

**THE LEVERHULME CENTRE FOR MARKET AND PUBLIC  
ORGANISATION**

The Leverhulme Centre for Market and Public Organisation, a Research Centre based at the University of Bristol, was established in 1998. The principal aim of the CMPO is to develop understanding of the design of activities within the public sector, on the boundary of the state and within recently privatised entities with the objective of developing research in, and assessing and informing policy toward, these activities.



Leverhulme Centre for Market and Public Organisation  
University of Bristol  
Department of Economics  
Mary Paley Building  
12 Priory Road  
Bristol BS8 1TN

*Tel: (0117) 954 6943*

*Fax: (0117) 954 6997*

*E-mail: [cmpo-office@bristol.ac.uk](mailto:cmpo-office@bristol.ac.uk)*

**The Role of Incentives in the Public Sector:  
Issues and Evidence**

Simon Burgess and Marisa Ratto

May 2003

Working Paper No. 03/071

# The Role of Incentives in the Public Sector: Issues and Evidence

Simon Burgess\*  
and  
Marisa Ratto

*\*University of Bristol and CEPR*

May 2003

## **Abstract**

Using incentive pay to improve public sector efficiency is an important component of the UK Government's public service modernisation agenda. In this paper, we review the important issues in performance pay in the public sector, and summarise the evidence on its effects. We consider how optimal incentives for public sector workers may differ from those in the private sector, and if so, what types of incentives are more appropriate for the public sector. We investigate the reasons for the infrequent use of explicit incentives in the public sector. We summarise evidence of particular relevance to the public sector, on issues such as the impact on output of incentive pay schemes, gaming and dysfunctional behaviour, multiple principals, intrinsic motivation and team-work. Finally, we comment on the design of new policies being introduced in the UK public sector in the light of the theoretical arguments and the evidence.

**Keywords:** incentives, public sector

**JEL Classification:** H4, J3

## **Acknowledgements**

We thank the Leverhulme Trust for funding this research. Thanks also to Margaret Stevens and a referee for comments.

## **Address for Correspondence**

Department of Economics  
University of Bristol  
12 Priory Road  
Bristol  
BS8 1TN  
[Cmpo-office@bristol.ac.uk](mailto:Cmpo-office@bristol.ac.uk)  
(0117) 954 6943

# 1. Introduction

The use of explicit incentives to improve the efficiency of the public sector is an important component of the UK Government's public service modernisation agenda. Explicit incentive contracts in the form of performance-related pay have always been more common in the private sector than in the public sector<sup>1</sup>, but the issue of incentivising the public sector is relatively recent. In November 1998, the Public Services Productivity Panel was created to "bring in outside experts ... to advise the Government on ways of improving the productivity and efficiency of government departments"<sup>2</sup>. A long term commitment in this direction was announced in the White Paper "Modernising Government", issued in March 1999, which emphasised the role of financial and other incentives to promote better performance. New performance indicators and systems of measuring and monitoring performance have been adopted. Particular emphasis has been laid on performance-based incentives, as a means to better motivate staff, and hence improve services to the public.

In this paper we review the important issues in performance pay in the public sector, and summarise the evidence on its effects. We investigate the reasons for the infrequent use of explicit incentives in the public sector. On the one hand, recent developments in incentive theory suggest that incentives in the public sector may optimally be absent or very low-powered. According to this view optimal incentives for public sector workers differ from those in the private sector. An alternative view is that the relative weakness of incentive schemes in the public sector is due to strong union opposition and weak government. This seems to be the view point of some of the public and policy debate on incentivising the public sector, which is based on the premise that one can simply read across from the private sector.

In section 2 we briefly review general issues in incentive design<sup>3</sup>. In section 3 we consider how optimal incentives for public sector workers may differ from those in the private sector, and if so, what types of incentives are more appropriate for the

---

<sup>1</sup> See Burgess and Metcalfe (1999a) for some data.

<sup>2</sup> Source: [http://www.hm-treasury.gov.uk/documents/public\\_spending\\_and\\_services/public\\_services\\_productivity\\_panel/pss\\_ps\\_p\\_backgrd.cfm](http://www.hm-treasury.gov.uk/documents/public_spending_and_services/public_services_productivity_panel/pss_ps_p_backgrd.cfm)

<sup>3</sup> Other recent surveys of related material include Prendergast (1999) on incentives in firms, Malcolmson (1999) on incentive contracts, Murphy (1999) on executive pay, Gibbons (1998) on incentives in organisations, Lazear (1999) on personnel economics, and Chiappori and Salanie (2003) on evidence on contracts. The most useful companion to this paper is Dixit's (2002) on public sector issues; he focusses on theoretical aspects (which we downplay here) and we highlight evidence.

public sector. Following Prendergast (1999) , we note the scarcity of evidence relative to theory, but summarise some of particular relevance to the public sector in section 4. In section 5 we briefly describe some of the new policies being introduced in the UK public sector. We comment on the design of the incentive schemes in the light of the theoretical arguments and the evidence provided in the earlier sections. Section 6 concludes.

## 2. General Issues in Designing Optimal Incentives

The core model used to analyse the design of incentives is the principal-agent model, and in particular the moral hazard (hidden action) model. The effort undertaken by the agent is not observable by the principal (the private firm or public organisation). We briefly discuss the standard case and then consider some refinements.

### (a) The standard Principal Agent case

In the standard models of moral hazard, incentives to exert the appropriate level of effort are delivered by linking the agent's compensation to his performance. Typically the reward schemes analysed in the theoretical models are linear functions of performance. When output is perfectly observable, as in the case of a monetary outcome, the value of output provides a perfect indicator of the agent's effort, hence paying an individual the full value of his output will induce the Pareto optimal level of effort. When output also depends on some random component, the principal is not able to infer the worker's effort precisely. If the agent is risk averse, the optimal scheme involves less sharp incentives. The reward system consists of a constant term - a risk premium - to compensate the agent for the risk intrinsic in the production process and a marginal reward proportional to the output produced, which measures the sharpness or intensity of the incentive scheme. The simple case shows that the incentives are less sharp (in this sense) the more noisy is the output measure, the higher is the agent's risk aversion, and the higher is the marginal cost of effort. This result encapsulates the classic trade-off between risk and incentives<sup>4</sup>. It also highlights

---

<sup>4</sup> Though see Gibbons (1998) and Prendergast (2002) who argue that this is now much less central to the study of the characteristics of pay-performance schedules in firms.

the key importance of measurement – poor measures of performance lead optimally to low-powered incentive structures.

Holmström and Milgrom (1987) show that while linear schemes are optimal under some specific circumstances, they are often also used in both the theoretical models and in practice for their simplicity and their applicability. Another type of reward schemes commonly considered in the theoretical models makes the compensation of the agent conditional on achieving a certain threshold of performance (step functions). Holmström (1979) and Mirrlees (1999) analyse the circumstances where these schemes can approximate the first best. Intuitively, these schemes are optimal when the output is very sensitive to the agent's effort in the neighbourhood of the threshold, so that if the agent slackens his effort even marginally, he faces a drastic penalty. As pointed out by Dixit (2002) step functions may be preferred by politicians to linear contracts. Politicians, in fact, are usually very risk averse, as they cannot easily diversify the risk of bad outcomes of public policies and agencies. Schemes which set a minimum standard of performance and a bonus for more are therefore attractive. The main disadvantage of such incentive schemes is that they are vulnerable to manipulation by the agents, as we show below.

