been excluded from this review as their basic design principles resemble one or more of the three basic forms of payment (FFS, capitation and salary) reviewed here (Christianson, 2007).
3.2.1 **Fee-for-service, salary, and mixed payment are ways to reimburse physicians**

Fee-for-service is a method of payment which reimburses individual physicians and other health practitioners for each service they provide to patients (Maynard, 2006). Medical procedures, tests, office and home visits are all examples of such services. The particular amount per service that physicians receive is usually determined through negotiations between payers and providers. It is possible that there are differences in fee schedules and these commonly reflect geographical or provider particularities.

Fee-for-service is commonly contrasted with capitation. Under capitation, individual physicians or groups of physicians receive a payment of pre-defined per-patient amount for all services they render over a given period (Maynard, 2006; McGuire, 2000). The exact amount of this fixed per-patient reimbursement is determined on the basis of socio-economic and other characteristics of provider’s location and their patient pool.

Salary is a payment per unit of time input by individual physicians (Maynard, 2006). It represents a fixed amount of money which is paid to physicians irrespective of the services rendered to patients or the number of patients treated. The amount is commonly negotiated between payers and providers, and may reflect geographical or other provider specificities.

Finally, mixed or ‘blended’ schemes combine two or all of the payment methods described above.

3.2.2 **Blended schemes are argued to be most appropriate to produce outcomes in an effective way**

**Fee for Service**

Fee-for-service (FFS) has been used widely in primary care and secondary outpatient care in a number of countries including Japan, Belgium, Germany and the US (Docteur and Oxley, 2003; WHO, 2007). Following from its design, FFS may weaken the payer-provider agency and somewhat strengthen the relationship between patients and providers. In the meantime, the supplier-induced demand problem may be exacerbated as providers are practically given unlimited spending allowances. Indeed, compared to other types of schemes, FFS creates incentives for higher service utilisation (Chaix-Couturier et al., 2000; Davidson et al., 1992; Gosden et al., 2000), even beyond what may be deemed as optimal (Maynard, 2006). Empirical evidence indicates that when tariff freezes were imposed in Canada (Quebec), primary care physicians increased by 6 per cent the quantity of services they provided (Chaix-Couturier et al., 2000). This clearly illustrates the effect of supplier-induced demand on service utilisation and suggests that this scheme of incentivisation may fail to deliver efficient provision of care. However, while higher utilisation may result in higher immediate cost for payers, it also may lead to lower costs in other parts of the system, such as emergency room services (Chaix-Couturier et al., 2000).

---

8 In some cases, it may be that patients have to pay to providers first, and then file a reimbursement claim to the payer. This particular mechanism of reimbursement may create significant inequalities in terms of access to care for patients who may be liquidity constrained. Albeit important, this particular issue lies beyond the scope of this work.
Imposing budget or revenue caps on FFS is a commonly adopted approach thought to induce increases in efficiency of provision. While it may strengthen insurer-physician agency and therefore is likely to result in lower total expenditure on healthcare service compared to ‘blank cheque’ FFS, it has little effect in practice (Maynard, 2006). Empirical evidence from Canada (Quebec) indicates that implementing revenue caps is associated with a modest re-distribution of patients among physicians (eg patients of higher activity physicians are reallocated to lower activity physicians) and decreases in total revenues from care (Chaix-Couturier et al., 2000).

A widely recognised downside of FFS with respect to the quality of services provided to patients is that it may induce treatment-to-the-test, where activities which are not incentivised are marginalised (Maynard, 2006). However, the degree of this particular problem is likely to vary among providers as it likely to depend primarily on the trade-off between provider preferences for patient health benefit and net income (Conrad, 2009).

**Capitation**

In contrast to FFS, capitation may strengthen provider agency with payers and weaken the relationship between providers and patients. Therefore, it may eliminate supplier-induced demand. This may come as a direct consequence of the cap on spending per patient which induces medical practitioners to reduce service utilisation (Davidson et al., 1992; Gosden et al., 1999; Maynard, 2006) even to levels below optimal. Moreover, if there is no link between the capitation payments received by primary and secondary care providers, there may be a strong incentive for providers to shift patient costs to others (Maynard, 2006). The reinforced incentive for creaming, or choosing to treat only healthier and therefore less costly patients (Ellis, 1998; Maynard, 2006), may undermine the provider-patient relationship. This is emphasised further by the lack of empirical support to theoretical constructs suggesting that capitation should provide for lower future costs of care by emphasising preventive care and health promotion (Gosden et al., 2000). Capitation is used frequently in primary care in US ‘managed care’ and a few other countries including Italy, UK and US (Docteur and Oxley, 2003; WHO, 2007). In secondary care, capitation in the form of payment by diagnosis-related group is widely used in OECD countries.

**Salary**

Salary payment is seen to be rather similar to capitation with respect to its effect on provider agency. This type of incentivisation is found to further reduce activity (Gosden et al., 2000) and induce cost shifting but has the benefit of providing for more effective overall control on costs than other methods (Maynard, 2006). Salaried systems have been shown not to be very effective for pursuing public health objectives, with immunisation being one example (Gosden et al., 1999). Therefore, salary incentivisation alone may not be the most appropriate incentive scheme when improvements in efficiency and quality of care are desired. All OECD countries make partial or full use of salary payment to remunerate physicians (WHO, 2007).

**Blended payment**

Finally, blended payment mechanisms are regarded as most appropriate for balancing the interests both of payers and patients (Conrad, 2009). Such schemes are very common in practice (eg almost all OECD countries have some form of blended payment system in the primary sector and secondary outpatient care) as they allow for greater control of costs than
fee-for-service and higher service provision than capitation or salary alone (Conrad, 2009; Krasnik et al., 1990; Maynard, 2006). While theory suggests that best behavioural outcomes are achieved through an equilibrium in which providers select the contract that maximises their utility (Jack, 2005), there is no empirical evidence of the best match between observable provider characteristics and blend of payment. However, empirical work, primarily on US primary care, suggests that certain procedures in the provider mix, such as immunisation and screening, are commonly reimbursed on a FFS basis while the rest fall under the cover of capitation and/or salary (Conrad, 2009; Maynard, 2006).

While generally limited, evidence from patient surveys in the US indicates that patient satisfaction of services provided by salaried and FFS primary care providers does not differ systematically (Gosden et al., 2000). The effects of any of the three payments, alone or in bundles, on patient treatment outcomes and health status remain unclear.

### 3.2.3 Efficiency and quality of care are likely to be costly

Blended payment schemes have a greater potential to provide for efficient and quality care than single mechanisms (Conrad, 2009; Krasnik et al., 1990; Maynard, 2006). While there is not a gold standard for exogenously determining the right degree of blending, once providers are given the opportunity to self-select into contracts, provider preferences are revealed (Jack, 2005). With the help of this type of information, it may be possible to develop effective policies that steer individual providers of care towards efficiency and quality goals.

More complex mechanisms for incentivisation come at a higher price as they require data collection, evaluation, monitoring, and reporting procedures as well as an administrative capacity to facilitate them. In particular, compared to salary payment mechanisms, for example, mixed payment schemes are likely to consume more resources as they also require monitoring and correction of unintended consequences and perverse incentives. There is insufficient empirical evidence, however, to either support or refute this theoretical conjecture.

Finally, because of the complexity of the healthcare system, payment schemes that are appropriate for primary care may be inadequate for secondary care. One of the key reasons for this is that the costs of establishing and administering a mechanism in secondary care could be so high that they may significantly outweigh the benefits of adopting it. Therefore, implementation requires careful evaluation of context and objectives followed by informed development of an appropriately tailored incentive mechanism.

### 3.3 Financial incentives

Bonus, stepped bonus and withholds are financial incentive schemes designed to reward or penalise particular provider behaviour. Unlike the payment mechanisms described in the previous section, the means of incentivisation discussed here are based on either one-off payments or one-off claw-backs over a particular period of time and are usually built on top of a particular payment system.

#### 3.3.1 Bonuses and withhold reward particular provider behaviour

Bonus incentives involve the payment of a lump-sum of money to providers, who are otherwise paid according to a particular payment mechanism, for performing in a certain
desirable way or achieving some pre-defined performance target. Bonuses could be provided at an individual or group level, could be of various sizes and can be paid out over various intervals (Conrad, 2009).

In comparison, stepped bonuses break away from the uniform structure of the simple bonus scheme and bind the level of achievement to a reward of particular monetary amount.

Bonuses are commonly used together with withholds, with the bonus amount being tied to the amount of the withhold. Under withhold, therefore, a bonus to providers is paid depending on their level of achievement of performance targets.

3.3.2 The effectiveness of bonuses and withholds is questionable

The effect of the bonus (and stepped bonus) incentive on supplier-induced demand and agency (and therefore on efficiency and quality) could vary according to the nature of the bonus payment and the underlying structure of the existing system of payment (eg FFS, etc). If the nature of the bonus is linked to cost savings, it is likely that the agency between payer and provider will be strengthened at the expense of the agency between patient and provider. Inversely, if the nature of the bonus is on meeting treatment performance targets, it is likely that the effect on agency will be the reverse.

Further, the existing system of incentivisation is likely to affect provider behaviour with respect to supplier-induced demand. A bonus targeted at cost-saving under capped FFS may reduce supplier-induced demand whereas a bonus targeted at achieving treatment performance goals is likely to increase it.

There is limited empirical evidence, coming primarily from the US, UK, Germany and the Netherlands, of the ability of bonus schemes to induce cost-saving and efficiency or improve quality of care in OECD countries. Work on immunisation in the US suggests that FFS has a somewhat stronger positive impact on achieving immunisation targets than bonus payment (AHRQ, 2004). Other evidence indicates that at individual physician level, bonus and particularly stepped bonus incentives improve children’s immunisation (Conrad, 2009). Equally, annual bonus was suggested (although not through formal testing) to have played a role, together with other management tools such as diabetes registry and group meetings, in reducing the proportion of patients with HbA1 less than 9.5 over one year from 13.9 per cent to 1.8 per cent in the US (New York) (Beaulieu and Horrigan, 2005).

Studies of the effect of bonus incentives at a group-level indicate that there was no statistically significant effect of such bonuses either on cancer screening for women of middle and advanced age nor on paediatric immunisation and well-child visit rates (Conrad, 2009).

In contrast to bonus incentives, the purpose of withhold is to claw back payment to providers and thus induce them to meet some pre-established performance targets. Once targets are met, providers receive the amount withheld commonly together with a bonus payment. If providers fail to meet targets, withhold is retained by the payer. These particular features of the withhold incentive imply that its effect on efficiency and quality of care may be similar (if not slightly stronger) compared to bonus payments.
Again, there is limited evidence of the effectiveness of withholds. The few studies from the US that provide some insight into the effect of withhold at a group level indicate that the potential of bonus payment and return of withholds introduced by Partners Community HealthCare in the US (Massachusetts) was associated with increased development of medical management programmes and improved diabetic care processes; however, no impact on cancer screening for women aged 50 and older, paediatric immunisation or asthma care was found (Conrad, 2009).

At an individual-physician level, studies conducted in the US revealed that a 5 per cent withhold of funding for primary care physicians rendering services to patients with diabetes had an impact on eye examination only in the first year after adoption; there were no significant differences between treatment and control groups in the four process quality measures observed in later periods (Young et al., 2007).

Because of the lack of evidence, the effect of bonuses and withholds on treatment outcomes and health cannot be established.

### 3.3.3 Alignment of incentives across the system is crucial for bonuses and withholds to be effective

There are several key points that emerge as a result of reviewing the literature on bonus and withhold incentivisation.

While bonus payments appear to be among the easiest incentive mechanisms to implement, their effects on efficiency and quality of care are likely to depend on the structure of the underlying incentivisation. Therefore, when such schemes are adopted, the incentives that they create should be aligned with all other incentives in the system. For example, it is important to ensure that the bonus serves as a complement rather than a substitute of income of providers. If designed properly and are targeted at the right level, bonuses could be an effective means for achieving desired goals.

Withholds on the other hand, may provide a somewhat better incentive for providers as the possibility of losing income may be valued higher by some providers than the possibility of gaining income (Kahneman and Tversky, 1979; Tversky and Kahneman, 1992). However, the size of the potential marginal loss has to be large enough to incentivise providers to achieve prescribed objectives.

### 3.4 Non-financial incentives

The literature provides somewhat limited information on the effect of single non-financial incentives, with existing evidence coming primarily from the US, UK and Australia. Because of this informational deficiency, the incentives in this section are grouped together.

#### 3.4.1 Non-financial incentives focus on less tangible performance outcomes

Unlike financial incentives which remunerate providers for achieving particular performance targets, all non-financial incentives focus on less tangible performance outcomes. Some incentives, such as education and guidelines, presume willingness to comply with norms and standards, and implicitly rely on reputation as an incentive, while others such as feedback and public reporting of relative performance use reputation
explicitly to induce desired responses. Non-financial incentives such as provider accreditation and re-certification, performance audit and reminders rely heavily on participating organisations to ensure compliance.

3.4.2 **Active delivery of education and audit with feedback could be effective incentives**

The empirical evidence on the effectiveness of non-financial incentives is somewhat limited. Where evidence exists, it is often not specific on contexts of implementation and specific effects as well as is very commonly coupled with a number of other initiatives, which makes identifying a clear causal relationship between incentive scheme and outcome difficult to establish.

Education delivered passively (e.g., delivery of educational materials by mail and conferences) was found to have no impact either on primary care physician behaviour or quality of care and health outcomes (Epstein, 2009; Grimshaw et al., 2001; Oxman et al., 1995). Guidelines delivered passively had no effect on performance or outcomes as well (Oxman et al., 1995). However, active forms of dissemination such as interactive workshops, outreach and educational sessions showed an overall positive but inconsistent effect on physician behaviour (e.g., prescribing) and quality of care in a few countries, including Australia and the UK (Epstein, 2009; Grimshaw et al., 2001; Oxman et al., 1995). Active dissemination approaches though were found to be associated with higher costs (Grimshaw et al., 2001).

In contrast, the evidence on the effect of implementing guidelines is mixed. Overall, secondary care physicians were found to be more responsive to guideline implementation than physicians in primary care (Grimshaw et al., 2001).

Studies on feedback and public reporting in the US indicate that these non-financial incentives have a somewhat positive effect on provider behaviour. Indeed, public reporting was suggested as one of the reasons for the observed decline in mortality associated with cardiac surgery in the US (New York) (Epstein, 2009; Shekelle, 2009), while low-performing hospital in the US whose scores were published were found to be more likely to adopt strategies for quality improvement (AHRQ, 2004).

While there is a lack of comprehensive evidence as to the effectiveness of accreditation and re-certification (Epstein, 2009), audit with feedback has been found to have a variable impact on quality of care. Early work from Australia, Canada and Denmark suggested that audit with feedback is not a particularly effective incentive for improving provider performance (Epstein, 2009). However, subsequent work showed that it had a modest effect on quality of care with greater impact when provider’s adherence to recommendations is low and when feedback is provided intensely (Grimshaw et al., 2001).

