The Economic and Social Review, Vol. 42, No. 2, Summer, 2011, pp. 113–134

What Have We Learnt About Pay For Performance? Geary Lecture Winter 2010

CANICE PRENDERGAST*
University of Chicago, USA

I INTRODUCTION

Governance is a central concern of economics. In much of economic activity, relevant principals (firms, voters, government agencies, and so on) delegate decisions to agents who may not have the interests of those principals at heart. Much of the literature in microeconomics of the last two decades has been concerned with the issue – what mechanisms can firms use to attempt to align interests? That such concerns are important hardly needs to be emphasised in the current Irish economic environment, where concerns of excessive lending by the banking sector have resulted in many claims about compensation and oversight. The purpose of this lecture is to overview the literature on one mechanism that has been proposed for solving agency concerns – pay for performance – but to place it in the context of other forms of governance.

It is useful to distinguish between two forms of governance – those that reward on outputs, and those that reward on inputs. By this I mean that there are many employees whose pay is largely independent of what they do, so long as they pretty much follow orders. The typical worker in the developed world has such a contract – they receive a salary (or an hourly wage), where the requirement to get the salary is to show up on time, and to do what one is told. Such oversight ranges from the obvious worker-boss relationship, to also include cases where boards of directors proscribe executives from certain actions. Here the threat for failing to follow "orders" is usually that the agent

*W. Allen Wallis Professor of Economics, University of Chicago Booth School of Business, USA.

is removed from the position. An alternative (or perhaps additional) form of oversight is where pay depends on some measure of performance – for example sales, stock prices, or loan yields – and is used (obviously) in cases where firms find that simply telling workers what to do will not work.

Another issue is the notion of control rights. Specifically, what are agents allowed to do? Over what types of arenas do they have either formal authority or real authority, to use the language of Jean Tirole and Matias Dewatripont? How these interact with the use of governance by inputs or outputs has become a central concern in the literature. To give an obvious and topical example, should the use of performance related pay for bankers result in allowing them more control over decisions, or should their decisions become more proscribed? This fundamentally depends on how well the performance measures work, and many of the examples of pay for performance plans that go awry often get this interaction wrong.

Before getting into the details of the literature, it is important at this stage to characterise what I mean by "pay for performance". In common parlance, it means where someone's pay is tied, through some formula, to some output measure, such as where a sales worker makes commission, or a banker is offered stocks or options in her company. Pay for performance in the academic literature is a more general notion, and refers to any case where somebody's compensation (or something else that they care about like a job title, vacation, a prize, etc.) depends on some measure of performance. Do not think of this as relating simply to say a sales worker getting commissions, or an executive being paid on the profits of her firm, but also includes for example, the use of merit based promotions, subjective bonuses given by supervisors, merit based raises, or even the possibility that good performance on a job results in a job offer from another firm.

I am interested in this question for two reasons. The first reason is the obvious one; I want to know whether it works and if so, in what kinds of situations. Second, a central theme of microeconomics over about a fifteen year period is essentially trying to work out various ways of making institutions efficient. Much of this literature has moved into addressing forms of governance beyond pay, such as formal oversight, or is about directly controlling the set of actions that agents are allowed to carry out. The only way in which these things become really central is if pay does not provide incentives in a reasonably appropriate way. So to some extent my interest in this is in actually understanding the alternative mechanisms that turn out to be important. Strangely, it turns out that even though there is so much work being done on other mechanisms, we do not have a really good answer to whether pay for performance is effective, and if so, when.

II THE BASIC ECONOMIC FRAMEWORK

Let me begin by offering a very basic framework to address these issues. Some action (which I will call effort) is taken by a worker – e at a cost $C(e)$, where higher e is more costly. Assume that the firm wishes to maximise e minus wage costs. The standard agency problem arises because the firm does not get to see e and rewards on some performance measure $y(e)$, where y is increasing in e. Specifically, pay is then given by $w(y(e))$ where w is (typically) increasing in y. For simplicity, let us say the relationship is linear in y where the agent is paid a fraction β of output (with some fixed payment also) so – ignoring risk aversion for the time being – the agent chooses e to maximise $\beta y(e) - C(e)$.

It turns out that there are two central issues in terms of whether you should or should not use pay for performance. The first of them concerns the relationship between the noisiness of the performance measure. Consider monitoring an executive with stock prices. If I see some increase in the stock price for the company, to what extent does this reflect what she did? It surely reflects not only what she did, but a lot of other things. If the measure is not just a reflection of what she does, we refer to this as noisiness. The early research on pay for performance, following seminal work by Bengt Holmstrom (1979) was almost entirely on the effect of this measurement noise on the effectiveness of pay for performance. So, for example, assume that the measure of output observed is given by $y(e) = e + \varepsilon$, where ε is normally distributed with mean 0 and variance σ^2. Hence the measure obtained is unbiased, but noisy.