### **(b) Multiple tasks, Measurement and Incentives**

The simple structure above is likely to be inappropriate in many, probably the majority of, cases. Agents perform different tasks in their jobs, and Holmström and Milgrom (1990, 1991) provide an analysis of how this multi-tasking affects the optimal design of an incentive scheme. A general insight is that interaction among the different tasks affects the power of the incentive scheme: if actions are substitutes the use of high powered incentive schemes may have undesirable effects on overall performance. This is because exerting more effort on one task increases the marginal cost of the task which is a substitute. Higher marginal incentives in one task will drive the agent's effort away from the tasks which are substitutes.

It is also likely that tasks contributing to multiple outcomes are measured with different errors. If each outcome could be rewarded in isolation, then the optimal incentive scheme would set higher incentives on the more easily measurable outcomes - as they provide a more accurate indicator of the effort exerted by the agent. However, in a context where there are multiple dimensions of output, this

would make the agent concentrate on the tasks which are more accurately measured. To avoid this misallocation of effort by the agent, the principal has to weaken the incentives on the more accurately measured tasks.

An interesting case is analysed by Marx and MacDonald (2001). They consider activities which are substitutes from the perspective of the agents (more time spent on one activity means less time on others), but complements from the perspective of the principal (the principal wants high performance in all of them). Therefore the agent is willing to devote more time to the less difficult activities, whereas the principal prefers him to devote time to all activities. They show that, if the principal is unsure about the agent's preferences over tasks, setting rewards on success on individual tasks may be sub-optimal in that it may induce workers to focus and specialise in the less costly tasks. The system of reward is typically non-monotonic, in that it defines different rewards according to the observed failure, partial success or full success on *all* tasks. Little specific compensation is awarded for each task.

### **(c) Intertemporal Aspects**

Most principal-agent relationships extend over a period of time, during which the agent takes actions several times and the principal observes output several times. This may induce the agent to manipulate the incentive scheme, and can give rise to the so called *ratchet* effect. By observing what outcomes the agent obtains in early periods, the principal can try to infer the difficulty of the task. If the agent does well in early periods, the principal will believe that the task is relatively easy and she will therefore decrease the incentive wage. An agent who is aware of this will respond by under-performing early on in order to make the principal think that the task is relatively difficult. Ratchet effects arise when the principal cannot commit to an incentive system that will not change as results become available. Gibbons (1987) and Laffont and Tirole (1988, 1993, Chapter 9) provide a valuable discussion and analysis of the ratchet effect in multi-period settings.

The other side of the coin is that if the principal is unsure, *a priori*, about the agent's ability. When the measured outcome of an agent depends on the agent's effort, some exogenous noise and also the agent's ability, repeated interactions make it possible for the principal to learn about the ability of the worker. In these cases the principal may not need to use explicit incentives to induce the appropriate level of effort by the

agent, but can exploit some implicit incentives not specified in explicit contracts, but relying more on the intrinsic motivation of the agent.

One example is career concerns. By exerting higher effort in the early stages, the agent can influence the principal's perceptions about his talent, in order to get higher future wages or better jobs opportunities: higher performance in the early stages will enhance the perception of the principal and will translate in better prospects of future rewards or job opportunities. The principal does not need to set explicit incentives to motivate the worker to work hard: the worker will exert high effort even if paid a fixed contract in order to influence the principal's perceptions. An interesting insight, as emphasised by Dixit (2002), is that the issue of career concerns can reverse a presumption of the moral hazard problems that greater uncertainty means weaker incentives. Greater uncertainty about abilities leads to greater uncertainty about the outcome for the principal. However, since effort in a career concern model is driven by talent uncertainty, greater uncertainty will induce the agent to exert greater effort in order to impress the principal about his talent. The principal can offer sharper incentives because this elicits larger effort from the young in the attempt to prove their ability. This argument is developed by Dewatripont, Jewitt and Tirole (1999), who extend Holmström's (1982) seminal model on career concerns to a multi-tasks context.

### 3. Features of Public Sector Agencies: Why is the Public Sector different?

In this section we investigate why optimal incentive structures in the public sector may differ from those in the private sector. Theory suggests some arguments against the use of high-powered incentive schemes, relating to specific aspects of how organisations are structured and on how output is produced and measured. In this section we review a number of such arguments: multiple principals, extreme measurement problems, intrinsic motivation, and the importance of teams in production. Although some of these features can also be found in large firms in the private sector, they are typical of public sector organisations.

### **(a) Multiple principals**

It is argued that one key feature of the public sector is that agents have to serve many masters; technically, that they work for many principals. Delivering incentives in these circumstances is more complex, in that principals are usually interested only in some dimensions of outputs and those interests are not aligned. Berheim and Whinston (1986) provide the seminal model on moral hazard problems in multi-principal settings. Further developments include Holmström and Milgrom (1988), who derive the optimal linear contract in a context with two principals, and Dixit (1996, 1997) who extends their analysis to a multi-principals setting. The main insight is that in these settings each principal will offer a positive coefficient on the element(s) she is interested in and negative coefficients on the other dimensions. This creates a negative externality on the other principals who have to face lower efforts in those dimensions. The aggregate marginal incentive coefficient for each outcome is decreasing in the number of principals. When the agent's efforts for the different principals are substitutes the agent will be more willing to allocate his effort for those principals who pay higher coefficients and this will reinforce the effect of the negative externality and hence will further weaken the aggregate marginal incentives. Dixit (1997) shows that the negative externality created by the interaction among the different principals may be internalised by separating the information regarding each outcome, so that each principal observe and reward only his own outcome, or by grouping together those principals whose interests are aligned. Which solution is more appropriate depends on a specific context. As noted by Dixit (2002), in some political context like the legislature or the executive, it may be impractical to restrain the actions of the principal from outside.

### **(b) Measurement and Monitoring**

Output measurement problems are very common in both the private and public sectors. Of course, there are many jobs that occur in both sectors (for example, secretaries, electricians, designers, purchasing managers etc.), and the measurement issues will also then be more-or-less common. But there are a number of occupations, which do divide strongly between public and private sectors – there are not many sales personnel in the public sector, and not many police officers (for example) in the

private. The key point is that for those uniquely in the public sector, there are two interlocking features. First, they are decision-makers – for example, welfare benefits assessors, police officers, tax inspectors, agency managers. Second, the organisations they work for do not have a single, clear goal. At its most basic, they are typically meant to achieve efficiency and some measure of equity. There are of course workers doing the same types of jobs (in terms of the same daily tasks) in the private sector – insurance claim assessors, security firms, tax advisors, managers – but they have a much clearer decision framework, value maximisation. This combination makes these individuals difficult to incentivise and difficult to monitor.