Reminders, as another non-financial incentive, were found to have at best a moderate effect on provider performance (Grimshaw et al., 2001; Oxman et al., 1995). In many cases, there was no effect.

3.4.3 **Multifaceted incentives are important to change to performance of providers**

The overarching conclusion about the effectiveness of non-financial incentives in healthcare is that it may take multifaceted incentivisation to achieve changes in the performance of providers (Grimshaw et al., 2001; Oxman et al., 1995). However, non-financial incentives, even when adopted in bundles, may produce inconsistent changes in
provider behaviour linked to improvements in patient health across cases (Oxman et al., 1995).

3.5 **Mixed schemes of incentivisation**

This section contains a description and discussion of comprehensive pay-for-performance (P4P) schemes as one example of mixed mechanisms of incentivisation. The section also offers some lessons on the implementation of P4P schemes in OECD countries.

3.5.1 **Pay-for Performance aims to reward providers who meet certain performance expectations with regard to quality of care or efficiency of care**

Compared to the previous incentive mechanisms discussed in this chapter, P4P is substantially more comprehensive and its primary objective is to reward providers who meet certain performance expectations with regard to quality of care or efficiency of care (RAND, 2010). The key advantage of P4P incentivisation over other mechanisms is that it is evidence-based, with modern P4P mechanisms incorporating patient outcomes along-side process measures (Conrad, 2009).

The incentivisation under P4P is realised with the help of three principal components:

- a set of performance measures;
- a set of performance goals;
- and a mechanism for performance evaluation and reward (Scheffler, 2010).

The performance of every provider is measured and scored against the goals and that of other providers. This evaluation and ranking provides the basis for the distribution of rewards (financial or non-financial) to providers. P4P schemes are flexible enough to provide for incentivisation both at individual and group level as well as in primary and secondary care.

3.5.2 **Comprehensive P4P schemes have varying effectiveness**

P4P incentivisation mechanisms vary significantly in size and scope within OECD countries. Practically all countries have adopted P4P incentivisation mechanism in either primary or secondary care, with P4P being particularly focused on care for the chronically ill. As the organisation and funding mechanism of the healthcare systems differ significantly across countries, P4Ps also vary in objectives and configuration. Moreover, in some countries, such as the US, where there is not a single payer for care, the specific set-up, size and scope of P4P commonly vary with the objectives of payers, which results in an array of P4P schemes. Due to these reasons, understanding what works and what does not in P4P schemes is a particularly daunting task. To keep the discussion focused, the chapter offers some general theoretical considerations which are then complemented with empirical evidence collected primarily from studies of such schemes operating in the UK and US.

Evidence-based performance assessment is one of the most notable building blocks of P4P schemes. A high degree of transparency in provision is likely to help reduce the information asymmetry between providers and payers and patients.
Furthermore, it is the case that P4P incentive schemes are targeting not only supplier-induced demand but also double agency by attempting to align the interests of all three parties involved in healthcare provision while asserting transparency and accountability. Therefore, from a design viewpoint, P4P offers a substantial promise to help deliver improvements in the efficiency and quality of care.

The empirical evidence on the effectiveness of P4P, however, is not entirely supportive of theoretical constructs. One reason for this is that while the majority of existing P4P target efficiency together with quality, the number and types of indicators employed to measure the latter are rather limited. They only focus on particular types of care (eg chronic care) or are not particularly well matched to outcomes. Another reason is the practical difficulty of fully implementing P4P that hinder P4P from realising its full potential.

Studies on P4P’s based mainly on financial incentives have failed to find conclusive evidence of its effectiveness to improve efficiency and quality of care (Epstein, 2009). Factors such as type of clinical condition, size and scope of incentive programme as well as timing of outcome measurement have been identified as possible drivers of these findings.

Evidence from studies on P4P in UK primary care (evaluated based on the Quality of Outcomes Framework) suggests that the quality (irrespective of value for money) of care for three conditions (ie angina, diabetes and asthma) covered by P4P continued to improve after the introduction of the scheme. Indeed, the quality of care for asthma and diabetes improved at a faster rate compared to the period preceding P4P with this improvement linked to P4P (Busse and Mays, 2008). However, P4P had no effect on angina. Further, once targets were reached, the improvement in the quality of care for all three groups of patients slowed and the quality of those aspects of care for asthma and heart disease that were not incentivised decreased. In addition, the scheme had a negative impact on the continuity of care, which was lower after the adoption of P4P (Campbell et al., 2009).

P4P was found to be associated with inequalities of care between deprived and non-deprived areas. These differences, however, were small on average and getting smaller over time (Busse and Mays, 2008).

In US primary care, P4P is widely adopted by commercial health management organisations paid on the basis of capitation (Rosenthal et al., 2006). Various forms of incentivisation are used and there are significant variations in the number and types of measures employed in P4P systems operating around the country. As in the UK, however, the overall findings on their effectiveness are inconclusive (Epstein, 2009).

There is more evidence on the effect of P4P in secondary care and specifically in hospital care. This evidence comes primarily from the US and indicates that hospitals where public reporting was combined with a P4P scheme showed “modestly greater improvements”, 2.6 to 4.1 per cent over two years, in four composite measures of quality (including measures of care for heart failure, acute myocardial infarction and pneumonia) compared to hospitals committed to public reporting only (Lindernauer et al., 2007).

---

9 See Appendix C.
Further, a comparative study between US hospitals involved in a voluntary quality improvement initiative and hospitals not part of the initiative revealed no significant differences in the quality of care provided to myocardial infarction patient. Incentivised hospitals performed better with respect to supplying patients with an aspirin at discharge and smoking cessation services (Conrad, 2009). Another study showed that as a result of the adoption of a voluntary incentive programme by Blue Cross Blue Shield of Michigan focusing on heart care outcomes, the hospital participating in the initiative modestly improved its adherence to heart care guidelines, which led to savings of approximately $12,000 to $30,000 per QALY (Nahra et al., 2006).

Overall, work on P4P programmes in secondary care, including chronic care, tends to show improvement both in processes and outcomes a result of their adoption. This is specifically the case with individual hospitals, insurers’ programmes and health care provided through integrated delivery networks (Busse and Mays, 2008).

There is substantially less evidence as to the effectiveness of P4P incentivisation on treatment outcomes and health. Existing work either does not test for statistical significance, finds no significant differences (Damberg et al., 2007) or provides inconclusive and condition-specific results. For example, P4P scores in UK primary care were found not to be associated with emergency admissions for coronary heart disease but to be negatively associated with emergency admissions for epilepsy (Roland, 2008).

3.5.3 Complex incentives systems require careful development and substantial resources

A significant body of work has focused on the lessons learnt from the application of comprehensive P4P incentivisation and has provided recommendations for improving future P4P initiatives both in primary and secondary care.

Studies have shown that continuous incentives may have a greater impact on providers than one-time bonuses (Busse and Mays, 2008), whereas financial incentives alone may be less effective than when combined with other types of interventions targeting quality (Roland, 2008). Public reporting in particular has been suggested as a successful instrument in promoting quality improvement in secondary care as hospitals tend to be conscious about their public image (Damberg et al., 2007). Rigorous performance monitoring and benchmarking has been suggested as another successful avenue for performance improvement in P4P schemes. The case of US Veteran Health Administration is a good example (Busse and Mays, 2008).

The level of incentivisation is equally important and should be determined in a way that helps to achieve the performance goals of P4P. The effect of financial incentives tends to be smaller at group level and somewhat larger at individual level. Hence, multi-level incentivisation has been proposed as one possible alternative (Damberg et al., 2007). If, for example, provision of counselling services is desired, incentives may be more effective when aimed at individual physicians (Rosenthal and Dudley, 2007). On the other hand, group-level incentivisation may be more effective for improving group-level efficiency such as adoption of information systems (Rosenthal and Dudley, 2007). Ultimately however, the level of incentivisation should be in accordance to the size, characteristics and activities of providers (Christianson et al., 2006; Roland, 2008).
The choice of measures of performance is another important facet to designing P4P mechanisms. While there is no clear rule on how measures of performance should be selected (Busse and Mays, 2008), providers may be able to determine what should be measured (Rosenthal and Dudley, 2007). Hospitals tend to prefer a limited set of measures that focus on quality improvement (Damberg et al., 2007) as well as standardised measures which minimise physician resistance to these measures (Damberg et al., 2007).

However, it is commonly the case that indicators which are easiest to measure and least costly to observe are the ones included in P4P schemes, although they may not be those which are associated with health improvement (Christianson et al., 2006). Furthermore, it may be that a “one size fits all” approach is particularly inappropriate. For example, providers in deprived areas may experience difficulties in responding to incentives and improving their performance “even when payment is based on simple and widely accepted performance measures” (Christianson, 2007).

The objectives of the incentive scheme and the type of thresholds used for benchmarking have raised some additional concerns among providers. Physicians require clearly specified objectives in P4P schemes (Busse and Mays, 2008). They also tend to prefer absolute performance thresholds (Conrad, 2009; Damberg et al., 2007), which reward all good performers (Damberg et al., 2007). To be an effective instrument, P4P schemes should reward high value care and avoid discriminating between high and low achievers. This could be accomplished by rewarding providers on the basis of marginal improvement such as quality care for an additional patient (Rosenthal and Dudley, 2007).

The size of the incentive under P4P is another significant feature. Effective communication with providers and consideration of their initial conditions are two important factors when designing rewards. The likely existence of heterogeneity among providers may result in the need for differential payment depending on provider characteristics (Christianson, 2007; Rosenthal and Dudley, 2007). Such type of payment may be especially necessary when providing care to disadvantaged populations. To improve the quality of care in such settings, it may help to offer larger incremental payments. Alternatively, other types of rewards such as capital grants, technical assistance, or training may be provided to promote the provision of desired quality of care (Rosenthal and Dudley, 2007).

Studies identify a number of challenges that P4P schemes face. The problem of small numbers is one of them. This results in unstable or inaccurate estimates of hospital performance and affects the performance-based incentive payments. While small and rural hospitals are especially susceptible to this problem, similar issues have occurred in other types of hospitals as well (Damberg et al., 2007).

P4P schemes require data, which may be inaccurate or costly to collect (Busse and Mays, 2008). Therefore, P4P incentive programmes are likely to demand a considerable amount of resources for the developing, monitoring, collection and reporting of data, as well as technical and administrative support (Damberg et al., 2007; Roland, 2004).

Perverse incentives and unintended consequences are some additional challenges for P4P schemes. Creaming has been identified as a problem with physicians who excluded a greater fraction of patients registering higher performance and therefore receiving higher rewards (Busse and Mays, 2008; Christianson, 2007). In contrast, although the presence of
unintended consequences has not been established by empirical studies of P4P in primary or secondary care, the most likely reason for this is that relevant data are not collected (Roland, 2008). Indeed, aspects of care which are not incentivised should be carefully monitored as a way to guard against unintended consequences. Adverse consequences of P4P adoption such as fragmentation and poor coordination of care have been raised as concerns as well (Roland, 2004).

Finally, on the basis of the existing evidence, it is impossible to establish the degree of impact of a specific aspect of P4P on the overall outcome in terms of efficiency and quality of care (Christianson et al., 2007; Conrad, 2009). The lack of comprehensive costeffectiveness or value for money information is another significant drawback in the existing literature (Conrad, 2009).

3.6 Conclusions

3.6.1 Multi-faceted incentivisation is likely to provide better outcomes in terms of quality and efficiency of care

There are a few key conclusions arising from this review on the use of incentives in healthcare.

From a theoretical viewpoint, there are three intertwined issues that initiatives targeting improvement of healthcare provision have to resolve. The first is the asymmetry of information in the healthcare market which gives rise to the second, supplier-induced demand. The last is provider double agency. All three have a direct impact on efficiency and quality of care. Reducing and eventually eliminating the issue of asymmetrical information and aligning the interests of payers, patients and providers is a necessary condition for the success of any incentivisation scheme.

Designing such mechanisms in healthcare is a complex problem as attested by the studies on effectiveness of P4P. Therefore, it is rather difficult to pin-point the precise building blocks of the incentive scheme that will help achieve the ultimate goal of efficiency and quality of care. The interaction between the incentive scheme and the underlying payment scheme further adds to the complexity in design. Still, this review has flagged a few key characteristics that may help improve overall outcomes.

Multifaceted incentivisation consisting of a bundle of financial and non-financial incentives is likely to provide better outcomes in terms of quality and efficiency of care.

3.6.2 When considering introducing incentive schemes to different contexts including middle- and low-income countries, it is important to consider that designing incentive schemes in OECD countries requires significant resources

Complex incentives schemes work best when:

- they are complemented by a carefully designed context-specific incentive mechanism aimed to promote a response at the appropriate level (eg organisation-wide, group or individual);
- the size of any financial incentive is sufficient to induce desired response while preferably crowding-in extrinsic (or moral) motivation (Frey, 1997);
- rewards are provided to all achievers;
• they are certain and over long enough intervals as to incentivise providers to continue improving;
• they carefully consider the choice of appropriate performance measures (i.e., measures of processes and outcomes) and the use of absolute benchmarks, which are crucial for monitoring improvement; and
• monitoring for perverse incentives and unintended consequence becomes a regular practice.

There is a further consideration, particularly relevant to middle- and low-income countries. Designing complex incentivisation schemes has been shown to require a significant amount of resources. Comprehensive mechanisms are expensive as they require a large amount of data associated activities such as monitoring, collection and reporting as well as a significant administrative capacity to coordinate and facilitate the entire process.
Reference List

Social Protection


Bruttel, O., Contracting-out the Public Employment Service and the consequence for hard-to-place jobseekers: Experiences from Australia, the Netherlands and the UK, conference paper, Wissenschaftszentrum Berlin für Sozialforschung (WZB), 2004.


Bruttel, Oliver, “Are Employment Zones Successful? Evidence From the First Four Years”, Local Economy, 20; 2005b, 4, 389 — 403


Cameron, M., Brown, K., Crooks, S., Young, N., Job Network Best Practice, Department of Employment and Workplace Relations, 2006.


Van Stolk, C., Rubin, J., Grant, J., Benchmarking the Use of Personal Advisers in Jobcentre Plus, 2006, RAND: Santa Monica CA.

**Healthcare**


Audit with feedback: Non-financial incentive in healthcare, whereby providers are audited every so often and offered feedback on performance (often also linked to disciplinary or release of public information).

Bonus payments: Bonus incentives involve the payment of a lump-sum of money to providers in healthcare for performing in a certain desirable way or achieving some pre-defined performance target.

Bonuses (stepped): Similar to bonuses in healthcare, but they break away from the uniform structure of the simple bonus scheme and bind the level of achievement to the reward of a particular monetary amount.

Cost reimbursement schemes: Payment system in employment service provision whereby service providers are paid on the costs they incur rather than outcomes.