Why should such noisiness matter? Noisiness simply means that sometimes agents are rewarded for things they are not responsible for, but equally they are penalised for things that they have no control over. There has to be a cost to such noise for it to matter. The early literature parameterised this through a risk premium on wages – put simply, workers who face randomness in their pay need to be compensated more to induce them to accept such jobs. The idea is simple – suppose that you have market opportunities that offer utility of say €50,000. Then a risk neutral person would be indifferent between that job and one that pays €25,000 in salary and another €25,000 in (say) commission. This logic is, of course, not true when workers do not like variability in their pay, i.e., are risk averse. As the commission component is usually random, they would prefer the €50,000 for sure elsewhere to the more random outcome above. Instead, the firm is required to offer compensation that exceeds €50,000 to make up for the variability in their compensation. This extra is the risk premium. Importantly, in simple commonly used parameterisations of this problem, the risk premium

– let us call it R – depends on both the extent of pay for performance, β, but also the noisiness of measurement. Specifically, the risk premium has the form $R(\beta, \sigma^2)$, where not only are both derivatives positive, but importantly, the cross partial is also positive. This cross partial – in words – says that the cost of using pay for performance increases when there is more randomness in the environment. The reason is pretty intuitive – using pay for performance induces variance in pay, and more uncertainty amplifies that variance in pay, hence increasing the risk premium.

This cross partial is the source of the first major empirical implication of this literature, namely, the trade-off of risk and incentives, and implies that when performance measures are noisy, they are very expensive to use because sometimes you just pay people for reasons that have nothing to do with them. You are just handing money to them for nothing – the company would have been successful or the product would have taken off anyway, for example. So the initial part of the story says that you do not use pay for performance if measures are poor, as this randomness results in high average compensation.

There is one implicit assumption that underlies this entire literature – that when firms offer variable pay of some form, they can reduce salaries. What risk aversion implies is that this reduction is not one for one, because of the risk premium. A striking thing that I often notice, however, with many pay for performance plans is that when they are introduced, firms do not change salaries, but instead simply add all the variable compensation on top. This is in itself not a problem for this approach, as it remains the case that pay rises with monitoring noise, as you are just paying more rents now to workers because they get rewarded for more things beyond their control. But it does cause the other major problem in firms with pay for performance plans, namely that bonuses, stock options, etc. become seen as "freebies" and something that are inherently sought after even if worker behaviour does not change (this is not true in the standard model). Such political problems in firms (such as the resentment for executives being given stock and options) often result from pay for performance being seen as a gift rather than being solely to reorient behaviour.

The second problem with pay for performance is that many jobs involve a number of tasks but firms only get to measure a subset of these tasks. So, to take a topical example, in evaluating teachers, it might be easy to measure the test scores of children but not their social skills. What effect does this have on pay for performance? In our language above, the issue here is not unbiased measurement error but instead that there is a misalignment between $y(e)$ and the firm's objectives. In the language of economics, this is what is referred to as multitasking and relates to the kind of harmful responses that can be

caused by using pay for performance. For example, in the teacher's case, there is evidence of teachers focusing their attention only on those things covered in tests ("teaching for the test") or only focusing on the children who are marginal for their bonuses. See Podursky *et al.* (2007), for details.

III BUT DOES PAY FOR PERFORMANCE CHANGE BEHAVIOUR?

These two issues – (i) how much will pay for performance cost? and (ii) will it cause harmful behavioural responses? are predicated on one assumption – that when you pay people to do something on the margin, they will do more of it. From an economist's perspective, effort always increases when you give more pay for performance. This is largely taken as a given, and from there the implications of the theories develop.

Yet a more basic question may be to simply identify whether paying on the margin does change behaviour, and if so, by how much? The answer to this question is not as obvious as it appears for two reasons. First, most firms have alternative ways of monitoring performance, so the issue is to what extent paying on the margin is of additional value. To put it another way, most firms that do not use pay for performance do not simply post you a cheque every week without checking what you have done for the week. Second, there is a growing literature in psychology that argues that paying people to do things can be demotivating in a number of ways, which I address below.

Given these issues, the most frequent "test" of agency theory has been to identify whether workers do indeed respond to monetary incentives of this type. This, of course, is not a test of the theory but rather a test of an assumption of the theory, but a useful one nonetheless. There is a very noticeable divergence between work in two literatures on this question. Specifically, the last two decades has seen a plethora of work in the economics literature that indeed shows that when paid to do x on the margin, workers indeed will do more x, and I think it is fair to say that the magnitudes of these responses are often larger than many imagined. Much of this is outlined in my earlier survey, Prendergast (1999). Yet psychologists have, at the same time, claimed the opposite where pay for performance reduces output.