If measuring outcomes is difficult, monitoring performance becomes important, and it seems likely that there might be different methods of monitoring depending on the type of organisation. Prendergast (2002) examines the problem of inducing efficient performance from bureaucrats in a case where the problem is most acute: types of public organisations defined by Wilson (1989) as “*coping organisations*”. Bureaucrats are characterised as having control over the transfer of an asset to a customer. In particular, their decision depends on factors which are not easily observable by the superiors. Police officers are a good example of such bureaucracies: the decision over arresting a suspect is based on the suspicion of guilt that the police officer has, which is not observable by his superiors. Unlike in non-bureaucratic settings, the decision to transfer the good cannot be based on a price paid by the customer, which is easily observable. Promoting efficiency in bureaucratic settings is therefore more difficult, above all when there are no easily available measures of performance, as is the case in coping organisations. In these organisations neither effort nor output is observable and the primary way of controlling behaviour is by costly audits of the details of cases handled by individuals.

These investigations are targeted to cases where a mistake is likely to have been made and are typically triggered by particular signals of a mistake, as consumer complaints. Prendergast shows that complaint mechanisms work less well in bureaucratic settings than elsewhere. But the optimal response to this produces an outcome, widely observed, where standards of consumer service are low, bureaucracies are largely unresponsive to customer complaints, are predisposed to turning down consumers requests and rarely have their decisions overturned.

Monitoring can be problematic even in situations where the outcomes are potentially observable even at individual level, but measuring performance requires sophisticated management information systems, which are too expensive for a public organisation. One practical result is that performance data is often only available at a more aggregate level than that at which production takes place. For example, outcomes are produced by individuals, but measured only at office level. In these cases only aggregate measure of performance are available, and a solution to the agency problem could be the use of team-rewards rather than individual rewards. This emphasis on the importance of monitoring and aggregated data suggests that a focus on teams may be particularly relevant to the public sector.

### **(c) Teams in production and rewards**

The opposite case to a multi-principal situation is a context where there is one principal dealing with several agents, who contribute to the same final output and work in a team. Holmström (1982) provides the seminal contribution to the theory of incentives in teams and shows that, in a setting where team members depend on each other to produce final output, i.e. where there are complementarities in production, if all the output of the team is shared among team members, team members are induced to free ride. This free-rider problem becomes greater in large organisations. The optimal compensation depends on how easy output is to measure, and on the size of the team. The greater the uncertainty in output measurement and the greater the size of the team the more complex is the design of an optimal incentive scheme and some form of monitoring may become necessary.

The free-rider problem is the standard economics view of the likely (lack of) success of team-based rewards. And yet we see many examples of team incentives chosen by firms and public organisations. There must presumably be benefits to the structure.

One key feature of team rewards is that they may induce peer monitoring. Linking the agents' compensation to group performance may induce team members to monitor each other and hence reduce the extent of free-riding. Kandel and Lazear (1992) analyse the interaction between peer monitoring and effort and show that more effort can be induced by peer monitoring when detection becomes more accurate as monitoring increases and the effect of being caught is substantial. This has an

important implication for the size of the team, in that in large teams the ability of workers to monitor or even inform on each other is reduced.

In Holmström's analysis the team is defined by the production function, in that team members are those agents whose effort contribute to the same outcome. However team-based rewards can also be used in context where there does not exist a "natural" production function to define the team. Team-based rewards may in themselves promote some positive externalities of teamwork. In some cases team-based rewards can be used to foster co-operation among team members. Milgrom and Roberts (1990) suggest that if co-operation within a group of individuals is important for the overall organisational objectives, than rewarding individual performance can detract from team performance by raising the marginal cost of effort in co-operating. The extent to which co-operation can be induced by group incentives depends on the strategic interaction between agents and their attitudes towards performing multiple tasks. Itoh (1991) analyses the relationship between financial incentives and 'helping' effort. He addresses the issue of whether it is always the case that, in moving from an individual based contract (i.e. one where individuals are paid only for their own output) towards one where rewards are based on teamwork, agents are induced to increase the level of helping effort.

In a context where an agent can choose how much effort to devote to his own task and how much to help another agent, Itoh finds that whether co-operation can be induced through financial rewards depends on the strategic interactions among agents and their attitudes towards performing multiple tasks. In particular, agents can be induced to provide help, even for a small change in the wage schedule, if they get positive benefit from both types of effort. If, instead, tasks are similar and agents only care about the total amount of effort, they are reluctant to provide even a small amount of help. In this case a large perturbation of the individual-based contract is required to induce any helping effort from the individual. Co-operation may also be spontaneously induced by long-term repeated interactions among team members. Che and Yoo (2001) explain co-operation as a self-enforcing behaviour, in a context where team members interact repeatedly and show that task assignment and work organisation together with group compensation can create implicit incentives which induce team members to co-operate. In particular, repeated interactions and mutual dependence provide team members with the motivation and means to exert peer sanctions and enforce co-operation.

Teamwork may also facilitate a process of communication and sharing of job experience. There is evidence that this is particularly important for organisations like the UK National Health Service. Wilcok and Headrick (2000), Freeman *et al* (2000) provide evidence on the role of multi-professional teamwork and shared learning in improving health care provision.

#### **(d) Intrinsic Motivation**

Given the nature of the output in the public sector (and parts of the private sector), managers and workers may care about the amount produced, and therefore be less inclined to shirk. For example they may care about the sick, the old, or the unemployed. If the welfare of the clients is the sole goal of the organisation itself, workers will internalise the objectives of the organisation. Setting financial rewards based on performance may actually be counterproductive in that it may send the signal that the relationship between the workers and the organisation is a pure market relationship. As pointed out by Kreps (1997), this can dilute the workers' intrinsic motivation and workers can develop a "distaste for the required effort". On the other hand, it may be that the welfare of the clients is only one goal of the organisation, and in this case the workers' intrinsic motivation will mean a misallocation of effort.

There is an important interaction between intrinsic motivation and another characteristic of public sector agencies, namely the absence of a residual claimant. Francois (2000) develops this argument and shows that if workers care about the level of service provided by their organisation, this can motivate their own effort with no need for high powered incentives, provided there is no residual claimant. The insight is that when a worker cares about the outcome of his organisation, he can be motivated not to shirk only if he believes that the level of service will fall if he does not provide the effort. This can happen only when the employer can credibly commit not to make up any shirking by adjusting other inputs. In private firms this is not possible because the owner of the firm is the residual claimant of net profits and she cannot credibly commit not to make up any shirking, as this will mean lower profits for her.

Only when there is no residual claimant can the employer credibly commit not to adjust other inputs and then the worker will not shirk since he knows that shirking

will affect output. In these situations, the principal can incentivise workers at a lower cost, as she can pay a lower incentive compatible wage. Hence in public organisations, where there is no owner or residual claimant, public service motivation provides a rationale for the use of low powered incentive schemes.