Capitation: Payment system in healthcare systems whereby payers place a cap in healthcare systems on the spending per patient, often aimed at reducing supply of services.

Creaming or Cream skimming: Externality of P4P schemes, whereby providers of employment services select easy to place jobseekers to achieve their performance targets.

Double agency: Common systemic problem in healthcare, which refers to providers having contractual relationships with payers and users that could imply contradictory priorities.

Effectiveness: Getting more (eg outcome) for the same amount of expenditure.

Efficiency: Getting the same (eg outcome) for less expenditure.

Employment Zone: A UK welfare to work programme in specific deprived areas that seeks to place jobseekers into employment.

Fee for service: Payment system in healthcare systems, whereby providers have few limits on the number of services they provide.

Fixed fee schemes: Payment system in employment service provision whereby service providers are paid on the basis of fixed fees and a specific number of services.
Fundholding: An incentive scheme in healthcare, which provides general practices with annual funding to purchase a bundle of services on behalf of a particular population.

Moral hazard in healthcare: Situation where consumers who are not responsible for the cost of provision will demand more services than optimal.

No cure no pay scheme: Straight P4P scheme in use in employment service provision whereby providers are only paid on the basis of outcomes.

No cure less pay scheme: Hybrid P4P scheme in use in employment service provision whereby providers are partly paid on the basis of outcomes and partly on costs.

One-stop-shop: Government employment advice agency that offers advice to jobseekers, helps them find jobs, and processes benefit applications.

Parking: Externality in P4P schemes, whereby employment service providers decide not to offer services to hard to place jobseekers in order to achieve performance targets more easily.

P4P: An umbrella term for a range of financial incentives that are concerned with budgeting, contracting or paying for outcomes/performance rather than costs/time.

Performance management frameworks: General frameworks that set outcome-based performance targets for government departments (an example is the Public Service Agreements in the United Kingdom)

Public employment services: Employment services offered by the government mainly through a dedicated government agency.

Public release of information: The release of performance information by providers in regular reports (mostly used in healthcare).

Quality of outcomes framework: Well-known comprehensive P4P scheme used in the UK National Health Service.

Salary: Payment system in healthcare, whereby cost control is exercised through salary payment.

Supplier-induced demand: An externality in a healthcare system whereby suppliers oversupply the market thus driving up systemic costs.

Trade-off in outcomes: Term used in employment provision when the positive outcome of placing a jobseeker in to work is associated with some negative factors such as the poor quality and instability of the job.

Welfare to work: Active labour market programmes aimed at taking those on benefits (back) into employment.

Withhold: Payments in healthcare, a bonus to providers is paid depending on their level of achievement of performance targets.
Appendix B: P4P examples in employment provision: Australia and the Netherlands

In order to look at the effect of P4P schemes, it is important upfront to reflect on the differences between these schemes. There are three fundamental aspects to the design of P4P systems (adapted from Bruttel, 2005a):

- what to incentivize and how much;
- what type of information systems to use, and;
- what control mechanisms to put in place

Australia and the Netherlands illustrate the differences in the use and design of P4P schemes in employment services. Table 4 provides a brief overview of some of these differences. These are elaborated on in the sections below.

Table 4: Overview of cases to illustrate the differences in the use of P4P in employment services

<table>
<thead>
<tr>
<th>Case</th>
<th>Type of contracts</th>
<th>Information system</th>
<th>Control mechanism</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>Hybrid</td>
<td>Highly sophisticated, Econometric model</td>
<td>Sophisticated monitoring and strict control systems</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>Pure P4P</td>
<td>Intermediate, no specific measures to link outcomes to incentive</td>
<td>Very low; provider autonomy prioritised</td>
</tr>
</tbody>
</table>

Source: RAND Europe compilation

**Australia**

The Australian government introduced the “Job Network” in 1998 as a means of bringing choice and competition into the provision of employment services, and thereby introducing financial incentives to traditional mechanisms. The structure of the system mirrors that outlined in Figure 2, separating the government “one stop shop” from private and third sector operators. (Productivity Commission, 2002)

The Australian government selected a hybrid version of pay for performance, where providers are rewarded both upon commencement and for outcome. Table 5 provides an overview. Commencement fees are the only non-outcome based payments in the Australian system.
Table 5: Fee structure of the Job Network (EC 2)

Job Matching:

i. Employment outcome
ii. Fixed bonus for long-term unemployed

Job-search training:

iii. Commencement fee
iv. Client gains employment within 13 weeks and
  retains for over 13 weeks with average benefit
  reduction of 70% or more

Intensive assistance:

v. Commencement fee (30% of contracted price)
vi. Primary interim outcome (job for 13 weeks, 70% of contracted price)

vii. Primary outcome fee (13 additional weeks)

Source: OECD 2001

Providers enter the Australian system through a process of competitive tender. This system, which has varied between contract rounds, has in essence moved from being solely outcome and price focused to include a higher number of quality indicators in later contract rounds (Bruttel, 2004). Jobseekers enter through the government agency Centrelink which uses the Job Seeker Classification Instrument to classify jobseekers into groups based on how difficult outcomes would be to achieve, and then randomly assign them to service providers, unless a choice is made by the jobseeker (OECD, 2001; Productivity Commission, 2002). Figure 4 gives a schematic overview.

Australia uses information systems in the form of benchmarking and monitoring. The benchmarking system uses econometric modelling to assign “star ratings” on providers, based on their ratio of successful outcomes adjusted for other significant variables such as area, local employment situation etc. (Bruttel, 2004; OECD, 2001). The star rating has been used as a mechanism in the re-awarding of contracts in later rounds, imposing a significant incentive on providers to achieve high adjusted outcome rates (Productivity Commission 2002). The monitoring system is computer based, and allows Centrelink to track every jobseeker in real time. Computerised risk assessments and desk monitoring exercises assign risk ratings to agencies and adjust monitoring visits accordingly (Bruttel, 2004).

Aside from using the Star Ratings as an ex-post control mechanism (not re-awarding contracts to underperformers), Australia has increasingly moved towards a system of fixed payments for services delivered, actually decreasing the importance of outcome payments after the third contract round (EC3) (Bruttel, 2004).
Figure 4: Comprehensive overview of the Australian system (EC2)

Centrlink streams unemployed jobseekers to service channels:

<table>
<thead>
<tr>
<th>Candidates for self-employment</th>
<th>Self-service clients</th>
<th>Clients requiring assessment of disadvantage (administered by Centrlink using JSCI) and, in some cases SNA or IATs)</th>
<th>Referrals according to level of disadvantage</th>
<th>Special categories of unemployed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Low</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CSP</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Disability Employment Services (refer to FaCS measures)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Other programmes</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Long-term benefit recipients (work for the Dole)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Youth (DETYA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>- Indigenous people (DEWRSB)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Source: Adapted from OECD 2001
The Netherlands

The system introduced in the Netherlands follows the institutional set-up described in Figure 2. The public Center for Work and Income (CWI) contains the one stop shop functions consisting of client profiling, running vacancy databases, collecting data and handling easy to place clients. Placement training services are contracted externally (Bruttel 2004).

The Netherlands have moved from a hybrid system with even splits between commencement payment and outcome payment as the most common reimbursement method to a more straight P4P system with solely outcome-based payments (Koning et al., 2010). This reflects a view of government as enablers rather than controllers and negative early experiences with high levels of regulation (Stuyven et al., 2005). This development makes it one of the most interesting cases for P4P. Contrary to Australia, the Netherlands has removed product-based contracting completely to give providers full freedom in how to spend funds (Bruttel, 2004).

Providers enter the system through a process of competitive tender based on “experience, offered outcome rate, price and method” (Bruttel 2004). Jobseekers enter the system through the CWI Kansmeter and in depth interview techniques (Kwint). They are classified into four different groups according to job readiness, with group 1 being assisted by the CWI and other groups passed on down to private providers through case managers (Bruttel, 2004).

The Netherlands have one of the least sophisticated control and information systems. The government opted for self-regulation which is done through a private industry organisation external auditors, and regular monitoring reports. There is no coherent information system (Bruttel 2004).
Appendix C: Some examples of studies on provider incentives used in healthcare

Audit with feedback in Canada (Ontario)
The effect of audit and feedback with education on the rate of prescribing benzodiazepines to elderly patients was investigated in a randomised control trial setting (Pilmott et al., 2003). The study looked at 1,624 primary care physicians who each wrote at least ten prescriptions for this group of drugs within a period of two months. Physicians were randomly assigned to an intervention or a control group. The study also identified a sample of residents of Ontario aged 65 years and over from the Ontario Drug Benefit claims database (Epstein 2009).

Every two months, within a six-month period, profiling for benzodiazepine prescriptions written for elderly patients together with evidence-based educational bulletin were mailed to the intervention group. The control group received feedback and educational materials about anti-hypertension drug prescribing for elderly patients. The authors looked at the proportion of physician’s total benzodiazepine prescriptions, prescriptions for benzodiazepines in combination with other psychoactive drugs and long-term benzodiazepine therapy.

There were 168 physicians in the intervention group and 206 in the control group with demographic and prescribing characteristics being similar across the two groups. By the end of the intervention, long-acting benzodiazepine prescribing decreased by 0.7 per cent in the intervention group whereas in the control group it increased by 1.1 per cent. The authors found no statistically significant difference between these two effects and concluded that “a program of confidential feedback and educational material offered to Ontario primary care physicians did not have a clinically significant impact on their benzodiazepine prescribing”.

Clinician education in the United States (Minnesota)
A randomized controlled trial with 37 community hospitals in Minnesota was conducted to study the effect of clinician education by local opinion leaders and performance feedback for improving the quality of care for patients with acute myocardial infarction (Soumerai et al., 1998). All patients with this condition admitted to study hospitals over 10 months before and after the intervention were considered.

As part of the intervention, opinion leaders at 20 experimental hospitals were identified who held small and large group discussions with peers, informal consultations, and
revisions of protocols and clinical pathways. The opinion leaders focused on three particular issues: evidence of drug efficacy, comparative performance, and barriers to change. Control hospitals in contrast, received performance feedback by mail.

Results of the study indicated a statistically significant increase by 17 per cent of the proportion of eligible elderly patients receiving aspirin and a statistically significant increase in other life-saving drugs. Harmful therapy was found to decrease by around 50 per cent in both groups. The study therefore concluded that use of local health opinion leaders together with providing performance feedback can speed the adoption of beneficial therapies for myocardial infarction.

P4P in the UK National Health Service
Challenged by a perceived shortage of general practitioners (GPs), the UK government negotiated and, in April 2004, implemented a new GP contract for general medical services. The principal goals of this contract were to contribute to an increased supply of physicians and help improve the quality of healthcare (Smith and Goddard, 2009).

An inseparable part of this new contract for medical services was a P4P incentivisation mechanism, the Quality of Outcomes Framework10 (QoF). The QoF consisted of 76 quality indicators in 10 clinical domains of care11, 56 indicators in organisational areas, four assessing patients’ experience and a number of other indicators for additional services (Roland, 2004; Shekelle, 2003). The framework also set targets for performance with the level of performance on each indicator yielding a score. The aggregate score of a general practice determined its quality bonus, with approximately a quarter of the practice’s income depending on its performance on the QoF (Smith and Goddard, 2009).

The QoF was acclaimed for a few of its features. The large number of indicators in multiple domains of care which are linked to methods of implementation was pointed out as one of its particularly attractive aspect. An additional positive feature was the attempt to align the incentives of the payer (ie the UK government) and the providers of primary care (Shekelle, 2003). Improved quality of care, rapid expansion of clinical computing systems, increased role of nurses in primary care, increased number of clinics specialising in management of conditions covered by the QoF, increased specialisation in primary care and biomedical orientation of primary physicians were also pointed out as positive and anticipated consequences (Roland, 2004).

However, unintended consequences of the framework were recognised as an issue of concern. Lack of monitoring in aspects of care, which were not incentivised, as well as a possible change in the relationship between doctor and patient were identified as threats to the quality of primary care provision (Roland, 2004; Shekelle, 2003). Increased administrative costs were also pointed out as a drawback of the new contract (Roland, 2004). Indeed, the cost for implementing the contract in the first three years were 9.4 per cent higher than initially intended (Smith and Goddard, 2009).

---

10 See Roland (2004) for a detailed discussion.

11 Clinical domains included coronary heart disease, stroke, transient ischemic attack, hypertension, hypothyroidism, diabetes, mental disorder, chronic obstructive pulmonary disease, asthma, epilepsy, and cancer.
Following the introduction of the QoF, on average, primary care practices achieved over 90 per cent of the maximum points available in every year since 2004, which is indicative of the provision of quality care to patients in the targeted domains of care. However, at the time of introduction of the QoF, there was already an underlying trend of improvement in primary care. It is difficult therefore to decouple the effect of the QoF from that of the overall tendency towards better quality care. (Smith and Goddard, 2009)
## Appendix D: Annotated bibliography on provider incentives in social protection