The literature has been moving in two polar opposite directions. Over the last ten to fifteen years there has been a rather large body of work carried out by economists and this empirical literature has essentially argued that pay for performance is startlingly effective in at least changing behaviour. There has been a series of studies that look inside organisations. Somebody's compensation is changed and the resulting change in performance measures is observed. These studies look at all types of occupations; these include

teachers, sports players, and fruit pickers. The surprising thing in this literature is not that productivity seems to go up when there is pay for performance, it is that it goes up by so much. It is not unusual to see productivity numbers going up by 25 per cent to 35 per cent. The most cited study that everybody talks about in this context concerns a company called Safelite. This study was conducted by Ed Lazear (Lazear, 2006). One evening Ed Lazear ended up sitting on a plane next its CEO. Safelite does one thing, if you are in a car accident and your windshield gets broken, they replace your wind shield. Over the space of three hours on the plane, the CEO decided to radically change the compensation of its windshield installers from monitoring by supervisors to commissions. Productivity went up by about 35 per cent. For other examples of such work illustrating effects on performance, see my 1999 survey.

On the other hand, there is a psychology literature that is going in a radically opposite direction, where it is argued that if you pay somebody on the margin to do something they will actually do less of it. As this is such a dissonant message to the economic's viewpoint, let me begin by explaining why I think the psychology evidence needs considerably more generality before these claims can truly be at a stage where it overturns the economic work above. The experiments remain very interesting and help us understand psychological influences surrounding pay, but I do not think that they translate to most of the cases that we care about.

There are two central ideas that argue that sometimes pay for performance backfires in its most fundamental sense. So let me start with the example: suppose that you are a golfer and you have to make a putt that is four feet. In the first case, if you make that putt you get €5, and the second case if you make that putt you get €5 million. In which of the two are you more likely to make the putt? So think about the economic example. The economist would say that because it is worth €5 million to me I concentrate more on making the putt and I am more likely to make it. What psychologists argue is the opposite. While they may agree that the golfer wants to do better when €5 million is on the line, they argue that he may not be able to translate that desire into performance due to a physiological response to the stakes. Specifically, there could be a nervousness effect or an arousal effect whereby he becomes so nervous because its worth €5 million that he misses, he is more likely to miss.

To put this in the terms of the earlier analysis, the objective function of the agent is then to maximise

$$\beta y(e, \beta) - C(e)$$

If we interpret e as effort in the sense of expended energy to exert a change, the problem now is that y_e may a decreasing function of β, so that while the agent may "try" harder, y may no longer increase in β if the ability to transform e into y declines.

The second version of what psychologists often talk about in the context of agency theory is what happens in circumstances where you intrinsically enjoy something. The idea here is simply that once you are being paid for doing something you inherently enjoy it less. So without pay for performance I am in 'I think I am doing this as it is fun' mode – when I get paid, now I am doing it for the money. If that attribution is important enough then what happens is people can end up working less hard. So in my language, this is the cost of working hard. Going back to the basic model, the objective function of the agent is then to maximise

$$\beta y(e) - C(e, \beta),$$

Here, the marginal cost of effort is increasing in the contract itself, as the contract affects the attribution of the cost of exerting effort. This theory says that the cost of working hard now depends on the contract itself, if you pay me on the margin for doing something I inherently enjoy it less, it is more costly for me to carry out that activity. This is fundamentally different from the way economists would see it.

Those are the two very direct confrontations of the psychology approach to pay for performance relative to what economists would say. Let me now turn to why these stories, appealing as they might be, may not actually be the central thing that is important in most examples that we likely care about.

Let me start with the example of the golfers. People have been looking for such arousal effects for a very long time. It is only over the last couple of years that these arousal issues have gained a lot of credibility. For many years people have been looking for this nervousness effect, looking at professional basketball players, soccer players, golfers and they rarely find anything harmful about increasing the stakes. However, over the last say five years, there have been a few studies using laboratory settings that seem to find such deleterious effects of pay for performance. The most highly cited of these comes from a study done in India. The reason it was done in India, I think, is because what they can do is have very, very high stakes' experiments without it being prohibitively expensive to the experimenters. Here is the idea: you ask people to do say puzzles, games that we would have done as teenagers. And what you do is you reward people based on the number of puzzles that they solve. There were many experiments but take a simplified version of two treatments: the first experiment is I give you say $1 for every puzzle that you

solve, the second experiment is I give you say $30 for every puzzle that you solve. You give people short periods of time to do these puzzles (say 30 minutes) and see how many they do. These experiments pretty convincingly show that as you increase the prize beyond some point, performance (in the sense of number of puzzles completed) starts to decline.