#### 4. Evidence on Incentives<sup>5</sup>

Many commentators have noted the relative premium on evidence compared to theory in this field. Consequently, this survey cannot have the feel of a meta-analysis, weighing the ‘bulk’ of the evidence in favour of this position or that – there is no ‘bulk’. Rather, we will highlight a number of studies that provide some interesting evidence to illustrate a number of the points made above. The evidence is typically focussed on one of two questions: ‘what were the effects of the incentive scheme on outcomes?’, and ‘does the choice of contracts in specific settings fit with the theory?’. One of the key problems is finding settings where we can be confident that the effect of the scheme has been identified.

##### **(a) A robust private sector study**

Lazear (1998) uses a very rich and detailed dataset on a single firm that phases in a performance pay scheme over a period of time, and offers probably the most robust estimate yet of both the motivation and the sorting effects of variable pay. Lazear’s investigation uses data on a large car windscreen fitter (*Safelite*) that changed compensation structure from straight salary to incentive pay (piece rates) over a period of 19 months. The same workers and the same firm are observed under both regimes and output is accurately measured. Because of the phasing-in period, productivity observations under both methods of pay are observed simultaneously. Exogenous productivity effects can thus be controlled for, and the identification strategy seems robust. The switch to piece rates was found to cause average output to rise by 44%. About half of this increase can be attributed to better workers joining the firm and half to motivation effects. Average wages also increased by about 10% following the switch. Thus for a given individual at a particular firm switching from time rates to piece rates caused his productivity to rise by 22%.

---

<sup>5</sup> This section draws on Burgess and Metcalfe (1999b).

### **(b) A successful public sector scheme**

There is much less evidence relating to the public sector than the private sector. One powerful study is that of Kahn, Silva and Ziliak (2001) who examine the impact on the Brazilian tax collection authority of the introduction of performance pay. The 1989 reform involved the payment of financial incentives based on individual and team performance in detecting and fining tax evaders. The amounts involved are very significant – bonuses were often worth more than twice mean annual salary. The authors use a panel of six years on ten tax regions, three years before the reform and three after. The findings show that the scheme had a dramatic effect: fine collections per inspection were 75% higher than in the estimated counter-factual. The authors also noted considerable heterogeneity, with the impact varying from 19% to 145% across regions. This is not a difference-in-difference study because the reform was implemented everywhere, but the authors investigate whether tax compliance was responding to other macroeconomic or policy changes. The pattern of changes in fines suggests that it is indeed attributable to the performance pay reforms.

### **(c) Dysfunctional behaviour, thresholds and timing**

There are now many examples of incentive pay schemes that are deemed to be failures, because the outcome for the organisation is not what it intended. Kerr's (1975) article famously entitled "On the folly of rewarding A while hoping for B" contains many examples, as does Baker, Gibbons and Murphy (1994). Typical cases may include employees aiming to hit quantity targets regardless of quality or failing to co-operate with other employees. Perverse incentive effects generally arise when a job requires workers to perform several tasks but only some are measured and rewarded. Unsurprisingly, in such circumstances, the worker will concentrate his efforts on the rewarded tasks to the detriment of the overall organisational objectives.

Such effects appear to be particularly related to schemes with thresholds and year-end effects. One clear example of this also arises from the JTPA programme. Courty and Marschke (1997, 1999, 2004) examine how the incentive structure of the JTPA programme affects the way in which the programme administrators report outcomes and the effect of their timing strategies on efficiency. The administrator can graduate a participant at any time between the end of training and 90 days thereafter and is awarded a bonus if an annual target is met for employment of programme graduates

by June 30 of each year. A simple incentive model predicts that throughout the year the reporting date relative to the end of training will take place as soon as the participant is employed or at the end of the 90 days. In June however, reporting behaviour is predicted to vary according to how well the agency has performed relative to its target. If the agency cannot possibly reach its target, or if it has already surpassed its target by the end of June, then it will graduate all participants who have completed their training. Otherwise, however, the minimum number will be graduated which enable the agency to achieve its quota. The evidence confirms that this model does indeed reflect how agencies act as the deadline approaches. Further, Courty and Marschke show that this timing strategy adversely affects the efficiency of training. It is found that earnings impacts in the last month of the year are lower than they are at other times of the year. It is hypothesised that this is a result of excess resources being diverted to accounting and away from human capital activities.

Asch (1990) examines a similar situation in the public sector showing how US Navy recruiters reacted to an incentive scheme which rewarded them if they achieved a quota of recruits before the cut-off date. Asch finds that the number of recruits was highest in the period immediately prior to the quota cut-off date and lowest immediately afterwards. Crucially, it was also shown that the average quality of recruits fell as the cut-off approached. The interpretation of this is that as the deadline approaches, the recruiters shade down their selection criteria in order to be able to hit their quantity target.

#### **(d) Teams**

There is a literature examining the impact of profit-sharing or employee share ownership schemes<sup>6</sup>. Empirical work investigating the impact of team-based incentive schemes where the teams are defined within the organisation appears to be very scarce (but see the next section). A recent interesting example of the former is that of Knez and Simester (2001) looking at a firm-level incentive scheme introduced by Continental Airlines in 1995. This promised bonuses to all hourly paid staff (some 35,000). One would expect free-riding to dominate any reaction to the scheme, since any one individual's influence on the outcome is bound to be negligible. Holmström (1982) provides the formalisation for the free-riding problem being greater in large

---

<sup>6</sup> See Blasi, Conti and Kruse (1996) for a survey.

firms. Econometric identification is achieved by comparing outsourced airports with non-outsourced airports, and the results show that performance did in fact increase by more in the latter. The authors argue that the results are robust to alternative hypotheses. This result is interpreted as the result of mutual (peer) monitoring. This is usually thought to be effective in small organisations such as partnerships (see below), not large firms. Knez and Simester argue that peer monitoring worked in Continental because employees worked in relatively small autonomous groups, within which monitoring and enforcement of group norms can be sustainable.

The importance of mutual monitoring and group norms in teams has also been identified by Hamilton, Nickerson and Owan (2003). The authors provide an empirical analysis of the relationship between team incentives, worker participation, worker heterogeneity and productivity. They use data at a garment factory (Koret) in California, which between 1995 and 1997 offered the opportunity to its workers to engage in team production. In particular, prior to 1995, sewing was divided into independent tasks and sewers were paid piece rates. Between 1995 and 1997 workers could choose to participate in autonomous work teams of typically six to seven workers, performing all tasks, and receiving a group piece rate. The analysis, based on weekly productivity data over the years 1995-1997 for 288 employees, shows that the introduction of teams was associated with a 14% increase in productivity, on average. Approximately one fifth of this effect was due to the fact that high ability workers at Koret were more likely to join teams. An interesting result is that teams formed early on attracted relatively high ability workers. These early teams also realised a larger team productivity effect. High ability workers were no more likely to leave the firm than low ability workers after joining a team. Moreover, some workers joined teams despite an absolute decrease in pay suggesting that teams offer non-pecuniary benefits to workers. These findings would suggest the presence of non-pecuniary benefits from teamwork, which might alleviate the free riding problem.