<table>
<thead>
<tr>
<th>No</th>
<th>Reference Details</th>
<th>Year of Publication</th>
<th>Type of Publication</th>
<th>Country/state</th>
<th>Programme</th>
<th>Types of Incentives</th>
<th>Effect of Incentives</th>
<th>Implementation details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Abello, D., H. MacDonald (2002), ‘Job Network: changing community sector values’, The Drawing Board: An Australian Review of Public Affairs, 3 (1), 51-63.</td>
<td>2002</td>
<td>Peer-reviewed journal; Case study; single case. Looking at the effect on non-profit organisations. Evidence used: focus group interviews, survey data and administrative data</td>
<td>Australia</td>
<td>Job Network</td>
<td>Hybrid of Pay for Performance and Cost-reimbursement; Procurement Process</td>
<td>The implementation of pay for performance seems to have radically changed the operation of non-profit agencies, towards a less holistic approach. A conflict was perceived with previous collaborative and open approaches in a new competitive environment. Non-profit organisations generally didn’t do very well in the programme, and approaching the third tender round, mainly religious organisations remained, with most community-based, secular organisations having failed to deliver on outcomes.</td>
<td></td>
</tr>
<tr>
<td>ID</td>
<td>Author(s)</td>
<td>Year</td>
<td>Study</td>
<td>Country</td>
<td>Incentives Model</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----</td>
<td>-----------</td>
<td>------</td>
<td>-------</td>
<td>---------</td>
<td>-----------------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Armstrong, D., Byrne, Y., Paton, L., Horack, S.</td>
<td>2009</td>
<td>Government contracted report, Case study: single case.</td>
<td>U.S., New York</td>
<td>Pay for Performance; Hybrid of Pay for Performance and Cost-reimbursement; Prime Provider model</td>
<td>Incentives are perceived as an efficient and streamlined way of implementing strategic goals of commissioners. Incentives have been pivotal in steering outcomes, demonstrated by shifts in outcomes achieved through changes in contracts from short-term employment focus to long-term milestones. Main problems of outcome focused pay for performance contracting have been the significant risk for private providers, creating barriers to entry to the market, and uncertainty around exogenous factors, causing several mid-term contract renegotiations. The study identifies three types of risks with the pay for performance system. Outcome related risks could mean that outcomes measures do not reflect intended outcomes. This was the case in NY when only job outcomes were specified, which was subsequently changed to long-term outcomes. Market related risks in forecasting supply of clients have also proved vast, leading to some instability and over/under supply at times. Finally, a client risk is identified, which arises when outcome goals are disconnected from users, causing a risk of poor user experience.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Bredgaard, T., Larsen, F.</td>
<td>2007</td>
<td>Peer-reviewed journal. Review of case studies: Comparative case study of three cases looking at institutional effects.</td>
<td>Australia; The Netherlands, Denmark</td>
<td>Pay for Performance; Hybrid of Pay for Performance and Cost-reimbursement; Procurement Process</td>
<td>The authors find that quasi markets aside from being a management model and institutional set-up shapes a significant proportion of the employment policy in the cases studied. Through front-line worker values and agency priorities, there is a shift from a &quot;human capital&quot; approach to a &quot;work first&quot; approach, driven by the incentives used. This is also true of NGO organisations previously committed to the &quot;human capital&quot; approach. They further argue that steering by incentives is a problematic approach that seems to call for extensive public regulation. They identify a pattern where services are first contracted freely on an outcome pay basis in the first tender round, to then be reined in and subjected to increasing amounts of regulation to offset adverse effects of the pure outcome focus, such as creaming and parking. This re-bureaucratisation process has progressed the furthest in Australia, which according to the authors has moved to a limited preferred provider model.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Brodkin, Z. E., Towards a Contractual Welfare State? The Case of Work Activation in the United States, in Sol, E., Westerveld, M., Contractualism in employment services: a new form of welfare state governance, Kluwer Law International: the Hague, The Netherlands, 2005 Book. Case study: Single case. Looking at outputs and unintended consequences. Evidence used: key informant interviews, discussion of previous literature and administrative data</td>
<td>U.S.; Chicago</td>
<td>Welfare-to-Work</td>
<td>Pay for Performance</td>
<td>The author cannot make a control group comparison because of lack of data. However, three noticeable effects of performance based contracts are discussed: creaming, goal displacement and trade-off challenges. Creaming was observed in the programme through officials screening clients for employability, and referring and offering intensive assistance to those perceived as most ‘employable’. Goal displacement was a significant frustration with local not for profit organisations that had previously focussed on human-capital building and long-term outcomes, but now changed their delivery models to suit incentives, offering short-term placements and minimising human capital building. The trade-off challenges perceived were that clients going through contracted agencies tended to gain employment that scored low on stability, access to benefits, regularity of work-hours and opportunities for learning and advancement.</td>
<td>Performance contracts in Chicago were very rudimentary in nature, only providing relatively short-term outcome measures, thus, poor contract design might be an issue in this case.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| 5 | Bruttel, O., (2004) Contracting-out the Public Employment Service and the consequence for hard-to-place jobseekers: Experiences from Australia, the Netherlands and the UK, conference paper, Wissenschaftszentrum Berlin für Sozialforschung (WZB) | 2004 Conference paper. Case study: comparative case study of 3 cases. Looking at effect of contracting-out, primarily effectiveness. Evidence used: administrative data, expert interviews, review and comparison of previous studies through discussion | UK; Australia; Netherlands | Employment Zones; UWV; Job Network | Hybrid of Pay for Performance and Cost-reimbursement; Pay for performance; Procurement Process; Client Choice; Performance pay | UK Employment Zones can be demonstrated to have had significant positive effects. Evidence remains weak in Australia, with only government reports that tend to show a slightly positive result. No evidence base exists for evaluating effectiveness in the Netherlands. In all three cases, expert interviews demonstrated a far more flexible work routine in the programmes than compared to previous provision. However, rather than services becoming increasingly heterogeneous as was expected, all cases demonstrated a homogenisation towards very generic services. In the Netherlands and Australia, the performance incentives also led to a decrease in schooling and training for jobseekers. Strict financing rules in the Netherlands left little space for innovation. There is evidence in all cases of performance incentives being passed on to case managers, who often worked on performance pay. This caused low-performing case managers to leave in the medium to long term. The author further argues that pure pay for performance encourages creaming and parking, and that in order to reach hard-to-place work seekers, it must be complemented with performance benchmarking, control and monitoring. | It seems important to combine the outcome focused performance incentives with loose regulation in order to encourage innovation and avoid the Dutch situation. Good information systems, benchmarking in relation to different kinds of job-seekers seem to be important to avoid creaming and parking. |

2005 Discussion paper. Case study: comparative case study of 3 cases. Looking at governance mechanisms related to contracting out with respect to moral hazard. Evidence used: administrative data, review and comparison of previous studies through discussion.

Hybrid of Pay for Performance and Cost-reimbursement; Pay for performance; Procurement Process; Client Choice; Performance pay.

In all three cases, incentives were the most important governance mechanism to align provider incentives with government priorities. All countries used some form of mix between outcome and fixed rates, reflecting a trade-off between optimal incentives and the risk aversion of contractors. Providers pass on incentives to case managers and sometimes even to clients, performance pay was a rule in all countries. In all cases, wage levels and educational outcomes were not prioritised, reflecting the fact that these outcomes were not incentivised, and causing employment services delivered to be unsuitable for some groups. In all countries quality standards were set to complement incentive mechanisms. In the Netherlands a change in contractual regimes towards a pure output focus was met with an increase in price from providers, confirming the trade-off between fixed rate and pay for performance as risk shifts between public and private. In Australia evidence shows that private case managers are less guided by rules in their work than corresponding public workers in other countries. Contracts have evolved towards more standard setting in Australia, reflecting previous problems with parking.

Providers pass on incentives to case managers and sometimes even to clients, performance pay was a rule in all countries. In all cases, wage levels and educational outcomes were not prioritised, reflecting the fact that these outcomes were not incentivised, and causing employment services delivered to be unsuitable for some groups. In all countries quality standards were set to complement incentive mechanisms. In the Netherlands a change in contractual regimes towards a pure output focus was met with an increase in price from providers, confirming the trade-off between fixed rate and pay for performance as risk shifts between public and private. In Australia evidence shows that private case managers are less guided by rules in their work than corresponding public workers in other countries. Contracts have evolved towards more standard setting in Australia, reflecting previous problems with parking.


Employment Zones; UWV; Job Network.

Hybrid of Pay for Performance and Cost-reimbursement; Pay for performance; Procurement Process; Client Choice; Performance pay.

The paper highlights the importance of balancing three governance mechanisms: incentives, control mechanisms and information systems. The importance of informational systems is underscored and classified into two models: benchmarking and monitoring. Australia has a very successful benchmarking system where econometric modelling is used to compare providers across areas and client groups. The monitoring system is likewise very advanced, giving central real-time follow-up on all job seekers through a web based system used by all providers. In the Netherlands and UK, benchmark comparisons are made using crude placement figures, increasing likelihood of parking and creaming. The Netherlands and Australia use different kinds of grouping to avoid creaming and parking problems. In the Netherlands, clients are subdivided into six general groups with regard to employability, and performance payments are adjusted accordingly. In Australia, clients are first centrally classified using the Job Seeker Classification Instrument to avoid the same issues.

The distinctive environment in which Employment Zone payment for performance incentives operate are minimum regulation, customer choice and provider competition. This combination makes it difficult to isolate the effect of incentives, but illustrates how they can operate in this environment.

8 Burgess, Simon; Propper, Carol; Ratto, Marisa and Tominey, Emma (2004) "Incentives in the Public Sector: Evidence from a Government Agency" CMPO.


UK; Australia; Netherlands.

Performance pay; Job Centre Plus.

There was a significant effect of performance pay in small and very small offices, with up to 10% increases in output on the quantitative measure (placements). The effect on qualitative measures is insignificant, due to fewer measurements and aggregate estimates rather than precise per individual/team measurements. The findings are consistent with the theory of free riding/peer-pressure, where small offices and teams will gain from peer monitoring and peer-pressure, whilst this trial was performed during one year in the equivalent of randomly selected offices. The control group was unselected offices with identical performance monitoring systems but without incentivised pay. The performance bonus that could be collected was up to 7.5% and allocated per team rather than individually.
| Working Paper Series No. 04/103 | surveys, job outcomes, administrative data, HR data and relevant macroeconomic data | larger offices suffer from free riding. The authors use the data to calculate costs and benefits of the incentive scheme and find that equivalent efficiency increases can be created as using traditional tools such as pay rise and increased staff numbers, but at a much lower cost. | This study looked at the comparative performance of incentivised sites, measured through 'star rating', which is an econometric model constructed to measure performance controlling for exogenous influences, and comparing this to site characteristics. Since low performers are not automatically given new contracts, the characteristics of high performing sites will probably increase. This is an example of a very well functioning benchmarking information system, where low-performers are dropped over time. |

| 9 | Cameron, M., Brown, K., Crooks, S., Young, N. (2006) Job Network Best Practice, Department of Employment and Workplace Relations | 2006 | Government report. Empirical study; multiple regressions of survey and national administrative data. Looking at quantitative and qualitative output and outcome measures of employment services. Evidence used: 30 case studies with survey data, extensive administrative data from multiple sources | Australia | Job Network (ESC 3) | Pay for Performance; Procurement Process; Performance Pay | The most successful sites within the Job Network scheme had strong focus on the Pay for Performance goals, using encouragement in office like ringing bells on successful outcome, weekly performance meeting and performance awards to staff. |

<p>| 10 | Considine, M. (2003). Governance and Competition: The Role of Non-profit Organisations in the Delivery of Public Services. Australian Journal of Political Science, 38(1), 63. | 2003 | Peer-reviewed journal. Case study; single case. Looking at outputs and outcomes of service delivery in relation to category of provider. Data used: focus group interviews, document analysis, survey data from three surveys (partly time series, n= 71, 365 and 187) | Australia | Job Network | Hybrid of Pay for Performance and Cost-reimbursement; Procurement Process | The author finds that the incentives introduced into service delivery drives a strong convergence between different types of agencies, and using (limited) time-series data of surveys can demonstrate a significant increase in caseloads and creasing behaviour in non-profits as service delivery strategies converge. Within years of implementation, many differences between private and non-profit agencies were no longer statistically significant. The conclusion from this would be that the increasing diversity of service provision envisaged has not materialised. | The distinctive features that could underlie the convergence are short contract terms, relatively high level of contractual regulation and a sole national purchaser of services. |</p>
<table>
<thead>
<tr>
<th>Provider incentives in social protection and healthcare</th>
<th>RAND Europe</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Reference</th>
<th>Study Type</th>
<th>Country</th>
<th>Provider Incentives</th>
<th>Payment Type</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dockery, A., &amp; Stromback, T. (2001). Developing public employment services: Preliminary assessment of the Australian experiment. International Labour Review, 140(4), 429.</td>
<td>Peer-reviewed journal. Review of reviews/evaluations. Looking at the outcomes, efficiency and effectiveness of the Job Network. Evidence used: discussion of previous reviews and evaluations</td>
<td>Australia</td>
<td>Job Network (ESC 1)</td>
<td>Pay for Performance; Procurement Process</td>
<td>The author finds that great efficiency savings have been achieved through pay for performance contracts and new modes of provision. However, significant parking has taken place where providers have been given flexibility to choose which services to provide. The author also argues that due to low outcome payments for those most difficult to place receive inadequate services.</td>
</tr>
<tr>
<td>Evaluation and Programme Performance Branch Research and Evaluation Group (2006), Customised Assistance, Job Search Training, Work for the Dole and Mutual Obligation - A Net Impact Study, Department of Employment and Workplace Relations, EPPB Report 1/2006</td>
<td>Government report. Empirical study: Net Impact Study, treated/untreated. The study looked at job outcomes of jobseekers. Evidence used: extensive administrative data</td>
<td>Australia</td>
<td>Job Network ESC 3 (Customised Assistance, Job Search Training, Work for the Dole and Mutual Obligation)</td>
<td>Pay for Performance; Procurement Process</td>
<td>The studied programmes were between 7 and 11% more effective in creating job outcomes. The study also concluded that there were significant deadweight losses associated with the programmes, as jobseekers that would have found jobs without intervention were a significant proportion.</td>
</tr>
</tbody>
</table>

These effects are very difficult to assign causality on the basis of this assessment, since it only looks at gross outcome results. Specifically it would be very difficult to assess the impact of incentives based on it.
<p>| 14 | Evaluation and Programme Performance Branch, Labour Market Policy Group. (2002) Job Network evaluation Stage three: effectiveness report, EPPB Report 1/2002, Department of Employment and Workplace Relations | 2002 | Government report. Empirical study. The report is very extensive, comprising a full evaluation of the Job Network. Data used: extensive administrative data, several surveys (&gt;10 surveys with n &gt; 1500), qualitative data, interviews | Australia | Job Network | Hybrid of Pay for Performance and Cost-reimbursement; Procurement Process | A majority of job-seekers were found to be more satisfied with service within Job Network than previously. A substantial proportion though expressed dissatisfaction because of lack of service. This is consistent with evaluation findings that the pay for performance incentive structure has made service providers very focused on securing outcomes for “easy” clients, thus parking harder-to-place clients. The study also notes significant deadweight losses arising from support given to clients likely to find jobs without support. The Job Network overall seemed to have created considerable efficiency gains. |
| 15 | Finn, D. (2007) Contracting out welfare to work in the USA: delivery lessons, DWP Report 466 (London/Sheffield: Department for Work and Pensions). | 2007 | Government contracted report. Review of reviews through discussion. The report looks at effects of contracting out welfare to work services with a specific focus on two cases. Evidence used: review of literature through discussion | U.S., New York; Wisconsin | Wisconsin Works | Hybrid Pay for Performance - Cost-reimbursement | In the W-2 programme, early incentive system designs resulted in adverse outcomes and large profits for delivery organisations. Contractual redesign, auditing and learning has amended these problems, and the market is now perceived to be working though proper evaluations and impact assessments are not available. |
| 16 | Griffiths, R., Durkin, S. (2007) Synthesising the evidence on Employment Zones, DWP Report 449 (London/Sheffield: Department for Work and Pensions). | 2007 | Government contracted report. Review of reviews through discussion. The report looks at the available evidence on effects of the introduction of Employment Zones. Evidence used: literature review, review of DWP administrative data, review of grey material | UK | Employment Zones | Pay for Performance | Employment Zones are significantly more effective in increasing rates of entry into work than comparable groups in government systems. All evidence points to financial incentives as having an overriding importance in achieving outcomes, confirmed by the fact that innovation within Employment Zones largely seems to mirror the development of the funding model. However, emphasis is given to the fact that incentivised funding works in conjunction with operational and financial flexibility. It is important to note that the value for money of the scheme does not seem to be higher than the comparison groups, indicating that increased effectiveness also comes at a higher price. |</p>
<table>
<thead>
<tr>
<th>Reference</th>
<th>Methodology</th>
<th>Data Sources</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>17 Hales, J., Taylor, R., Mandy, W. &amp; Miller, M. (2003) Evaluation of Employment Zones: Report on a Cohort Survey of Long-Term Unemployed People in the Zones and a Matched Set of Comparison Areas, National Centre for Social Research, DWP Report 176 (London/Sheffield: Department for Work and Pensions).</td>
<td>Government contracted report. Empirical study: cohort study using regression analysis, n &gt; 1000. Looking at the rates of entry into work and related indicators (i.e. quality of employment). Evidence used: cohort survey, administrative data, macroeconomic data</td>
<td>The evaluation shows that a private provision pay for performance scheme increased rates of entry into work by 8-11 % compared to a relevant control group in the government run Job Centre Plus. Participants in this scheme however seemed to have a lower qualification profile. The researchers see increased flexibility as the primary mechanism.</td>
<td>The distinctive environment in which Employment Zone payment for performance incentives operate are minimum regulation, customer choice and provider competition. This combination makes it difficult to isolate the effect of incentives, but illustrates how they can operate in this environment.</td>
</tr>
<tr>
<td>18 Hasluck, C., Elias, P. &amp; Green, A. (2003) The Wider Labour Market Impact of Employment Zones, Warwick Institute for Employment Studies, DWP Report 175 (London/Sheffield: Department for Work and Pensions)</td>
<td>Government contracted report. Empirical study: modelling, time-series regressions, multivariate analysis on micro-data. Looked at impacts of Employment Zone from micro and macro perspectives, i.e. effect on spells of unemployment, comparisons etc. Evidence used: administrative data, macroeconomic databases</td>
<td>Four different statistical tests are used that produce consistent results indicating that there is a significant positive effect of the programme on the rates of entry into work. The programme also seems to assist clients with shorter unemployment history better, and increase job retention compared to the control group.</td>
<td></td>
</tr>
</tbody>
</table>

- **Year**: 2007
- **Type**: Peer-reviewed journal. Case study: comparative case study of four different contract periods. Compares effects between different contractual periods in one state. Evidence used: document analysis, regression using administrative and macroeconomic data.