The inference that has been drawn from this literature is that there is a physical arousal that occurs when stakes get too high, and this arousal can swamp the usual effects that economists like to focus on. Notably, this arousal appears only to arise when cognitive tasks are being carried out – for instance it is unlikely you would see it when asking people to carry heavy weights for money. The conclusion of this literature – that one should not use pay for performance in creative settings – has developed considerable credence over the last few years. My own view on this is that most of the lessons that we learn from this body of work, while useful, are actually very specific to the experimental design that is being discussed here.

Let us do this in a number of steps. First, think of various settings that involve creativity. Think of my occupation, academia. I like to think of research as a reasonably creative activity, so should we not use pay for performance for research? Should we, for example, allocate professorships randomly rather than giving them based on measured productivity? It does not smell right. Take another example, where the stakes are high – brain surgeons, do we think they do badly because the stakes are so high? I doubt it. (Although the stakes here are not monetary, few could claim that the stakes to their performance are not high.) Finally, think about these professional golfers: do they perform worse on match winning putts?

What is different about each of these cases from those in the experiment? The experiment is largely about giving a group of people a task which is far removed from what they usually do, where the pay-off from that unusual event is – as described in the papers – a "chance in a lifetime". It is a very unusual event, it is not the kind of thing that people do day-in, day-out. But most examples of pay for performance involve precisely the kinds of things people do all the time. Yet when you do things all the time, you cannot stay nervous (or aroused) the whole time. You simply habituate to things. So what I learnt about these studies is, I think they are really interesting studies to learn about what happens when there is an unusual event with a very high payoff – a crisis. In really high stakes' environments people do get nervous and if they get nervous sometimes they do badly. This is an entirely intuitive idea – people panic in crises, and I would be willing to bet I would be as capable of this as anyone. For example, if I had to take a penalty to win the World Cup, I am pretty sure I would do worse than even my limited ability would suggest. But if I got to take penalties for those stakes most of the day, I am guessing I would get used to it. As a result, I think these experiments are of rather

limited value for most employment settings, even those that involve creative tasks.

What would be useful extensions of these experiments is to see their more general validity? Most desirably (but sadly least plausibly), one would like to run them over weeks not minutes. For example, if one could rerun the experiments above where you give subjects two weeks to solve puzzles rather than an hour, I suspect the arousal effects would go away after a bit. Nobody can stay nervous for two weeks.

The second thing about these studies that is important is selection. Let us think of another example, I think part of the reason why professional golfers do well is that they have selected into that occupation because they can handle the stress. Another example is traders, traders on the stock floor face these high stakes all the time, they are selected into these occupations because they can handle it. This does not happen in the experimental evidence cited above. So for this reason also, I am not very persuaded of arousal as a first order constraint on the efficiency of pay for performance. A useful experiment to address this would be to ask people to carry out their chosen occupations under relatively high stakes settings. So for instance, one could ask me to teach an MBA class with varying pay-offs to good ratings by my students. Would I do worse as the stakes got higher? I doubt it, but I can see a possibility of becoming nervous and choking. It would be worth finding out.

Now we look at the second type of study that psychologists do, this is the idea that if you pay people to do something, they will enjoy it less. These studies originated from experiments that were carried out on children many years ago. Suppose you have a group of seven year olds who have toys to play with. For half these children, I pay them for playing with the toys (with sweets for example), and for the other half, I pay them nothing. The idea about intrinsic motivation arises from observing the behaviour of the children after the experimental period is over. If you look at who continues to play with the toys afterwards, it is only the group that did not get paid. The group that got paid they just drop the toys and walk away. The conclusion that people typically draw with respect to this psychology literature is: because the group that is not paid continues to do the activity, the expression that psychologists use is that extrinsic motivation crowds out what we call intrinsic motivation.

These studies have an intuitive appeal – when something becomes our job, we think of it as our job, and perhaps that attribution means we may not work as hard. There is, however, in my mind a significant limitation here. Think of the following extreme example – I do an experiment on the people in a classroom and instead of you playing with toys what I am going to do is pay you to eat doughnuts, in particular half the class gets €1,000 for every doughnut they eat, and the other half gets nothing. Say we do the experiment

for an hour. I can guarantee you that at the end of the hour nobody on my right will continue to eat another doughnut. But the reason has nothing to do with intrinsic motivation, it is because you have eaten so many of them during the hour that you are sick of them. In the language of economics that is called diminishing marginal utility and it is a very difficult issue to overcome.