Another result, which confirms the presence of mutual learning and group norms in teams is that teams with greater spread in ability are more productive. This, in fact, would suggest that superior skills get transmitted to the other members of the team and also that high ability workers are able to impose a higher team norm level of effort.

Another part of this literature looks at (US) medical practices. Encinosa, Gaynor and Rebitzer (1997) construct a model of the extent of performance pay in compensation

contracts, allowing for group norms. This model is tested on medical group practices and it is found, as predicted, that the size of the group has a significant influence on the distribution of compensation. More recently, Gaynor, Rebitzer and Taylor (2001) look at incentives for doctors in HMO (health maintenance organisation) practices. They find significant effects of contracts that they characterise as providing \$1 of income to the doctor for every \$10 reduction in medical expenditure. They find an average of 5% reduction in medical expenditures following from the contract. Differences in HMO size are key to identifying the impact of incentives, these being assumed random. So the variation in team size is what is driving differences in the degree of sharpness of incentives.

#### **(e) Intrinsic motivation**

Heckman, Smith and Taber (1996) use the JTPA programme data and investigate the response of training centres to these incentives. Under the JTPA system, local training centres receive monetary rewards based on the employment levels and wage rates attained by graduates of the programme. This creates an incentive for the manager to 'cream-skin' the most employable of the applicants into the programme. Heckman, Smith and Taber estimate the probability of acceptance into the programme using predicted earnings levels and earnings gains, calculated from observed human capital variables, as independent variables. In fact, they find that people with lower expected earnings levels are significantly more likely to be accepted into the programme. Weaker evidence suggests that those with larger expected gains are more likely to be accepted. These results contrast with the cream-skimming prediction, suggesting instead that for these bureaucrats at least, preferences for helping the disadvantaged overrode pecuniary incentives.

#### **(f) Teachers**

Possibly the one area where there is a set of papers on the impact of incentive schemes in the public sector is that of performance pay for teachers. The use of performance pay for teachers has waxed and waned in the US with fears for national productivity – being introduced in great numbers after Sputnik, and later in the 1980s, only to decline in use between these dates. There is a modest literature on this, surveyed in Burgess, Croxson, Gregg and Propper (2001). There is not, however, a

great deal of robust, credible evidence. A few papers are worth considering. First, Ladd (1999) considers an incentive programme for schools in Dallas, and uses panel data on schools to test for effects on test scores and student drop-out rates. The scheme, introduced in 1991/2, is school-based rather than individual teacher-based and provides monetary rewards to all teachers (indeed, all staff, including cleaners and secretaries) in successful schools. The details of the programme reflect quite a sophisticated approach to measuring the outcome, namely test score gains. The test score for each individual student is regressed on that student's personal characteristics (race, gender, and eligibility for free school lunch) and the residuals calculated. These residuals were then compared to the residuals the previous year and estimates of gains produced, so controlling for linear pupil effects. The mean of these across all subjects and all enrolled pupils is the school's score. The scheme used multiple measures of student outcomes from a variety of different tests in an attempt to minimise problems of "teaching to the test"; student attendance and drop-out were also factored in. Given the overall school score, around 20% of schools each year win bonuses, worth around \$1000 to teachers.

Ladd's study uses a panel of school-level student test score gains across six large Texas cities, over the period 1991-1995 (non-availability of comparable data prevents any "before/after" comparison). The output measure used is the pass rate on maths and reading tests, thus emphasising the bottom end of the ability distribution. The panel regressions control for common time effects and for city fixed effects rather than school fixed effects. There are also a number of school characteristics, such as racial mix and percent disadvantaged. The results are generally positive, in that pass rates appeared to increase faster in Dallas than in other cities. However, the results are somewhat complicated by the fact that a positive Dallas effect is also found for the year before the scheme was introduced. Effects differ by sub-groups, being most positive for Hispanics and whites, and insignificant for blacks. The study does not investigate how these improvements came about, but interestingly Ladd notes a substantial increase in turnover of school principals once the scheme was in place.

Lavy (2002a) and (2002b) has produced two recent papers reporting on two different incentive experiments in schools in Israel. The first looks at a school tournament incentive scheme in which schools competed on their average performance and the bonuses were distributed equally to all teachers in the winning schools. This was shown to have a significant impact on pupil performance. The second paper evaluates

another (unrelated) experiment which is a teacher-level incentive scheme. This is again a tournament model, with the standardised teacher scores being compared and the winning teachers awarded a high-stakes money bonus. The criterion for success is based on regular high-stakes exams, and adopts a sophisticated approach to measuring exceptional achievement. A test score residual for each pupil is computed taking account of previous test scores (so focussing on test score gain) and socio-economic characteristics. This is then added up over all the pupils in the teacher's class. Teachers in the top quarter of the distribution of this statistic are awarded the bonus, equal to around 25% of mean teacher salary. The paper deals with the identification issues (namely, which schools were included) in a credible fashion and finds a significant impact on performance for teachers in the scheme. Lavy also finds two other interesting facts. First, there appears to be no obvious negative spill-overs on to teaching other subjects – these exam performances do not suffer as they might if teacher energy were simply diverted from other subjects. Second, observable personal characteristics of the teachers do not predict the probability of being a 'winning' teacher.

#### **(g) Choice of incentive schemes**

Theory makes several predictions concerning the choice of method of pay for an employee. Performance-related pay (PRP) is less likely when output is a noisier signal of effort, when the employee is more risk averse, when output is more costly to measure because of firm characteristics and when the output of a job is multi-dimensional. Subjective evaluation of performance is more likely in cases of multi-tasking. A number of authors have tested these predictions<sup>7</sup>.

Brown (1990) uses data on a large cross-section of US, predominantly manufacturing, establishments to empirically examine the determinants of firms' choice of method of pay. Size is found to have a positive influence and occupational dispersion is found to be negatively correlated with the likelihood that an establishment operates a PRP scheme for any employees. Drago and Heywood (1995) perform a similar investigation using Australian establishment data and find again that size has a significant positive influence on PRP likelihood. They also verify the prediction that

---

<sup>7</sup> See also Chiappori and Salanie (2003) for a review of empirical testing of contract theory in general.

the number of supervisors as a proportion of the workforce is negatively correlated with PRP likelihood.

In addition to these broad cross-sections, two case studies are of interest. Fernie and Metcalf (1998) investigate the pay systems and characteristics of four individual workplaces. Those investigated are a bookmaker, an (ex-civil service) executive agency, a contracted-out (from local authority) unit collecting parking fines and the advertising section of a daily newspaper. These four are chosen because they are all call centres, with the exception of the bookmaker which is used as a comparator. They find that the pattern of the choice of method of pay fits with some of the theoretical imperatives above. In particular, the balance of incentives and monitoring is interesting. Paarsch and Shearer (1997) investigate the determinants and productivity effects of different methods of pay in the tree-planting industry of British Columbia over a 6 month planting season. A model of contractual choice is constructed specific to this context, and a structural model estimated. This shows that the firm appears to select contract form optimally – that is, to make both the choice of piece or time rate, and the value of the piece rate dependent on planting conditions.