**U.S.; Wisconsin**

- **Program**: TANF; Wisconsin Works
- **Pay for Performance**

**Wisconsin's experience over time demonstrates significant difficulties in formulating contracts in the face of uncertainty, such as changing caseloads or budgetary situations. However, the study shows significant effect on service provision and priorities based on changes in the Performance for Pay system, such as adjusting weights of performance indicators to better suit political priorities. The study further notes the importance of investment in monitoring and auditing capacity, and also that significant organisational learning seems to take place. Regression analysis shows that there is a significant error component to outcomes not explained by performance indicators, indicating that using this tool effectively might be challenging.**


- **Year**: 2006
- **Type**: Government commissioned report. Empirical study: qualitative. Aims to deepen understanding of the effects of the programme. Evidence used: >250 interviews, document analysis, administrative data.

**UK**

- **Program**: Multiple Provider Employment Zones
- **Pay for Performance; User Choice**

**Incentives are passed on to staff, in evolving structures. In early EZ implementation, individual performance targets were often set, which seemed to be effective, but led to some gaming. Incentives have evolved into team targets, and diversified the Pay for Performance incentives into new areas such as reputation and marketing, the approach also included some process incentives. Evidence from interviews suggests that the introduction of competition has removed 'comfort zones' and rigidly that had started to arise in previous delivery, reinvigorating provision.**


- **Year**: 2002
- **Type**: Academic journal. Review through discussion. Summarises evidence on the Job Network with relation to mature workers. Evidence used: review of previous research.

**Australia**

- **Program**: Job Network
- **Hybrid of Pay for Performance and Cost-reimbursement; Procurement Process**

**The study argues that the pay for performance scheme used poses significant disadvantage to delivering employment services to mature workers since outcomes are harder to achieve amongst them. Simultaneously, getting more mature workers into work is a priority of the government - illustrating the possible unintended effects of pay for performance regimes.**

---

This describes the second stage of the Employment Zones scheme: Multiple Provider Employment Zone, where competition is being phased into the market.
<table>
<thead>
<tr>
<th>Provider incentives in social protection and healthcare</th>
<th>RAND Europe</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Koning, Pierre and Heinrich, Carolyn,</strong> Cream-Skimming, Parking and Other Intended and Unintended Effects of Performance-Based Contracting in Social Welfare Services. IZA Discussion Paper No. 4801</td>
<td>2010</td>
</tr>
<tr>
<td><strong>McConnell, S., Burwick, A., Perez-Johnson, I., Winston, P.</strong> (2003) Privatization in Practice: Case Studies of Contracting for TANF Case Management, Mathematica Policy Research, Inc.</td>
<td>2003</td>
</tr>
<tr>
<td><strong>Organisation for Economic Co-operation and Development.</strong> (2001), Innovations in labour market policies: the Australian way OECD, Paris, France</td>
<td>2001</td>
</tr>
</tbody>
</table>

---

**Notes:**
- **Provider incentives in social protection and healthcare** by RAND Europe
- **Organisation for Economic Co-operation and Development.** (2001), Innovations in labour market policies: the Australian way OECD, Paris, France
<p>| 26 | Productivity Commission (2002), Independent Review of the Job Network, Report No. 21, AusInfo, Canberra | 2002 | Government audit report. Empirical study: comparisons, regression, cross-tabulation etc. The study is an extensive (550 pages) evaluation of the full programme. Evidence used: administrative data, surveys, interviews, document analysis etc. | Australia | Job Network (ESC 1 &amp; 2) | Hybrid of Pay for Performance and Cost-reimbursement | The overall effect of the programme seems to have been increased cost-efficiency but effectiveness is largely the same. There are weaker incentives to provide help to those who are less responsive to assistance with the outcome-based incentives, causing problems with parking. The high level of commencement fees limits the incentives for providers in achieving additional outcomes (apart for those that would have happened anyway). A large part of the provider incentives in this model seems to come from the 'star rating' model, where providers are benchmarked against each other, and which informs the next round of tenders. Jobseekers perceived a very differing level of service, explained by the pay for performance structure causing parking and creaming behaviours. | The commission identifies six primary advantages with provider incentives through pay for performance: forces clarification of objectives, better focus on those objectives, greater flexibility, quantifiable results, avoids disipation or activity that wouldn't yield outcomes, easier to monitor and shifting risk to private sector. It also identifies some downsides: inequality through parking/creaming, possible focus on short-term payable outcomes, loss of accountability for outcomes, complex payment structures, need to define default provisions, complicated contracts and poor information about what works. It also notes that a higher outcome focus would cause serious liquidity problems for the industry that would be hard to overcome. |
| 27 | Struyven, L., Steurs, G., (2005) Design and redesign of a quasi-market for the reintegration of jobseekers: empirical evidence from Australia and the Netherlands, Journal of European Social Policy 0958-9287; Vol 15(3): 211–229 | 2005 | Peer-reviewed journal. Theoretical study using quasi market framework. The study looks at observed outcomes in relation to theory. Evidence used: review of previous research | Netherlands, Australia | UWV, Job Network | Pay for Performance; Hybrid of Pay for Performance and Cost-reimbursement; Procurement Process | Results so far have been modest. In Australia, large cost-efficiency gains have been made but effectiveness is more or less the same. There has been erosion of the level of service, and incentives have led to short-term placements at the expense of lengthier training. The Dutch system is no more effective than in the past and no evidence of increase responsiveness exists. Programme redesigns suggest that unintended consequences can occur, for example concerning training, when incentives for that are weak. | Significant institutional differences exist in the two systems, influencing the structure of supply and demand in the system, as well as its implementation. The Australian system tends towards a holding-company led by central government which strives towards retaining strong control over providers. This structure in conjunction with the fact that the first years of experience have yielded problems with risk selection (creaming and parking), means that Australia is moving towards more regulations and standards. The Netherlands have a variation in principals, and a different view of the market, seeing themselves as enablers of a functioning market. This in conjunction with the fact that regulations have led to little innovation in the first years have caused the Netherlands to move towards a straight pay for performance regime. The author demonstrates that implementation is highly contextual and dependent on previous experience and structure. |</p>
<table>
<thead>
<tr>
<th>No.</th>
<th>Source</th>
<th>Year</th>
<th>Description</th>
<th>Country</th>
<th>Program</th>
<th>Incentive Model</th>
<th>Incentive Focus</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>28</td>
<td>Department of Employment and Workplace Relations., (2008), The Future of Employment Services in Australia: A Discussion Paper.</td>
<td>2008</td>
<td>Government discussion paper. Review of evaluations; stakeholder input. The paper is the result of consultation on experiences with the Job Network. Evidence used: previous evaluations, stakeholder input.</td>
<td>Australia</td>
<td>Job Network</td>
<td>Hybrid of Pay for Performance and Cost-reimbursement; Procurement Process</td>
<td>The discussion paper identifies several incentive problems in the Job Network that are to be addressed in reorganisation of the scheme. These are primarily the widespread creaming in the system, that is to be tackled with more targeted incentives to those hard-to-place. It also identifies significant underinvestment in education and skills enhancement, causing negative externalities in lower productivity and sub-optimal outcomes. This is to be addressed through an added bonus incentive-system for outcomes reached after accredited training. A lack of incentives for employer focus is also identified, that is to be addressed through higher outcome payments for provider brokered outcomes.</td>
<td>This discussion paper is the result of a consultation process with all stakeholders and is the formation for Australia's new employment services programme.</td>
</tr>
</tbody>
</table>
## Appendix E: Annotated bibliography on provider incentives in healthcare

<table>
<thead>
<tr>
<th>No</th>
<th>Reference Details</th>
<th>Year</th>
<th>Type of Publication</th>
<th>Country</th>
<th>Types of Incentives</th>
<th>Effect of Incentives</th>
<th>Other Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Epstein, A., &quot;Performance Measurement and Professional Improvement,&quot; in Performance Measurement for Health System Improvement, Smith, P., E. Mossialos, I. Papanicolas, and S. Leatherman, eds. Cambridge, UK. Cambridge University Press, 2009, pp. 613-640.</td>
<td>2009</td>
<td>Review of reviews, single and complex interventions; over 50 studies with various study designs; single and complex interventions</td>
<td>Developed Countries (particularly US, England and Australia)</td>
<td>Education, guidelines, outreach; Audit, practice profiling and feedback; Accreditation and recertification; Public reporting; P4P; Quality measurement</td>
<td>Passive education approaches are largely ineffective and unlikely to change physician practices significantly (with respect to quality); passive dissemination of guidelines has little impact on quality improvement; active educational activities such as interactive workshops, outreach visits or educational sessions by charismatic opinion leaders show positive outcomes (as to quality) but effectiveness varies; multifaceted interventions are more effective than single interventions (with respect to quality); Early studies found that audit with feedback was neither a consistent nor a particularly effective intervention/incentive for improved performance (with respect to quality), later studies showed that audit and feedback on performance has a small to moderate impact on quality, greater improvements when baseline adherence to recommended practice is low and when feedback is provided more intensely; lack of information about effectiveness of accreditation and recertification; public reporting has variable success and some areas (e.g. mental health) have proved intransigent, there are concerns about unintended responses such as false reporting, overlooking aspects of care which are not in the spot-light, inequality of and access to care; P4P - mainly financial incentives - for UK; mixed results with respect to effectiveness, no strong positive effect on quality of care, unintended effects such as adverse selection and improved documentation; for US: equally inconclusive results with monetary incentives provided to hospitals not resulting in a substantial difference in quality; P4P - overall, lack of conclusive data on effectiveness with multiple factors having an impact on the success of efforts to drive improvement with financial incentives (factors include: nature of clinical conditions, size and form of incentive programme, time lag between the initiation of the programme and the measurement of care)</td>
<td>Primary and Secondary Care; general and with some focus on care for specific (chronic) conditions</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Year</th>
<th>Review of reviews, systematised but not systematic; over 50 studies with various study designs; single and complex interventions</th>
<th>Developed Countries (Particularly US, England, Germany, the Netherlands)</th>
<th>Various, including financial and non-financial, rewards and penalties, size and timing, performance thresholds</th>
<th>Health plan to provider organisation: Fund holding in the UK has not impacted on specialist referral or hospital admission rates among general practitioners but produces consistent reductions in drugs per prescription; the shift from fee-for-service to fund holding led to fewer referrals for elective surgery and to private clinics; capitation payments reduced the number of hospital days by up to 80%; General incentives: fee-for-service providers saw their patients more, more often and patients had fewer emergency room visits; for children: compared to fee-for-service, capitation patients had fewer primary care visits, fewer visits to non-primary care office-based specialists and fewer emergency visits; for elderly: capitation was linked to fewer physician visits and in-patient stays and marginally better self-reported general health and well-being; for GPs in Copenhagen, a change from capitation to mixed fee-for-service/capitation payment resulted in only temporary rise in face-to-face consultations while referrals to specialist and hospital admissions declined, and rate of diagnostic and curative services increased; a change from fee-for-service to capitation was not found to result in significant changes in hospital utilisation rates.</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Physician group-level incentives: evidence is mixed on the effects of quality incentives on physician behaviour; positive effects of explicit quality incentives were found on flu immunisation rates, cognitive service interventions by pharmacists; studies also showed no statistically significant effect of group bonuses on cancer screening for women aged fifty and more or on paediatric immunisation and well-child visit rates; potential of bonus distribution and return of withholds was associated with increased development of medical management programmes and improved process of diabetic care but with no impact on paediatric asthma care; UK QOF: on balance, performance incentives were related to a modest increase in the improvement rate of quality of care.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Insurers to providers; individual physician; within-group physician; within hospital</td>
</tr>
</tbody>
</table>
Hospital-based incentives: studies based on comparison between public reporting and public reporting with a pay-for-performance component show that P4P hospitals improved significantly more on process measures for acute myocardial infarction, heart failure, pneumonia and a composite measure of performance; baseline performance was inversely associated with improvement; comparison of care received by acute myocardial infarction patients of US hospitals participating in voluntary quality improvement initiatives and hospitals not participating revealed no significant differences in overall improvement between the incentive and control hospitals. Incentive hospitals were found to adhere better to two dimensions of care: aspirin at discharge and smoking cessation counseling; no treating to the test was found in incentivised hospitals. Other work on the effect of P4P compared to public reporting hospitals found weak effects of the incentives programme with only significant differences established among congestive heart failure patients and no significant differences for those with acute myocardial infarction or pneumonia; an evaluation of the cost effectiveness of a voluntary incentive programme adopted by Blue Cross Blue Shield of Michigan on heart care outcomes showed savings of roughly $12,000 to $30,000 in costs per QALY.