Some attempts have been tried – usually with time delays – to deal with this. An example would be to wait say a week before seeing whether they take up the puzzles again, the idea being that surely diminishing marginal utility effects would disappear by then. Waiting a week seems to show the same effects as looking straight after the experiment. There are two problems with this. First, who knows over what time frame diminishing marginal utility kicks in? Eating 50 doughnuts in an hour could put me off them for life. More importantly, the identifying assumption of these exercises is that these attribution issues should decline at a rate slower than our usual diminishing marginal utility, and again I see no reason for this to be the case. Once again, while I find these studies interesting, I ultimately see less relevance for pay for performance than is sometimes posited.

This is not meant to be interpreted as a claim that little has been learnt from approaching pay for performance through a behavioural lens. I think the introduction of psychology is immensely valuable in economics and I think there are some areas, even in my field where it is truly important. So let me give you the two examples where I think money can really backfire in the sense of making people do less of the activity that you want them to do.

The first is when you are trying to impress an audience about some characteristic of yourself. There is one study that I do find entirely persuasive. This addresses people who collect money for charities. Consider two scenarios – (i) where those collecting are paid nothing for their efforts, and (ii) where they are paid a little. Convincingly, you will find that the people who are not paid at all tend to collect more money. This result seems entirely credible to me. But for me, what it has to do with is the inference that people are drawing about your altruism. We all care about being perceived as altruistic. If somebody has the view that you are only doing it for the money, that is a reason why you may not do it. But the key point about that example is we are trying to impress an audience. Interestingly, the relevant audience we are trying to persuade may be ourselves. Introspection tells me that this may well be true. Interesting, though, these studies also find that for large enough pay for performance, these effects are reversed.

The other example, which I think is not something that has been focused on by psychologists but is potentially important, concerns selection. There is considerable evidence in many "standard" occupations, that pay for performance results in better quality workers applying for jobs. Now consider how this works in settings where people inherently care about what they do – say,

collecting money for charities. One of the things you may worry about when using pay for performance is that you select people into that occupation who do not inherently care about it but actually are only doing it for the money. So, for example, there would be a danger if you start to pay social workers on the margins for things that they do or maybe perhaps teachers because you might end up selecting a different person. The famous example is actually blood donation, which occurred in Britain during the 1950s. They began to pay people to give blood whereas previously they did not. They found that the quality of blood they were receiving was of poor quality largely because they began to select those whose interests were solely in the money – it was largely due to the selection of people that they were getting.

Let me now return to economic research to my comparative advantage. On the basis of this I think that the preponderance of the evidence is: if you pay people to do something on the margin they will do more of it. Given this premise, the literature has focused on two issues:

(a) Performance pay may indeed change behaviour, but not in the way that the employer would like.
(b) Performance may not improve enough to warrant the additional compensation costs.

This first issue has become known as "multitasking" whereas the usual moniker for the latter is the "trade-off of risk and incentives". Let me briefly deal with each in turn.

IV MULTITASKING

Multitasking refers to the phenomena where you only get to observe a subset of what the individual actually does or alternatively you choose to contract on a subset of what they do. The central question then becomes "are you better off contracting on the subset, or doing nothing"? This ultimately depends on how efforts on one activity are complements or substitutes for those on other activities.

The central examples of multitasking concern cases where some subset of activities can be contracted upon – say the test scores of a teacher's students – but not their social skills. Then paying on test scores will clearly result in more emphasis on the test component of what a teacher does (given our discussion above) but whether it affects the other parts of what they do depends on whether effort on tests is complementary with social skills tests or a substitute.

The academic literature is replete with examples of how paying on one measure results in substitution away from other activities. Let me give you two examples. The best example of this recently uses a policy that has been used in the United States called – No Child Left Behind. The dispersion in the quality of public schools in the United States is unbelievably high. The best public schools are spectacularly good and the worst public schools are truly spectacularly bad. George Bush introduced No Child Left Behind, which involves, among other things, giving incentives to teachers. Suppose that you are a teacher of a group of nine year olds. The way the contracts typically work is that you have to get say 40 per cent of students to exceed a reading standard by the end of the year – if you do so then you get a bonus. So that is standard pay for performance. Now if you look at test scores they have improved as a result of this policy. However, there are a number of things that teachers have done that have not been in the interest, I suspect, of the Department of Education. In the No Child Left Behind example, you only get paid on the fraction of children who take the test. So in some instances, teachers actually tried to persuade the academically poor children to drop out of school, not exactly the objective you are looking for. Realising that they only had to get 40 per cent of the children over the hurdle, the more common action that teachers took was that they basically ignored the children who were going to sail through because they did not get paid on the margin for those. They also ignored the children who were not going to get through at all and they spent all their time on the marginal children.