The evidence from these case studies and from the cross-section studies lends support to the idea that measurement costs significantly influence the likelihood of performance pay in an establishment.

Turning more specifically to the public sector, Burgess and Metcalfe (1999a) use cross-section establishment data from the 1990 Workplace Industrial Relations Survey to investigate the pattern of existence of incentive schemes across private manufacturing, private services and public services. They confirm that public service establishments are much less likely to operate any incentive schemes than private firms, and then investigate whether this is optimally the case or whether the public sector is simply inefficiently incentivised. Controlling for variation in size, union density, and workforce composition variables, which identify general measurement issues, the study compares the presence of performance pay across both sectors by 8 occupation groups. Disaggregating the presence of schemes by occupation identifies the importance of multi-tasking and allows an insight into core public / private differences. The hypothesis is that within each occupation, employees in public and private service jobs perform broadly the same tasks in terms of complexity and

measurability. Thus if multi-tasking issues are the core determinants of the optimal type of incentive structure and the public sector is efficiently run, then we should expect to see no difference in the pattern of existence of incentive schemes between public and private services. If the issue of multiple principals is important, then we should expect these problems to only arise for decision makers. An unskilled manual worker such as a cleaner faces no accountability difference if she works in a university or in a private software company; the issue of working for multiple principals will, if at all, affect managers and professionals only. If this is the case, then we should expect to see no difference in the existence of performance pay for manual workers but fewer establishments operating such a scheme for managers and professionals. If the public sector is simply inefficiently incentivised, then we should expect to see fewer incentive schemes for all occupations. The key finding is that there is no difference in the likelihood of an establishment operating a performance pay scheme for manuals but public service establishments are much less likely to operate a PRP scheme for non-manual workers.

The choice of scheme is also considered. It is found that merit pay (based on subjective assessments), as predicted, is more likely than objective performance pay for non-manuals (whose output is harder to measure) and the reverse is true for manual workers. Looking at differences across the public/private services divide, we see a big difference between sectors in the likelihood of performance pay for non-manual workers. This may be because of a different occupational composition within the professional / technical group. Salespeople, for example, are a classic case where performance pay works well and these come into the same occupational category as nurses and teachers. If this is the case, then merit pay ought to optimally be used more prevalently for non-manuals in the public sector. In fact, however, it is found that merit pay is also significantly less likely in the public sector for all occupations. The study concludes that the lack of incentive schemes, particularly subjective performance evaluation may be evidence in favour of inefficiency in the public sector.

## 5. New Policies, New Research

Despite this lack of a firm evidence base, there has been much activity in recent years introducing more serious performance pay schemes into the public sector in Britain. In this section, we briefly review these schemes. We also note a number of in-progress studies evaluating these reforms and attempting to fill in some of the evidence gaps. This specificity of approach meshes well with Dixit's (2002) views: "Empirical research on past reforms should not look for blanket verdicts on their success or failure, but should relate the probability of success to specific characteristics of the agencies or organisations, and theoretical research on the design of new reforms should likewise pay attention to these specifics."

### **(a) Performance Pay for Teachers**

For the first time in recent history, the current UK government has introduced a national scheme for individual based performance related pay for teachers. Teachers' pay has been on a single scale, with nine basic increments. Under the new scheme, teachers who have reached the ninth increment are eligible to apply to pass a Performance Threshold. Their personal performance is assessed against various specific categories, including professional development and pupil attainment. If successful, they receive an annual bonus of £2000, which they will continue to receive until the end of their career, without needing to reapply. They also move on to a new, upper pay scale where they will be eligible for further performance-related increments. The first round of this scheme began in 2000. This constitutes an individual performance related pay scheme for teachers.

A number of UK commentators are highly critical of the notion of performance pay in education in general, and the current performance threshold scheme in particular. Teaching unions have also shown outright hostility to these proposals. Objections to the individual performance pay revolve round three overlapping concerns, relating to some of the public sector characteristics noted above. First, teaching is multidimensional and aimed at much wider outcomes than exam results or test scores. Second, teaching involves team-based co-operation that is inconsistent with an individual PRP scheme. Third, teachers are professionals and do not require financial incentives to induce effort.

Research is underway to evaluate the impact of the Performance Threshold on pupil attainment (see Burgess and Croxson, 2001).

### **(b) Performance Pay in the Major Government Agencies – the Makinson Report**

John Makinson, then Group Finance Director of the media company Pearson PLC, was recruited by the Public Sector Productivity Panel to investigate the idea of incentivising civil servants. Specifically, he looked at 150,000 front-line staff in the major government agencies working with the public: the Benefits Agency, HM Customs & Excise, the Employment Service (now Jobcentre Plus) and the Inland Revenue. His report in 2000 contained specific proposals tailored to those four agencies. These have been used to design some pilot schemes for Jobcentre Plus, the Child Support Agency and HM Customs & Excise.

The key features of these schemes are as follows. First, they are neither individual nor organisation-wide schemes, but team-based. Some of these teams are at office level and may comprise up to 100 people or so; some of the teams consist of entire divisions or regions and include many thousands of people. Second, the schemes have threshold targets for bonuses – they are not linear incentive schemes. The ‘stretch’ factor to win the bonus is typically a given percentage increase on the minimum standard target, and related to the Public Service Agreement targets. The bonuses are a percentage of salary, and give an opportunity to earn up to 5% more. Third, they have a range of targets, usually 5, which mix quantity and quality targets.

We<sup>8</sup> are engaged in an evaluation of these schemes.

### **(c) Performance Pay in the NHS**

The NHS Plan (2000) devotes a chapter to the design of incentives for better performance. The incentive scheme is designed as a system of rewards based on performance and consisting of (a) greater autonomy, in the form of lighter monitoring by Regional Offices, less frequent monitoring by the Commission for Health Improvement and greater freedom to decide the local organisation of services, (b) national recognition and (c) financial incentives. Considerable emphasis is given to financial incentives, in particular to extra rewards for team-production. As the document states “...*Incentives will also be developed for joint working between*

---

<sup>8</sup> Burgess, Propper, Ratto, Tominey.

*primary care groups, NHS Trusts and Social Services to achieve improvements in rehabilitation facilities for older people.... The Performance Fund will enable NHS trusts and primary care trusts to offer greater incentives to staff in clinical teams and primary health care teams linked to their contribution to service objectives.”* Therefore a distinction is made between joint production across NHS organisations and joint production within the same NHS organisation, although a clear definition of a team is missing. The design of team rewards is postponed pending the results of some pilot schemes on team bonuses in a number of NHS trusts that are currently being evaluated.

#### **(d) Issues of Scheme Design**

The incentive schemes designed for JobcentrePlus, Child Support Agency, HM Customs and Excise and the NHS all highlight the importance of teamwork and include team-based compensations. In the public agencies where the new policies are being introduced, in fact, output is hardly observable at an individual level, so that the use of individual-based incentives would not be feasible. Moreover in these organisations there are complementarities in production and workers do contribute to the same outcome. However, the very broad definition of a team is a potential weakness in the design of these incentive schemes.