Individual physician-based incentives: stepped bonus improves children's immunisation status but enhanced fee-for-service incentive shows no significant effect; both were found to provide positive effects in a later study; sexually active females were screened more often by physicians receiving salary in conjunction with a quality of care incentive than those paid in other ways; annual bonus for exceeding predetermined target in medical care for diabetes patients were found to improve care; absolute and percentage improvements in care process were inversely related to baseline performance; a study of the effect of a 5% withhold for primary care physicians caring for patients with diabetes revealed that the programme had an impact on eye examination only in the first year and no significant differences between treatment and control were found after that.

Unintended consequences: cream-skimming in treating substance abuse, increased documentation rather than increase in preventive care, various gaming routines: miscoding of diagnoses, making up of discharge diagnoses, cream-skimming in hospitals.

Nature of incentives: both negative and positive rewards induce provider responses in the expected direction; health plans tend to favour positive incentives in lieu of penalties in their P4P programmes.
<p>| | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Target entity: the level which exerts more powerful effects on performance is ambiguous; overall, where system failure is the major source of quality problems, group incentives would be expected to dominate over those for individuals.</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Type of incentive: studies find that behaviour is influenced by general payment system-level incentives (FFS, per case and capitation), selective P4P and indirect incentives of public reporting; system-level capitation incentives produce somewhat lesser use of hospital and other expensive resources than do indemnity payments based on FFS.</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Extrinsic incentives and effects on intrinsic motivation: weak evidence that extrinsic rewards crowded out intrinsic motivation; also, non-measurable domains of clinical quality are found not to decline even as certain rewarded types of performance improve.</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Nature of behaviour subject to incentives: process measuring was the core of first generation P4P; that is no longer the case as outcome measures have become standard; blending outcome measures and outcome and process incentives is the current norm; absolute performance thresholds are more favoured by physicians.</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Size of incentive: no particular size; 2% to 9% is the range which is assumed to provide statistically significant but modest change in provider behaviour.</td>
</tr>
<tr>
<td>2</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Certainty, frequency and duration of incentive: incentives should be predictable and endure over a timeframe sufficiently long to prompt providers to make sustained investments in improved clinical infrastructure and care process.</td>
</tr>
<tr>
<td>3</td>
<td>Shekelle, P., &quot;Public performance reporting on quality information,&quot; in P. Smith et al., eds., Performance Measurement for Health System Improvement, Cambridge, UK: Cambridge University Press, 2009, pp. 537-551.</td>
<td>2009</td>
<td>Review of reviews, systematised but not systematic; over 40 studies of single and complex interventions</td>
<td>US (predominantly)</td>
<td>Public reporting</td>
</tr>
</tbody>
</table>

<p>| 4 | Campbell, S., D. Reeves, E. Kontopantelis, B. Sibbald, and M. Roland, &quot;Effect of Pay for Performance on the Quality of Primary Care in England,&quot; The New England Journal of Medicine, Vol. 361, No. 4, 2009, pp. 368-378. | 2009 | Empirical analysis (interrupted time series analysis); 42 representative family practices (about 200 people in each practice); single intervention | England | P4P based on targets for quality | P4P accelerated improvements in quality for asthma and diabetes in the short term but not for heart disease. However, once targets were reached, the improvement in the quality of care for patients with all three conditions slowed, and the quality of aspects of care for asthma and heart disease that had not been linked to incentives decreased. Continuity of care was reduced after the introduction of the scheme. | Primary Care; care of patients with asthma, diabetes, or coronary heart disease; patients' perceptions of access to care, continuity of care, and interpersonal aspects of care |</p>
<table>
<thead>
<tr>
<th>Provider incentives in social protection and healthcare</th>
<th>RAND Europe</th>
</tr>
</thead>
</table>

| 5 | Busse, R. and N. Mays, "Paying for Chronic Disease Care," in Caring for People with Chronic Conditions. A Health System Perspective., Nolte, E. and M. McKee, eds. Maidenhead, UK: Open University Press, 2008, pp. 195-221. | Theoretical review of incentives schemes; non-systematic review of work on reimbursement mechanisms; examples from practice; single and complex interventions | UK and US | Financial (P4P, quality based purchasing, QOF) | US Quality-based purchasing: mixed results with bulk of the work reviewed focusing on preventive care; no consistent relationship between magnitude of incentive and response was found; On balance, FFS was found to provide positive results while bonus payments were found to have no effect on quality; Generally, incentives to achieve performance were found to be more effective when indicator to be followed required less patient cooperation (e.g. receiving vaccinations or answering questions about smoking) than when significant patient cooperation was needed (e.g. quit smoking); Overall, few unequivocal data on which to base a quality-based purchasing strategy but some evidence that both payment and reputational incentives can work; US P4P: Generally positive findings on P4P at the level of individual hospitals, insurers’ programmes or large, integrated health care delivery networks; often, it is impossible to conclude which aspect of the intervention created the advantage over the comparators and no information is available about the cost-effectiveness of the intervention; cream-skimming was identified as a potential problem while there was no evidence on whether once incentives are removed the effects of intervention persist; Important issues for P4P: financial incentives for quality have to be designed with care as there is evidence of ‘pervasive responses to incentives; incentive schemes have to have clearly defined objectives; data collection may be costly and data itself may be inaccurate; incentive programmes may be costly as providers may expend considerable resources in their attempts to earn incentive payments; choice of performance measures have to be done with care, although no clear rule is available; continuous incentives may be more influential than for example an end-of-year bonus; financial incentives alone do not work and have to be combined with other interventions targeting quality; although the effect of financial incentives tends to be smaller at the provider level and somewhat larger at the level of individual professionals, incentives at team level may be a good compromise (e.g. multidisciplinary teams tend to produce better outcomes); further, US VHA evidence indicates that provider group incentives without large monetary incentives for physicians can be effective in the presence of rigorous performance monitoring and benchmarking between provider groups. | Chronic disease care; Primary and secondary care; chronic care |

<p>| 5 | - | - | - | - | - | UK QOF: Evidence suggests that for the quality of care for three common conditions in the QOF (angina, diabetes and asthma) quality (irrespective of value for money) continued to improve after the introduction of QOF and at a faster rate for asthma and diabetes than in the period prior to the intervention; improvement was found to be statistically significant in areas of performance linked to incentives in QOF and for those areas that were not, only a small causal effect of QOF on quality was established; However, some evidence of inequality of care between deprived and non-deprived areas was found, though these differences were small in magnitude and there was evidence of a catch-up over time. | Chronic disease care; Primary and secondary care; chronic care |</p>
<table>
<thead>
<tr>
<th>Page</th>
<th>Study Title</th>
<th>Year</th>
<th>Study Design &amp; Description</th>
<th>Country(s) &amp; Intervention Type</th>
<th>Key Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>Sturm, H. A., Austvoll-Dahlgren, M. Aaserud, A. D. Oxman, C. Ramsay, A. Vernby, and J. P. Kösters, “Pharmaceutical Policies: Effects of Financial Incentives for Prescribers,” Cochrane Database of Systematic Reviews No. 3, 2007.</td>
<td>2007</td>
<td>Systematic review of empirical studies; 13 studies; single and complex interventions</td>
<td>UK, Ireland, Germany (Financial (drug budgets))</td>
<td>UK: Real drug expenditure of fund holders relative to the expected level of expenditure was lower than that of non-fund holders. A similar effect was found for measures such as drug expenditure per patient and total prescribing cost. Further, fund holders were likely to prescribe less drugs than non-fund holders with the magnitude of this effect decreasing over time; use of generic drugs was higher in fund holders but not statistically significantly different from non-fund holders; use of more expensive drugs for gastric ulcer and depression was relatively lower in fund holders; fund holders had lower relative referral rate to NHS outpatient care; Ireland: absolute decrease in prescribing cost but not statistically significant change after introduction of indicative budgets; a relative reduction in number of prescribed items was established for intervention group; Germany: drug use decreased after the introduction of the German drug budget; inconclusive results as to effect on health care utilisation.</td>
</tr>
<tr>
<td>7</td>
<td>Rosenthal, M. and R. A. Dudley, “Pay-for-Performance: Will the Latest Payment Trend Improve Care?,” JAMA, Vol. 297, No. 7, 2007, pp. 740-744.</td>
<td>2007</td>
<td>Non-systematic review and discussion; use of real-life examples; single and complex interventions</td>
<td>US (P4P (Five key design elements))</td>
<td>Pay-for-Performance as Individual vs. Group Motivator: for performance issues that can be improved most efficiently through group action (e.g., by adding an information system that improves prescribing accuracy for all providers), incentives should be directed towards the group. For behaviours under the individual physician’s control, such as counselling on smoking cessation and documentation, incentives may be more effective when targeted at individuals; a varying mix of incentive approaches, combined with other management techniques such as feedback, will probably be optimal; Paying the Right Amount: No specific dollar amount or percentage will be the right amount for every circumstance; Selecting High-Impact Performance Measures: providers should seek a central role in deciding what is measured; Making Payment Reward All High-Quality Care: To be a more effective lever for change, however, P4P programmes should be recast from a programme that rewards “top quality” providers to one that rewards high value care, provided by anyone. Payers could do this by paying all providers an additional fee for each appropriately managed patient or for each recommended service. With this approach, every provider has an incentive to deliver the best care to each patient seen; Prioritizing Quality Improvement for Underserved Populations: One approach could be to offer larger incremental payments for providing high-quality care to populations that are disadvantaged or more costly to treat effectively. Alternatively, capital grants, technical assistance, or special training (e.g., in cultural competence) could be provided to hospitals and physicians who treat disadvantaged patients under P4P contracts.</td>
</tr>
</tbody>
</table>

*British fund holders; German drug budgets; effects on drug expenditure, drug utilisation and healthcare utilisation; effect on quality of care; savings; health outcomes*
<table>
<thead>
<tr>
<th>#</th>
<th>Author(s)</th>
<th>Year</th>
<th>Study Design</th>
<th>Country</th>
<th>Intervention</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>Lindenauer, P. K., D. Remus, S. Roman, M. Rothberg, E. Benjamin, A. Ma, and D. Bratzler</td>
<td>2007</td>
<td>Empirical analysis (based on measures of quality at 613 hospitals over 2 years); Single intervention</td>
<td>US</td>
<td>PR and P4P</td>
<td>Hospitals engaged in both public reporting and pay for performance achieve modestly greater improvements in quality than did hospitals engaged only in public reporting; P4P hospitals showed greater improvement in all composite measures of quality, including measures of care for heart failure, acute myocardial infarction, and pneumonia and a composite of 10 measures. Baseline performance was inversely associated with improvement.</td>
</tr>
<tr>
<td>9</td>
<td>Frolich, A., J. Talavera, P. Broadhead, and R. A. Dudley</td>
<td>2007</td>
<td>Theoretical study accompanied by a systematic review of 21 papers (involving 18 trials; only a total of 9 RCT trials were included in the discussion); Review of single interventions</td>
<td>Developed Countries</td>
<td>Financial and non-financial incentives</td>
<td>Structure of Incentive: By comparing graduated bonus payment to enhanced FFS, with approximately the same total maximum income potential for providers in each group, Fairbrother et al found no difference between bonuses and FFS, though both were superior to receiving no incentive; Patient factors: incentives to achieve performance are more effective when the indicator to be followed required less patient cooperation (e.g. answering inquiries about smoking) than when significant patient cooperation is needed; Reputational incentives: Hospitals whose performance was publicly reported were more likely to adopt quality improvement programmes than those receiving confidential reports or no report. The impact of PR was especially large among those hospitals whose performance was worse than average.</td>
</tr>
</tbody>
</table>