The second example of multitasking actually occurred in a company that fixes cars called Sears. Sears decided at one point that they wanted to introduce pay for performance for its mechanics. It was a disaster. The reason is that one of the first rules of pay for performance is that you never offer pay for performance in circumstances where a person both diagnoses and cures the problem. What was happening was people were bringing in perfectly fine cars and they were telling them that everything was wrong with them. This is true for doctors too, you do not want to use pay for performance for doctors if the person who diagnoses your problem is the one who benefits from curing you. The reason the Safelite example described above worked is there is no diagnoses problem there – you know if your windshield is broken.

V RELYING ON INTERESTED PARTIES FOR INFORMATION

The idea behind multitasking above is that for some exogenous reason, it is hard to measure some part of a worker's performance. But information on performance typically arises from the actions of potentially interested parties.

For most firms, by far the most important source of information that they get on performance comes from consumers. The most obvious example of this is that consumers do not buy your product if the quality of agent behaviour is poor. Signals of performance can easily be interpreted through the lens of revenues or profits. But in many other circumstances the way that firms find out whether people are doing the right or the wrong thing is that people make complaints. So the first question I want to ask is to what extent can you trust the information that you get from consumers?

Consider the following two examples. First, think about a company like FedEx. The way they get information on the performance of their employees is that somebody complains if a package does not arrive. They say "where is the package, I sent it yesterday it was supposed to be arrived by 12 o'clock, the thing is not here?". Do the same exercise in a restaurant, in a restaurant suppose the food is late or the food is not good, you complain to the manager. This is a really good source of information for using performance pay but there is a simple reason for that – the interests of the firm are the same as the interest of the consumers. FedEx wants the package to arrive and the restaurant wants the food to be good. The consumers complain in exactly the right way.

Now compare that to an example that I have studied – police officers in Los Angeles. Over the last two decades, there have been various attempts to transfer the use of these kinds of compensation plans into the public sector. The most extreme example of this I know is when the city of Los Angeles increased oversight of its police officers. Overseeing police officers is a first order problem – much of what they do is unseen by superiors, and there are no obvious measures of "output" on which to reward them. Yet there are many concerns about the motivations of police officers, and fears of abuse of their powers. How should such agents be overseen? A common complaint about police officers is police departments are notoriously poor about responding to complaints about officers. There was a series of scandals in Los Angeles during the 1990s, and as a result of the uproar, the Los Angeles police department agreed that if anybody made a complaint against a police officer, it would go to an external review board and if that external review board found that the police officer had done something wrong, he or she would either be suspended, fired or in extreme cases taken to court. Now that is essentially the same idea as the FedEx one, what you are doing is you are using consumers to oversee the police force. It backfired. The reason relates to the interests of "consumers" in this case. Specifically, the group of people used to oversee the police officers were the people who police officers come in contact with all the time, mainly suspects. Let us take an example: the case of confronting a suspected drug dealer. In these instances, there is always some possibility of a physical

confrontation with the suspect. Here the concern is that the suspect makes the complaint and as a result there would be some disciplinary action taken against the police officer. What is so different about this case to the FedEx one? The central issue here is that suspects often have exactly the opposite interests to the population. Specifically, society wants the guilty to be arrested, but the guilty just want to be set free. In the FedEx example, consumer and firm interests align, as they both want the package to arrive on time. This misalignment of interests with consumers has two problem. First, from a purely monitoring perspective, society finds out little information when these consumers incorrectly are given benefits – what suspect turns himself in when the police has let him go? Second, there is a harmful behavioural response that can arise – if police officers are being given incentives to keep consumers happy, a simple way to do that is to cut down on arrests in any case where problems might arise. This is exactly what happened: arrests went down by 35 per cent in the space of three years and the murder rate tripled.

What is the more general point here? The more general point is that one of the features that limits the ability of firms to solve agency problems is the relationship between firms and their consumers. When are consumer interests typically aligned? Largely when they pay for goods. The central problem in many public sector settings is that consumers get significant rents – applicants always want visas, welfare benefits, suspects do not want to be arrested, and so on. Because these rents are so large, consumers do not get to choose whether to get them, but instead agents choose their allocations. (So to take a trivial example, a social welfare officer decides whether I qualify for unemployment benefits – I do not get to make that decision myself.) It is this central feature of the public sector that limits its ability to solve agency problems; consumers will never point out errors that garner them these rents. Because of this, oversight (of this type) has to be limited, as otherwise agents will simply capitulate to the desires of their clients.

Let me say this another way – using consumers to oversee gives your employees an incentive to make consumers happy. You have to ask the question – do you want your consumers to be happy? If you are a police officer the answer is probably no. Your job is actually to deny benefits to people whereas in a restaurant, it is great. So, can you trust the information that you are getting from your clients? Sometimes yes but there is a lot of information that you are not going to get in public sector settings.