A clear definition of a team is very important for setting an optimal incentive scheme. As we have considered above, teams can be identified by the production process or can be created by the compensation system, in order to promote some positive externalities of teamwork. When there is a natural production function defining a team, the team members and their output are automatically identified, and the choice of the compensation scheme depends on the degree of measurability of output. When, instead, the team is not defined by the production function, the selection of team members, and hence the identification of joint output, depends on the reason for promoting teamwork. The incentive mechanism and hence the reward system, has to be designed differently, according to the motivation it acts on. Therefore, identifying the rationale behind the existence of a team is crucial for designing an effective incentive scheme. If for example the aim is to promote peer monitoring, the team members will be chosen so that they are able to observe each other and the reward system will be contingent on their joint performance.

If teams are defined very broadly it is difficult to identify the reason for their existence and hence difficult to design an effective incentive scheme. Moreover in large teams the contribution of each team member to the joint output may be very little and the team members may find it very hard to identify their team. This can encourage free-riding.

The similarities in these scheme designs presumably arise because they have been designed centrally, by the government, with little knowledge of the actual production process or the organisational structure of each single agency. Yet a clear message we get from the theory is that incentives should be tailored specifically to each agency in the public sector. In particular, as argued by Dixit (2002), performance-based incentives do have important roles in public sector agencies but they "... need selective application to specific agencies or tasks".

## 6. Conclusions

In the light of the increasing attention devoted to performance measures by the current government, we have considered the issues involved in designing performance-related incentives in the public sector, and the evidence on their structure and effects.

In particular we have investigated how optimal incentives for public sector workers may differ from those in the private sector, and if so, what types of incentives are more appropriate for the public sector. A clear message we get from the theory is that the use of performance related incentives, and in particular of performance related pay, is more problematic in the public sector than in the private sector. This is due to aspects like multi-tasking, multiple principals, the difficulty of defining and measuring output, and the issue of the intrinsic motivation of workers. In these circumstances the theory predicts that low powered incentive schemes are optimal and task assignment and work organisation become crucial in promoting better performance and may sometimes be substitutes for high powered financial incentives. But "the" optimal incentive scheme for public sector agencies depends ultimately on the type of the organisation and on which of the above aspects are more important. For example, if multi-tasking is a major issue, task assignment could be an alternative incentive device. By grouping together complementary tasks and assigning them to the same agent, the principal could limit the opportunities of dysfunctional behaviour .

Moreover, if measurability issues are particularly relevant in a multitasking environment, the principal, where possible, should engage in monitoring the agents and use more subjective compensation schemes like merit pay.

Team-based rewards may also be preferred to individual compensation schemes, in contexts where co-operation is important for the outcome of the organisation or where only aggregate measures of performance are available. Investing the bonuses gained from better performance in creating a better working environment (e.g. new equipment and facilities for workers) could also be an optimal alternative to individual bonuses in organisations where workers have a strong intrinsic motivation.

The evidence on the effects of incentive schemes on outcomes and on how the choice of contracts in specific settings fits with the theory suggests that incentives do act on individuals' behaviour. In particular, people do respond to a change in their compensation system, but they may not respond in the way intended by the scheme's designer. There is evidence that workers do concentrate their effort on the rewarded tasks, and engage in dysfunctional behaviour. This is particularly important for schemes with thresholds.

In comparing the choice of contracts between the private and public sector, there are big differences on the way workers are rewarded, above all in the case of non-manual workers: a public organisation is less likely to operate a performance related pay scheme than a private organisation. This is in line with the predictions of the theory: manual workers in the private and public sector perform the same tasks and face the same complexity. The output they produce is subject to the same measurability issues. Hence we should expect the same behavioural response to incentives for these workers, and expect to see the same incentive schemes between public and private services. The theory would suggest a widespread use of merit pay to overcome the problems of objective performance-related pay, but this is not borne out by the evidence. Hence there is some (weak) evidence for thinking that incentive pay is not set optimally in the public sector.

The theoretical arguments make it clear that we cannot simply read across from the private sector, and that optimal incentive schemes in the public sector will be different. Currently, there is not a large body of evidence to assess the theoretical arguments. The evaluation of the current pilot schemes in the UK public sector will start to fill in the gaps between the economic analysis and policy makers.

## 7. References

- Asch, B. "Do incentives matter? The case of Navy recruiters" *Industrial and Labor Relations Review*, 43, 1990, pp.89-107.
- Baker, G. , Gibbons, R. , Murphy, K. J. "Subjective performance measures in optimal incentive contracts" *Quarterly Journal of Economics*, 109, 1994, pp.1125-1156.
- Bernheim, B. D. , Whinston, M. "Common Agency", *Econometrica*, 54(4), 1986, pp.923-942.
- Blasi, J., Conte, M., Kruse, D. "Employee stock ownership and corporate performance among public companies", *Industrial and Labor Relations Review*, 50 (1), 1996, pp.60-78.
- Brown, C. "Firms' Choice of Method of Pay", *Industrial and Labor Relations Review*, 43(3), 1990, Spp.165-182.
- Burgess, S. , Croxson, B. "the impact of Teacher Pay Reform on pupil attainment: An outline of the CMPO research project on the performance threshold", CMPO Working Paper 01/036, 2001.
- Burgess, S. , Croxson, B. , Gregg, P. Propper, C. "The intricacies of the relationship between pay and performance for teachers: Do teachers respond to performance related pay schemes?", CMPO Working Paper 01/035, 2001.
- Burgess, S., Metcalfe, P. "The use of incentive schemes in the Public and Private sectors: Evidence from British Establishments", CMPO Working Paper 99/015, 1999a.
- Burgess, S. , Metcalfe, P. , "Incentives in organisations: A selective overview of the literature with application to the Public Sector", CMPO Working Paper 99/016, 1999b.
- Cabinet Office "Modernising Government" Command Paper Cm 4310, March 1999
- Che, Y-K. , Yoo, S-W "Optimal incentives for teams", *American Economic Review*, 91(3), 2001, pp.525-541.
- Chiappori, P.A. and Salanié, B. "Testing Contract Theory: a Survey of Some Recent Work", in *Advances in Economics and Econometrics – Theory and Applications*, Eight World Congress, M. Dewatripont, L. Hansen and P.