Hospitals (secondary care); comparison between PR with P4P and PR only; comparison over measures of care for heart failure, acute myocardial infarction, and pneumonia and a composite of 10 measures.
<p>| 10 | Damberg, C., M. Sorbero, A., Mehrotra, S., Teleki, S., Lovejoy, and L. Bradley, An Environmental Scan of Pay for Performance in the Hospital Setting: Final Report, Santa Monica, Calif.: RAND Health, 2007. | 2007 | Empirical Study including a comprehensive review of the literature (theoretical and empirical) and key informant discussions and; 9 empirical studies on hospital incentivisation were used; single and complex interventions | US | P4P | P4P. The strongest evidence on the impact of hospital P4P to date has been shown through the Lindenauer (2007) study of the impact of PHQID relative to the Medicare RHQDAPU programme. These studies, while showing a positive effect of P4P, reveal that the additional effects of P4P are somewhat modest relative to public reporting and other quality interventions that are occurring simultaneously. Improvements in hospital performance have been observed in response to feedback reports and public reporting, with a financial incentive for submitting data; it is argued that to accomplish sustained quality improvement, interventions should be multifaceted and focus on different levels of the health care system. This suggests that to be most effective, P4P should be coupled with other activities such as public reporting and internal quality improvement activities that also encourage quality improvement for the same clinical area; P4P and patient outcomes: less evidence, with studies either not reporting statistical significance of results or finding no statistically significant differences; Findings from key informant discussions: Measures: alignment of measures for reporting efforts to be coordinated and use of evidence-based standardised measures to minimise physician push-back; use of limited set of measures on which hospital can focus quality improvement efforts; Payment Structure: support for use of absolute thresholds and reward all good performers; preference for multilevel incentives; Data infrastructure: technical support and shift to electronic medical recording; Public reporting: hospitals pay attention to how their institution looks publicly and that public reporting has forced their boards to monitor quality and provide resources for quality improvement more closely; Barriers and Challenges to P4P: Small numbers problem: A sizeable number of hospitals have only a small number of events or cases to report for one or more measures. A small number of events to score results in unstable estimates of performance as a basis for determining performance-based incentive payments. While this is a more acute problem for small and rural hospitals with a small number of patients per year, the problem also occurs in some medium- and large-size hospitals depending on their service mix, the details of measure specifications, and the use of sampling during data collection; Burden of data collection; Ensuring the validity of data used to make differential Payments. | Hospitals (Secondary Care); financial incentives in inpatient or outpatient care; effect of paying for quality, patient experience, and safety or resource use |</p>
<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>Study Type</th>
<th>Location</th>
<th>Financial Incentives</th>
<th>Financial Incentives directed at improving quality: Relationship between financial incentives and quality of care, where the incentives were intended to improve quality: In some studies there was improvement in selected outcome measures; however, in many of the studies financial incentives are combined with other quality improvement efforts, making it impossible to determine the contribution of the incentives in achieving the reported results; Even where there is relatively more research to draw on, such as the use of financial incentives to improve the delivery of preventive services by physicians, the few studies with strong research designs find small, if any, effects on payments to providers; Payer initiatives that reward providers for quality improvements or the attainment of quality benchmarks: mixed results; few significant impacts are reported, and it is often the case that payer programmes included components in addition to incentive payments, again making it difficult to assess the independent effect of the financial incentives; Financial incentives and secondary impacts on quality: no consistent evidence that financial incentives employed to reduce utilisation have a negative impact on quality; however, these findings are influenced by data availability and study designs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roland, M., Comment on 'Financial Incentives, Healthcare Providers and Quality Improvements'. As of 26 May 2010: <a href="http://www.health.org.uk/publications/commentary.html">http://www.health.org.uk/publications/commentary.html</a></td>
<td>2008</td>
<td>Commentary and limited review of single and complex interventions</td>
<td>Primary and secondary care</td>
<td>P4P is not a panacea: Many evaluations of P4P show small or negligible impacts of financial incentives. The context in which incentives are introduced may be very important, and detailed context needs to be reported in evaluations of P4P schemes; P4P should probably be regarded as one of a range of interventions that need to be considered as part of a package of quality improvement measures; P4P un incentivised aspects: the impact of P4P on un incentivised aspects of quality is generally not reported, and this should be part of the routine monitoring of P4P schemes in order to guard against unexpected or unintended consequences of financial incentives; Financial incentives to change other behaviours (e.g. reduce hospital utilisation by changing patterns of referral) could have a negative impact on quality. Although there is little evidence that this has occurred, this is largely because the data are not usually collected; More recent findings: P4P impact on health outcomes: mixed and condition specific: QoF scores are not associated with emergency admissions for coronary heart disease but are negatively associated for epilepsy; Impact on health inequalities: statistically significant but small differences in QoF scores between practices in affluent and deprived areas; larger relative differences for some indicators; Impact on un incentivised aspects of care: no indications of such problem, however lack of sufficient data.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Primary and secondary care</td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Methodology</td>
<td>Year</td>
<td>Location</td>
<td>Incentive Type</td>
<td>Findings</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>------</td>
<td>----------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>Christianson, J., &quot;Evaluating Pay-for-Performance in Medicaid through Real-World Observation,&quot; Health Affairs, Vol. 26, No. 4, 2007, pp. 528-531.</td>
<td>Non-systemic review of qualitative evidence accompanied by discussion of emerging issues</td>
<td>2007</td>
<td>US</td>
<td>P4P</td>
<td>P4P Gaming: Statistically significant evidence that physicians who excluded greater proportions of patients had higher performance scores and consequently received more reward money; P4P Success Factors: factors such as effective communication with physicians and need to consider the initial conditions of physicians when designing rewards; physicians in resource-constrained practices serving low-income patients may have difficulty responding to financial incentives and improving their performance even when payment is based on simple and widely accepted performance measures.</td>
</tr>
<tr>
<td>Young, G. J. et al., &quot;Effects of paying physicians based on their relative performance for quality,&quot; Journal of General Internal Medicine, Vol. 22, No. 6, 2007, pp. 872-876.</td>
<td>Empirical study; retrospective cohort study using pre/post analysis; outcomes: adherence to 4 diabetes performance measures; sample size:334 primary care physicians</td>
<td>2007</td>
<td>US (New York state)</td>
<td>Bonus and withholds</td>
<td>The study found that the absolute performance levels of physicians increased across all measures of performance immediately following implementation of the new system of incentivisation. However, there was no difference between the post- and pre-intervention trends which suggested that the overall increase in performance was largely a result of an underlying trend. The study also found evidence of a 7 percentage point improvement in physician adherence for eye examination in the year immediately following the year of implementation of the programme. This one-time increase was attributed to the new incentives mechanism.</td>
</tr>
<tr>
<td>Petersen, L. A., L. D. Woodard, T. Urech, C. Daw, and S. Sookanan. &quot;Does Pay-for-Performance Improve the Quality of Health Care?&quot;, Annals of Internal Medicine, Vol. 145, No. 4, 2006, pp. 265-272.</td>
<td>Systematic review of single intervention empirical studies; 17 studies were selected (randomized, controlled trials and controlled before-and after studies)</td>
<td>2006</td>
<td>US</td>
<td>Financial incentives</td>
<td>Financial Incentives at the Provider Group Level: Evidence of a predominantly positive effect of financial incentives directed at provider group level along several dimensions of quality, although the size of the effect was small (statistically significant effects were found for cervical cancer screening while improvements in mammography screening rates and haemoglobin A1c testing were not statistically significant); Financial Incentives at the Physician Level: Partial or positive effects of incentives directed at individual physicians; evidence is based on immunisation and considered effect of bonus and feedback, enhanced fee-for-service and feedback, and feedback only incentive schemes; highest improvement was found in the bonus and feedback group; Unintended Effects of Incentives: adverse selection; gaming the system; Design of incentives: Have to be designed carefully; documentation, rather than actual use of the preventive service, improved statistically significantly with a financial incentive; adverse selection may have occurred with performance-based contracting in settings where providers can avoid sicker patients; combined incentives for both overall improvement and achievement of a threshold should be considered; basing incentives on the combination of a process-of-care measure (for example, documentation of smoking cessation advice) and the outcome of interest may prove beneficial; size of bonus may be important but no clear rule on how much it should be; remains unclear whether financial incentives are cost-effective.</td>
</tr>
<tr>
<td>Year</td>
<td>Study Title</td>
<td>Type of Study</td>
<td>Country</td>
<td>Payment Scheme</td>
<td>Findings</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>---------------</td>
<td>---------</td>
<td>----------------</td>
<td>----------</td>
</tr>
<tr>
<td>2006</td>
<td>&quot;Incentives in Health Care: The Shift in Emphasis from the Implicit to the Explicit,&quot; in Human Resources for Health in Europe, Dubois, C., M. McKee, and E. Nolte, eds. Maidenhead, UK: Open University Press, 2006, pp. 140-154.</td>
<td>Theoretical review of reimbursement schemes</td>
<td>Developed Countries, Europe</td>
<td>Implicit (e.g. trust) and Explicit (e.g. regulation, payment)</td>
<td>Explicit incentives; FFS schemes may induce unnecessary activity and in poorly managed systems where efficient self-regulation and budget caps are missing may distort activity and expenditure. In some countries, FFS is coupled with a budget cap. Under FFS it is likely that what is not incentivised is marginalised. Capitation does not induce over-activity but may create under-activity, cream-skimming and cost shifting from primary to secondary care. Capitation salary payments facilitate the maintenance of expenditure control but they also may induce undertreatment and shifting of costs. Mixed payment systems are quite common and their aim is to provide improved performance while limiting the downsides of each payment scheme. Career pay scale is another incentive to practitioners.</td>
</tr>
<tr>
<td>2006</td>
<td>&quot;Cost-Effectiveness of Hospital Pay-for-Performance Incentives,&quot; Med Care Res Rev, Vol. 63, 2006, pp. 49S-72S.</td>
<td>Empirical study; cost-effectiveness analysis of incentive system for adherence to heart-care-related clinical guidelines in 85 Michigan hospitals</td>
<td>US</td>
<td>P4P; direct financial incentive as an add-on to the diagnosis-related group reimbursement</td>
<td>Hospitals are evaluated on (1) the provision of aspirin orders at discharge for patients diagnosed with acute myocardial infarction, (2) the prescription of beta-adrenergic blockers at discharge for patients diagnosed with AML, and (3) the prescription of angiotensin converting enzyme inhibitors at discharge for patients diagnosed with congestive heart failure. Hospital performance is measured as the proportion of all eligible patients—not just patients with Blue Cross Blue Shield of Michigan policies—that receive a particular treatment. The study estimated that 24,418 patients received improved care, resulting in a range of QALYs from 733 to 1,701, depending on assumptions about clinical effectiveness. Cost per QALY was found to be between $12,967 and $30,081, a level well under consensus measures of the value of a QALY.</td>
</tr>
<tr>
<td>2006</td>
<td>&quot;Physician Pay-for-Performance. Implementation and Research Issues,&quot; J Gen Intern Med, Vol. 21, 2006, pp. S9-S13.</td>
<td>Theoretical study including limited supporting empirical evidence; primarily a discussion study</td>
<td>US</td>
<td>P4P</td>
<td>P4P and Physicians: Concerns and Challenges: Measures that are the easiest and least costly to observe tend to be included as P4P benchmarks. However, behaviours that are easy to monitor are not necessarily those that will yield the greatest improvements in health. Cost of Implementation: costly monitoring is required under any payment arrangement that seeks to reward specific behaviours; also there are costs associated with infrastructure to achieve P4P goals; Mixed Messages: multiple different P4P initiatives can muddle physician incentives and increase physician reporting costs; Variation versus Standardization: “one size fits all” P4P programmes are not likely to be efficient in accomplishing change. Payments will be more than needed for some physicians and too small to affect the behaviour of others; Individual or Group-Level Incentives: level of incentivisation should depend on the size and organisational characteristics of the practice.</td>
</tr>
</tbody>
</table>