Another obvious source of information on performance is from supervisors. This form of oversight, generally called subjective performance evaluation, would seem to dominate tying pay to numbers, which have the problems of being manipulated in the way described above. Yet there is considerable

evidence that supervisors often hide information from interested parties. This could be due to some inherent bias (Prendergast and Topel, 1996), because supervisors are the recipients of rent seeking behaviour (Milgrom and Roberts, 1988) or outright bribery (Tirole, 1986). Alternatively, they may simply have an incentive to under-report performance to reduce their wage bill. Said another way, there are many, many examples where supervisors know full well who is doing a good job and who is doing a bad job but these are not reflected in ultimate evaluations. As such, it becomes another reason why the interests of the providers of information on performance need to be addressed when designing compensation. Furthermore, it is also a reason why many firms do tie pay to objective measures of performance, despite the problems that this gives rise to.

VI PAY FOR PERFORMANCE AND COMPENSATION COSTS

A second reason why firms may not use pay for performance is that it costs too much money. A useful way to address this is to consider two workers who essentially do the same job, but where one happens to be paid directly through performance. So suppose I am a sales agent, out on the road dealing with clients. I compare them to people who sell but who are in the office most of the time, they pick up the phone and consumers call and they take the answers. The former is paid by commission and the latter much more likely to be paid primarily by salary. The person who is out on the road makes on average about 40 per cent more. CEOs of *Fortune 500* companies in the United States are largely rewarded via stocks and options. Yet if we consider the CEO of similarly sized utilities – those who run electricity companies or gas companies that are the same size – they are not primarily paid in this way. Instead the majority of their compensation is in salary. To see the effect of using stocks, note that the CEO of the *Fortune 500* companies makes about five times more. Finally, when McDonalds opens an outlet, they spend a lot of time deciding whether to own a restaurant or franchise it. About three-quarters of McDonalds outlets are franchised. The managers of company-owned stores are paid primarily by salaries, whereas franchisees are primarily paid on profits generated. After you take account of all the capital costs, the franchisee make about three times as much as the company manager.

There are a couple of points I want to make from this. The first is obvious, but important, namely, that pay for performance is costly in terms of compensation costs. The second relevant point is that there is wild variation

in how much pay for performance costs. I think that one of the major problems that we have when people have discussions about pay for performance, is that few people have good models in their head for calibrating what these numbers should be. Instead, pay for performance in many cases becomes synonymous with overpaying people, turning it from the kind of efficiency calculation that underlies the academic literature, to a political one. For what it is worth, McDonalds does this well. Specifically, the extra 200 per cent that McDonalds compensates franchisees largely reflects three issues that reflect what they are contributing to McDonalds – extra effort (franchisees work many more hours than company-owned managers), selection (better applicants apply to work at franchisees) and risk (as franchisees face more variable pay). Put another way, McDonalds is attempting to make these franchisees as well off as they would have been had they not become a franchisee.

The theory says that pay rises when compensation depends on performance. Two further and related implications arise from the analysis. First, the increase in pay should be more pronounced when performance measures are more variable. This has largely been empirically verified. Second, when environments are such that it is difficult to measure performance, pay for performance should be muted, for the obvious reason that it is more expensive to use it. This has become known as the trade-off of risk and incentives. However, empirically there is little support for this, for a reason I now turn to.

VII EVALUATING PAY IN THE CONTEXT OF ALTERNATIVES

Firms have many possible avenues for inducing their employees to exert effort. Some involve aspects that have little to do with pay – such as making jobs more inherently enjoyable, or attempting to persuade workers of the importance of the outcomes of their actions – yet even within the class of compensation issues, there are many alternatives. One of the failures of the literature on pay for performance has been a tendency to treat these various methods in isolation. This is important both because it leads to academics overstating the importance of formal pay for performance, but in addition, has resulted in some potentially inappropriate implications that have been drawn from the theory. To see this, consider three alternatives:

1. *Monitoring Inputs*

The most common way that worker performance is overseen is, in the language of economics, monitored inputs. What does that mean? For 80 per cent of workers in the United States, they do not have any formal form of pay

for performance, where their pay is tied to measures of productivity. Instead, they are (loosely) told what to do and if they do what they are told to do, they get their pay packet at the end of the week. For many workers, it is the most effective way for employers to pay.

So when should you deviate from this, and instead pay based on observed output? Tautologically, when it does better. There are two issues that become important in the context of monitoring people's actions. The first of them is you get to see what they do, so let me give you two examples on this. Part of the reason why sales agents who are out of the office get paid for performance, and get paid 40 per cent more, is for the simple reason that you do not get to see what they are up to, they are out on the road the whole time. Because I cannot see what they are doing, the only way I can influence their behaviour is by giving them pay for performance. The second example is McDonalds, so as I told you McDonalds franchisee rate is about 75 per cent. Franchisees get large amounts of pay for performance (as they keep most marginal revenues), whereas company-owner managers do not. In cities, McDonalds typically own the restaurants but in rural areas they franchise them. Why? Because it is easy to check what people do in cities, they can monitor you incredibly cheaply whereas if you are off in the sticks it is more expensive to do.