- Turnovsky, ed., *Econometric Society Monographs*, Cambridge University Press, Cambridge, 2003, 115-149.
- Courty, P. , Marschke, G. “Measuring Government Performance: Lessons from a Federal Job-Training Program”, *American Economic Review*, 87, 1997, pp.383-388.
- Courty, P., Marschke, G. “An empirical investigation of gaming responses to performance incentives”, London Business School, 1999.
- Courty, P. , Marschke, J. “An empirical investigation of gaming responses to explicit performance incentives”, *Journal of Labor Economics*, 2004 (forthcoming).
- Department of Health “The NHS Plan”, Command Paper Cm 4818-I, July 2000. London:HMSO.
- Dewatripont, M., Jewitt, I., Tirole, J. “The economics of career concerns, Part II: Application to missions of accountability of Government Agencies”, *Review of Economic Studies*, 66, 1999, pp.199-217.
- Dixit, A. “ The Making of Economic Policy: A Transaction-Cost Politics Perspective.” Cambridge, MA:MIT Press, 1996.
- Dixit, A. “Power of incentives in private vs public organisations”, *American Economic Review Papers and Proceedings*, 87, 1997, pp.378-382.
- Dixit, A. “Incentives and organisations in the public sector: an interpretative review” *Journal of Human Resources*, 37(4), 2002, pp.696-727.
- Drago, R. , Heywood, J. “The choice of payment schemes: Australian establishment data”, *Industrial Relations*, 34, 1995, pp.507-531.
- Encinosa III, W. , Gaynor, M. , Rebitzer, J. “The sociology of groups and the economics of incentives”, NBER Working Paper 5953, 1997.
- Fernie, S. , Metcalf, D. “(Not) Hanging on the telephone: Payment systems in the new sweatshops”, LSE Working Paper 891, 1998.
- Francois, P. “‘Public service motivation’ as an argument for government provision”, *Journal of Public Economics* 78, 2000, pp.275-299.
- Freeman, M. et al. “The impact of individual philosophies of teamwork on multi-professional practice and the implications for education”, *Journal of Interprofessional Care*, 14, (3), 2000.
- Gaynor, M., Rebitzer, J. B. , Taylor, L. J. “Incentives in HMOs”, *Economics Working Paper Archive* 340, Levy Economics Institute, 2001.

- Gibbons, R. "Piece-Rate Incentive Schemes", *Journal of Labor Economics*, 5(4), 1987, 413-429.
- Gibbons, R. "Incentives in Organisations", NBER Working Paper 6695, 1998.
- Hamilton, B. H., Nickerson, J.A., Owan, H. "Team incentives and worker heterogeneity: an empirical analysis of the impact of teams on productivity and participation", *Journal of Political Economy*, 111(3), 2003, pp. 465-497.
- Heckman, J. Smith, J. Taber, C. "What do bureaucrats do? The effects of performance standards and bureaucratic preferences on acceptance into the JTPA program", NBER Working Paper 5535, 1996.
- Holmström, B. "Moral hazard and observability", *Bell Journal of Economics*, vol.10, no.1, 1979, pp.74-91.
- Holmström, B. "Moral hazard in teams", *Bell Journal of Economics*, 13, 1982, pp. 324-340.
- Holmström, B., Milgrom, P. "Aggregation and linearity in the provision of intertemporal incentives", *Econometrica*, 55, 1987, pp.303-328.
- Holmström, B., Milgrom, P. "Common Agency and Exclusive Dealing", Yale University, School of Management, working paper, 1988.
- Holmström, B. Milgrom, P. "Regulating trade among agents", *Journal of Institutional and Theoretical Economics*, vol. 146(1), 1990, pp.85-105.
- Holmström, B. Milgrom, P. "Multi-task principal-agent analyses: Linear contracts, asset ownership and job design", *Journal of Law, Economics and Organisation*, 7, 1991, pp.24-52.
- Itoh, H. "Incentives to help in multi-agent situations", *Econometrica*, 59, (3), 1991, pp. 611-636.
- Kahn, C. M., Silva, E. C. D., Ziliak, J. P. "Performance-based wages in tax collection: The Brazilian Tax Collection Reform and its effects", *Economic Journal*, 111(468), 2001, pp.188-205.
- Kandel, E. Lazear, E. "Peer Pressure and Partnerships", *Journal of Political Economy*, 100 (4), 1992, pp. 801-817.
- Kerr, S. "On the folly of rewarding A while hoping for B", *Academy of Management Journal*, 18, 1975, pp.769-783.
- Knez, M., Simester, D. "Firm-wide incentives and mutual monitoring at Continental Airlines", *Journal of Labor Economics*, 19(4), 2001, pp.743-772.

- Kreps, D. M. "Intrinsic motivation and extrinsic incentives", *American Economic Review*, 87(2), 1997, pp.359-364.
- Ladd, H. F. "The Dallas School Accountability and Incentive Program: An evaluation of its impacts on student outcomes", *Economics of Education Review*, 18(1), 1999, pp.1-16.
- Laffont, J.-J. Tirole, J. "The Dynamics of Incentive Contracts", *Econometrica*, 56, 1988, pp. 1153-1176.
- Laffont, J.-J. Tirole, J. "A Theory of Incentives in Procurement and Regulation". Cambridge, MA: MIT Press, 1993.
- Lavy, V "Evaluating the effect of teachers' group performance incentives on pupil achievement", *Journal of Political Economy*, 110(6), 2002a, pp.1286-1317.
- Lavy, V "Pay for performance: the effect of teachers' financial incentives on students' scholastic outcomes", forthcoming, 2002b.
- Lazear, E. "Performance pay and productivity", Stanford University, 1998.
- Lazear, E. "Personnel economics: Past lessons and future directions", *Journal of Labor Economics*, 17, 1999, pp.199-236.
- MacDonald, G. Max, L. M. "Adverse Specialization", *Journal of Political Economy*, 109 (4), 2001, 864-899.
- Makinson, J. "Incentives for change. Rewarding performance in national government networks", Public Service Productivity Panel, 2000.
- Malcomson, J. "Incentive contracts in labor markets" in *Handbook of Labor Economics*, Vol.3, O.Ashenfelter and D.Card (eds), North Holland, Amsterdam, 1999.
- Marx, L.M. MacDonald, L. "Adverse specialization", *Journal of Political Economy*, 109(4), 2001, pp.864-899.
- Milgrom and Roberts "The economics of modern manufacturing: technology, strategy and organization", *The American Economic Review*, 80(3), 1990. pp.511-528.
- Mirrlees, J. A. "The theory of moral hazard and unobservable behaviour: Part 1", *Review of Economic Studies*, 66(1), 1999, pp.3-21.
- Murphy, K, "Executive compensation", in *Handbook of Labor Economics*, Vol. 3, O.Ashenfelter and D. Card, eds North-Holland, Amsterdam, 1999.
- Paarsch, H. Shearer, B. "Fixed wages, piece rates and incentive effects: Statistical evidence from payroll records", Society of Labor Economists Conference Paper, May 2-3 1997.

- Prendergast, C. “The provision of incentives in firms”, *Journal of Economic Literature*, 37, 1999, pp.7-63.
- Prendergast, C. “The tenuous trade-off between risk and incentives’, *Journal of Political Economy*, 110(5), 2002, pp. 1071-1102.
- Wilcock, P. and Headrick, L. “Interprofessional learning for the improvement of health care: why bother?”, *Journal of Interprofessional Care*, 14 (2), 2000, pp.111-117.
- Wilson, J. “Bureacracy”, Basic Books, 1989.