Primary and secondary care
<table>
<thead>
<tr>
<th></th>
<th>Authors</th>
<th>Title</th>
<th>Year</th>
<th>Study Design</th>
<th>Country</th>
<th>Setting</th>
<th>Key Findings</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>19</td>
<td>Rosenthal, M.et al.,</td>
<td>&quot;Pay for Performance in Commercial HMOs,&quot;</td>
<td>2006</td>
<td>Empirical paper; descriptive and regression (logistic) regression; sample: 252 HMOs</td>
<td>US</td>
<td>P4P</td>
<td>The study found that pay for performance is used in more than 50% of the HMOs, representing more than 80% of persons enrolled. Almost 90% of the 126 health plans with pay-for-performance programmes had programmes for physicians and 38% had programmes for hospitals. Geographic region, use of primary care providers (PCPs) as gatekeepers, use of capitation to pay PCPs, and whether the plans themselves received bonuses or penalties according to performance was significantly associated with geographical characteristics.</td>
<td>Primary care (HMOs)</td>
</tr>
<tr>
<td>20</td>
<td>Beaulieu, N. D., and D.R. Horrigan,</td>
<td>&quot;Putting Smart Money to Work for Quality Improvement,&quot;</td>
<td>2005</td>
<td>Empirical study; pre/post design for 476 diabetes patients whose physicians participated in the performance pay program; 21 physicians</td>
<td>US (upstate New York)</td>
<td>Bonus and other management tools (registry, group meetings)</td>
<td>&quot;In 2001, a managed care organization in upstate New York designed and implemented a pilot programme to financially reward doctors for the quality of care delivered to diabetic patients. In addition to paying a performance bonus, physicians were also supplied with a diabetic registry and met in groups to discuss progress in meeting goals for diabetic care. Primary data on diabetes care at the patient level were collected from each physician during the 8-month period, April 2001–January 2002. Physicians and patients achieved significant improvement on five out of six process measures, and on two out of three outcome measures (HbA1c control and LDL control). 13 out of 21 physicians improved their average composite score enough to earn some level of financial reward. Of the eight physicians not receiving any of the three levels of reward, six improved their composite scores. Financial incentives for physicians, bundled with other care management tools, led to improvement on objectively measured quality of care for diabetic patients. Self-selection by physicians into the pay pilot and the small sample size of participating physicians limit the generalisability of the results.&quot;</td>
<td>Managed care organisation</td>
</tr>
<tr>
<td>21</td>
<td>Gagne, M. and E. Deci,</td>
<td>&quot;Self-Determination Theory and Work Motivation,&quot;</td>
<td>2005</td>
<td>Theory-focused paper without empirical application</td>
<td>N.A.</td>
<td>Extrinsic versus Intrinsic Incentives</td>
<td>Cognitive evaluation theory, which explains the effects of extrinsic motivators on intrinsic motivation, received some initial attention in the organisational literature. However, the simple dichotomy between intrinsic and extrinsic motivation made the theory difficult to apply to work settings. Differentiating extrinsic motivation into types that differ in their degree of autonomy led to self-determination theory, which has received widespread attention in the education, health care, and sport domains. This article describes self-determination theory as a theory of work motivation and shows its relevance to theories of organisational behaviour.</td>
<td>N.A.</td>
</tr>
<tr>
<td>22</td>
<td>Jack, W.,</td>
<td>&quot;Purchasing health care services from providers with unknown altruism,&quot;</td>
<td>2005</td>
<td>Theoretical study; explores conditions for optimum design of medical contracts; no empirical application</td>
<td>N.A.</td>
<td>Incentives resulting from mechanism of reimbursement</td>
<td>This paper derives conditions for the optimal nonlinear cost-sharing mechanism in the presence of asymmetric information about altruism and concludes that if contracts cannot be designed explicitly to consider provider characteristics, purchasers need to develop incentive mechanisms that induce providers to reveal their types.</td>
<td>N.A.</td>
</tr>
<tr>
<td>References</td>
<td>Title</td>
<td>Type</td>
<td>Country</td>
<td>Incentives</td>
<td>Consequences</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>-------</td>
<td>------</td>
<td>---------</td>
<td>------------</td>
<td>--------------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kane, R.L., P.E. Johnson, R.J. Town, and M. Butler, Economic Incentives for Preventive Care. Summary, Rockville, MD: Agency for Healthcare Research and Quality, AHRQ Publication No. 04-E024-2, 2004.</td>
<td>Structured literature review (9 articles included in the provider review and 47 in the consumer review); single interventions</td>
<td>US</td>
<td>Financial incentives both for providers and consumers</td>
<td>Providers: There is little evidence available to support the idea that explicit provider financial incentives in the form of bonus payments are effective; Consumers: some evidence that economic incentives are effective in the short run for simple preventive care and well defined, distinct behavioural interventions; Size of incentive: no clarity for providers; relatively modest for consumers.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Beer, M. and M. Cannon, &quot;Promise and Peril in Implementing Pay for Performance,&quot; Human Resource Management, Vol. 43, 2004, pp. 3-48.</td>
<td>Discussion paper involving case studies of incentives in other sectors of the economy; single and complex interventions</td>
<td>US</td>
<td>N.A.</td>
<td>P4P in other sectors of the economy: Despite the popularity of pay for performance programmes, very little research has examined the dynamics and dilemmas associated with implementing these programmes. The study looks at the implementation of thirteen experiments in pay for performance that were initiated by local management in a high-commitment company (Hewlett Packard). It examines Hewlett Packard documents and interviewed managers to understand their experience with implementing these programmes. Managers reported a relatively unfavourable cost-benefit assessment of programmes and difficulty in designing and maintaining them, especially in a fast changing business environment. Managers at each site eventually concluded that they could attain greater performance benefits through alternative managerial tools like effective leadership, clear objectives, coaching or training, and therefore discontinued their pay for performance programmes.</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reference</td>
<td>Type</td>
<td>Author(s)</td>
<td>Year</td>
<td>Description</td>
<td>Country</td>
<td>Source</td>
<td>Notes</td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>------</td>
<td>-----------</td>
<td>------</td>
<td>-------------</td>
<td>---------</td>
<td>--------</td>
<td>-------</td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>AHRQ, Strategies to Support Quality-Based Purchasing: A Review of the Evidence, Rockville, MD: Agency for Healthcare Research and Quality, 2004.</td>
<td>2004</td>
<td>Literature review of single interventions (results of 9 RCT designs were considered)</td>
<td>US</td>
<td>Financial (FFS and bonus) and non-financial incentives</td>
<td>Financial: Recipient of incentive: no clarity in the effect of target group on outcome of intervention; Magnitude of the incentive: no consistent relationship between the magnitude of the incentive and response (for immunisation); Fee-for-service vs. bonus: stronger evidence of positive impact of FFS as an incentive scheme compared to bonus payment; Reputational incentives: hospitals with low performance scores are more likely to engage in quality improvement activities which is especially true for hospitals whose performance was released to the public.</td>
<td>Primary and secondary care but primarily preventive care including outcomes such as screening for smoking and smoking cessation</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>Grimshaw, J. M., L. Shirran, R. Thomas, G. Mowatt, C. Fraser, L. Bero, R. Grilli, E. Harvey, A. Oxman, and M. A. O'Brien, &quot;Changing Provider Behaviour. An Overview of Systematic Reviews of Interventions,&quot; Medical Care, Vol. 39, No. 8, Supplement 2, 2001, pp. II-2-II-45.</td>
<td>2001</td>
<td>Overview of systematic reviews (41 systematic reviews were included)</td>
<td>US</td>
<td>Non-financial incentives</td>
<td>Passive dissemination (e.g. mailing educational materials to targeted clinicians) is generally ineffective and is unlikely to result in behaviour change when used alone; active approaches may be more effective but are associated with higher costs; Interventions of variable effectiveness include audit and feedback, and use of local opinion leader; Generally, effective strategies include educational outreach (for prescribing behaviour) and reminders. Multifaceted interventions based on assessment of potential barriers to change are more likely to be effective than single interventions; audit and feedback is found to have a modest effect, at best; the effect of implementing guidelines is mixed; they are found to spur behaviour in secondary care but not as much in primary care.</td>
<td>Primary and secondary care; medical education, dissemination and implementation of guidelines, preventive care, prescribing, referrals, test ordering, end-of-life care; effectiveness of interventions</td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Zweifel, P. and W. G. Manning, &quot;Moral Hazard and Consumer Incentives in Health Care,&quot; in Handbook of Health Economics, Volume 1, Culyer, A. J. and J. P. Newhouse, eds.: Elsevier Science, 2000, pp. 409-459.</td>
<td>2000</td>
<td>Theory paper supported with empirical evidence from single interventions</td>
<td>Developed countries with most of empirical evidence drawn from the US</td>
<td>Consumer incentives</td>
<td>This paper discusses some likely effects of moral hazard on demand of health care. Depending on the type of financing of the health care system (capitation, FFS, etc), moral hazard may produce higher or lower demand for care and hence affect care with respect to utilisation, outcomes and cost. The effect of changes in the direct cost to consumers on their demand of health care is also discussed.</td>
<td>Primary and secondary care; preventive care</td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>McGuire, T. G., &quot;Physician Agency,&quot; in Handbook of Health Economics, Volume 1, Culyer, A. J. and J. P. Newhouse, eds.; Elsevier Science, 2000, pp. 461-536.</td>
<td>2000</td>
<td>Theory paper supported with empirical evidence from single interventions</td>
<td>Developed countries with most of empirical evidence drawn from the US</td>
<td>Physician related incentives</td>
<td>The review reveals three mechanisms physicians may use to influence quantity of care provided to patients: quantity setting of a non-refundable service, influencing demand by setting the level of a non-contractual input (&quot;quality&quot;), and, in an asymmetric-information context, taking an action to influence patient preferences. The third mechanism is known as &quot;physician-induced demand.&quot; The empirical literature on this topic is reviewed. Theories based on alternatives to profit maximization as objectives of physicians are also reviewed, including ethics and concern for patients, and the &quot;target-income&quot; hypothesis.</td>
<td>Primary and secondary care; preventive care</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>Johnston, G., I. K. Crombie, H. T. O. Davies, E. M. Alder, and A. Millard, &quot;Reviewing Audit: Barriers and Facilitating Factors for Effective Clinical Audit,&quot; Quality in Health Care, Vol. 9, 2000, pp. 23-36.</td>
<td>2000</td>
<td>Comprehensive review of 93 publications on single and complex interventions</td>
<td>primarily UK; other countries (unnamed)</td>
<td>Audit (clinical)</td>
<td>Health professionals' attitudes: Benefits of Audit: Professional benefits: evidence that clinicians feel they have benefited from audit through improvements in communication between professional groups and increased professional satisfaction and knowledge; Patient care and service delivery: Patient centred benefits were found to be improvements in patient care, improved patient satisfaction, and better patient feedback; Doctor centred benefits included increased knowledge, satisfaction, performance, and teamwork. Overall, attitudes to audit are largely positive and were related to audit experience; the more audit experience the more positive the attitude; Health professionals' attitudes: Disadvantages of Audit: Increased workload: evidence consistently suggests that clinicians feel that audit detracts from clinical work at the expense of patient care, and that collaboration on large projects leads to a reduction of practice based activity; Restriction of clinical freedom: evidence suggests that physicians have mixed feelings about the about the effect of audit on their professional freedom; on the one hand, physicians indicate that audit was beneficial and on the other constrained them in providing individualised and the best care possible for patients; Professional threat: Negative attitudes are associated with suspicion about motives, fears of intimidation, and ridicule; beliefs that it caused discord among professionals; and a feeling that it was being used as a government plot to discipline doctors and stifle individuality.</td>
<td>Primary care</td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Gosden, T., F. Forland, I. Kristiansen, M. Sutton, B. Leese, A. Giuffrida, M. Sergison, and L. Pedersen, “Capitation, Salary, Fee-for-Service and Mixed Systems of Payment: Effects on the Behaviour of Primary Care Physicians,” Cochrane Database of Systematic Reviews, No. 3, 2000.</td>
<td>2000</td>
<td>Systematic review of single interventions (8 papers were included, 2 RCT and 2 before and after design)</td>
<td>Developed countries, primarily UK and US</td>
<td>Capitation, Salary, Fee-For-Service and Mixed Systems of Payment</td>
<td>Evidence suggests that the quantity of primary care services provided by primary care physicians (PCPs) under FFS payment was higher than that provided by capitated and salaried PCPs. The evidence concern the impact of capitation payment on the number of hospital and specialist visits compared with FFS is mixed. There is evidence that salaried payment results in a lower number of primary care visits compared with capitation while capitation payment resulted in higher costs compared with FFS payment. Further, theory suggests that to maximise income PCPs under FFS would aim to provide a high volume of low cost activities, whereas capitation PCPs would deliver care that reduced future costs (eg preventive care, health promotion). There is no conclusive empirical evidence of this. Empirical evidence also suggests that the overall satisfaction of patients of salaried PCPs did not differ from those of FFS PCPs; the lack of evidence did not allow for conclusions to be made on whether the three payment systems resulted in changes in clinical outcomes and patients' health status.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Primary care; health professional outcomes; health professional process; health services utilisation; health care costs; and patient outcomes  

| 32 | Chaix-Couturier, C., I. Durand-Zaleski, D. Jolly, and P. Durieux, “Effects of Financial Incentives on Medical Practice: Results from a Systematic Review of the Literature and Methodological Issues,” International Journal for Quality of Health Care, Vol. 12, No. 2, 2000, pp. 133-142. | 2000 | Systematic review of the literature (89 studies of which 8 with RCT design); single and complex interventions | Developed countries, primarily UK and US | Financial incentives | Fund holding: no effect on the workload of GPs while a moderate reduction and prescription cost and total number of drugs per prescription; shifting from FFS to fund holding reduced the number of referral for elective surgery and private clinics; when an incentive of size close to fund holders is provided to non-fund holders, they slightly reduced prescriptions and shifted to generic drugs; FFS with tariff freeze: evidence suggests an increase in the quantity of services produced; Ceiling on annual revenue: modest redistribution of patients from more to less active physicians and reduction of total medical revenues; Shifting from FFS to capitation affected only most elective of procedures; Salary: salaried physician referred their patients less frequently and had lower level of activities compared to FFS; FFS doctors chose home visits more often than salaried (yet effect is not statistically significant) and when forced to go on salary traded revenue for leisure; Managed care (US): Effect on health resources: managed care resulted in lower costs mainly through reduction of length of stay; managed care was linked to lower hospital admissions; capitation significantly decreased the number of physician visits and hospitalisations while FFS was found to be associated with more visits per patient, more frequent visits and better continuity of care; Prospective payments: the effect of prospective payment on hospital care is a reduction in utilisation of services (both length of stay and admissions) which the magnitude of effect varying by disease category (greatest effect is found for psychiatric disorders); managed care was found to reduce the cost per admission with greater reductions found in those hospitals that promote sharing of resource use with physicians and dissemination of guidelines on the process of care; Outcomes of care: poor and elderly patients treated in FFS practice had better outcomes than those treated in managed care. |  

Primary care; health professional outcomes; health professional process; health services utilisation; health care costs; and patient outcomes
<table>
<thead>
<tr>
<th>Reference</th>
<th>Year</th>
<th>Study Type</th>
<th>Countries</th>
<th>Payment Type</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gosden, T., L. Pedersen, and D. Torgerson, <em>How Should We Pay Doctors? A Systematic Review of Salary Payments and Their Effect on Doctor Behaviour,</em> Q J Med Vol. 92, 1999, pp. 47-55.</td>
<td>1999</td>
<td>Systematic review of empirical studies (23 papers were included; 1 RCT, 1 before and after, 2 interrupted time series and 15 case control)</td>
<td>US, Norway, Germany, Canada and South Africa</td>
<td>Salary Payments</td>
<td>Salary payment reduces activity compared to FFS; capitation appears to have a similar but smaller effect. If cost containment is a key policy, then salaried payments are more likely to achieve that compared with FFS and possibly more effectively than capitation. However, cost containment by itself may be inefficient if it results in the provision of sub-optimal care; the use of salaried systems may not be ideal for pursuing public health policy such as population immunisation.</td>
</tr>
<tr>
<td>Giuffrida, A., T. Gosden, F. Forland, I. Kristiansen, M. Sergison, B. Leese, L. Pedersen, and M. Sutton, <em>Target Payments in Primary Care: Effects on Professional Practice and Health Care Outcomes,</em> Cochrane Database of Systematic Reviews, No. 4, 1999.</td>
<td>1999</td>
<td>Systematic review of single interventions (2 studies met inclusion criteria, 1 RCT and 1 interrupted time series)</td>
<td>US and Scotland</td>
<td>Target Payments</td>
<td>Weak to no evidence that the introduction of target payments increases immunisation and vaccination rates.</td>
</tr>
<tr>
<td>Ellis, R.P., <em>Creaming, skimping, and dumping: provider competition on the intensive and extensive margins,</em> Journal of Health Economics Vol. J17, No. 5, 1998, pp. 537-555.</td>
<td>1998</td>
<td>Theoretical study; looks at reimbursement mechanisms and the incentives they create for providers to cream patients; no empirical support</td>
<td>N.A.</td>
<td>Reimbursement incentives</td>
<td>The study concludes that cost-based reimbursement results in creaming of all types of patients. prospectively paid providers cream low severity patients and skimp high severity ones. If there is dumping of high severity patients, then there will also be skimping.</td>
</tr>
<tr>
<td>Reference</td>
<td>Year</td>
<td>Study Type</td>
<td>Setting</td>
<td>Incentive Type</td>
<td>Study Description</td>
</tr>
<tr>
<td>-----------</td>
<td>------</td>
<td>------------</td>
<td>---------</td>
<td>----------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Frey, B., Not Just for the Money, Cheltenham, UK and Northampton, MA, USA: Edward Elgar, 1997.</td>
<td>1997</td>
<td>Theoretical study, no empirical support</td>
<td>N.A.</td>
<td>Extrinsic versus intrinsic</td>
<td>Theoretical discussion on how financial incentives may affect the intrinsic and extrinsic motivation. The author argues that external financial incentives may crowd out intrinsic motivation and therefore create unintended consequences; therefore, the author claims, to design incentives appropriately attention should be paid to the nature of motivation and behaviour to be changed, so that the incentives scheme designed and implemented crowds in rather than crowds out intrinsic motivation.</td>
</tr>
<tr>
<td>Oxman, A. D., M. A. Thomson, D. A. Davis, and R. B. Haynes, &quot;No Magic Bullets: A Systematic Review of 102 Trials of Interventions to Improve Professional Practice,&quot; Can Med Assoc J, Vol. 153, No. 10, 1995, pp. 1423-1431.</td>
<td>1995</td>
<td>Systematic review of single interventions (102 trials were included)</td>
<td>Developed Countries, with an emphasis on the US</td>
<td>Non-financial incentives</td>
<td>Educational materials: printed materials only failed to demonstrate changes in performance or health outcome, a finding that has been associated also with the distribution of guidelines; Conferences: passive dissemination has no effect; active dissemination including outreach and local opinion leader show positive effect on physician behaviour; Patient-mediated interventions: Patient education effected statistically significant improvements in the management of diabetes mellitus that were probably clinically important. These effects were enhanced when patient education was combined with physician education; Audit, feedback and reminders: evidence suggests nil to moderate effect; Multifaceted interventions and local consensus processes: The use of a variety of interventions, such as audit and feedback, reminders, outreach visits, patient-mediated interventions or opinion leaders, has demonstrated changes in professional performance and, less consistently, changes in health outcomes.</td>
</tr>
<tr>
<td>Davidson, S. M.et al., &quot;Prepayment with office-based physicians in publicly funder programs: results from the children's Medicaid program,&quot; Pediatrics, Vol. 89, No. 4, 1992, pp. 761-767.</td>
<td>1992</td>
<td>Empirical study; experimental setting; regression analysis; sample size: 4627 children</td>
<td>US</td>
<td>FFS versus capitation</td>
<td>This paper is a report of the results of a study for collecting empirical evidence regarding the effects of alternative approaches to paying physicians for serving children in the Medicaid programme. Two approaches were examined: (1) visit fees set at twice regular Medicaid fees in return for physician agreement to manage utilisation and (2) capitation and financial risk-sharing along with the same physician agreement to manage utilisation. Participating physicians were assigned randomly to either of the two payment groups. Comparisons of utilisation and expenditures were made between these two plans and the regular Medicaid programme (fee-for-service, low fees). Results showed no adverse effect of capitation payments on primary care visits to office-based physicians. Capitation physician referrals to specialists decreased relative to all other groups studied, consistent with the theory that the financial incentives in capitation will lead primary care physicians to reduce referrals to specialists.</td>
</tr>
<tr>
<td>----</td>
<td>--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1990</td>
<td>Empirical study; experimental setting, RCT (100 practitioners); outcome measures: number of consultations (face to face and by telephone) and renewals of prescriptions, diagnostic and curative services, and specialist and hospital referrals per 1000 enlisted patients in one week</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Denmark</td>
<td>FFS (with a component of capitation) versus capitation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Introducing a partial fee for service system had the effect of reducing referral rates by general practitioners. &quot;The concept of a &quot;target income&quot; which doctors aim at, rather than maximising their income seemed to play a part in adjustment to changing the system of remuneration.&quot;</td>
<td>General Practices in Copenhagen city and Copenhagen county</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>