This is rather prosaic. Something a little more subtle arises when considering firms operating in very random, fast moving environments. It is well known that using pay for performance in these settings can be very expensive, for the simple reason that firms often reward their employees for things that they are not responsible for (such as the market rapidly expanding): a firm introduces a new product and the product takes off. It is nothing to do with an individual worker but they end up getting paid for it. The first lesson of economics should be that if something becomes more expensive you do it less often. It turns out that, empirically, it is exactly the opposite; the times when firms actually use pay for performance are those circumstances where it is actually most expensive.

Why is this? It relates to the difficulty of monitoring inputs. Specifically, in very uncertain settings, it is very hard to tell people what to do because you do not know what they should be doing. Things are changing so rapidly that you delegate control rights to them – you let them take the decisions. Said another way, you can only monitor inputs when you know what inputs they should be using. In rapidly changing environments, its hard to tell managers what to do, and so pay for performance, expensive as it is, becomes more common. See Prendergast (2002), for details of these issues, and relevant references.

Such insights are also useful to address executive compensation in the United States. I am far from an apologist for the way executive compensation is designed in the United States, but one point is relevant here – pay for performance (usually tied to stock prices or earnings) is perhaps the only game in town for providing incentives for the simple reason that the most likely alternative (telling the CEO what to do) suffers from the problem that the board likely does not know what the CEO should be doing.

2. *Promotion*

The two other things that most firms have to be aware of is that often some incentives come for free. First, most firms have some hierarchies so most people are working to try to get promoted. To the extent that promotions are based on merit, firms already have these incentives in place (Lazear and Rosen, 1981). There is a lot of work that has been done by economists over the last fifteen years or so, to show that people respond to these kinds of incentives in exactly the same way as they do to the incentives that I have been talking about up to now.

Incentives in tournaments are affected by two issues: the likelihood of winning and the size of the return if promoted. Jennifer Brown (Brown, 2010) has recent work looking at the effect of Tiger Woods on golf tournaments. Consider the effect of Tiger Woods on the performance of other golfers in a tournament. So suppose you look at close competitors' performance when Tiger Woods is playing, compared to when Tiger Woods is not playing. If he is the typical close competitor to Tiger Woods, he performs considerably worse when Tiger Woods is playing than when he is not. So when people see that they are more likely to have a shot, they try harder. This type of analysis has been looked at more generally in many sports setting, specifically by looking at the betting odds for games with two competitors. Pretty consistently it is shown that when the odds of winning are close to 50:50, performance measures are better than when one team is heavily favoured. Said another way, this standard economic model seems to work reasonably well in terms of providing incentives, so pay for performance becomes less necessary in these settings.

3. *Career Concerns*

A second example of how incentives come "for free" can be seen by looking at MBA graduates in top US programmes. A number of our MBA graduates go off to work for consultant companies. Consulting companies do not have much formal pay for performance. Instead the primary incentives to their associates – besides the prospect of making them partner – is that most of them ultimately work for a client. So as far as the big consulting companies

are concerned, they do not have to give incentives it comes for free. They want to impress these clients anyway. So in many circumstances firms are not going to get a great deal from performance pay because they are already working particularly hard. In effect, sometimes markets can solve the incentive problem. This logic was elegantly shown in Holmstrom (1982), and its empirical implications have been largely born out in later empirical work by Gibbons and Murphy (1992).

VIII WHERE IS THE LITERATURE HEADING?

Given the issues described above, what are likely future avenues for academic research on the provision of incentives. I outline a series of areas that I think may ultimately be fruitful.

1. *What Instruments Can Be Used When Pay Fails?*

If we have learnt one thing about the provision of incentives, it is that using money to align interests does not always work, either because it rewards the wrong thing, or it is simply too expensive to be practical. A central problem with the literature until recently is that economists had little to say about what to do in these instances. However, over the last decade or so, a number of interesting avenues have been addressed. First, as mentioned above, agents have preferences sometimes such that they care about what they do – perhaps they do not care in exactly the way that the firm would like, but there remains some intrinsic motivation. One issue that firms can control is authority, or more generally control rights. Recent work by Aghion and Tirole (1997), and others has begun to address how firms can give control to workers in order to align interests. Another potential avenue is through the design of jobs. Beyond simply making jobs more complementary with effort (for example, it has been argued for a long time that seeing a job brought to completion induces more effort due to a sense of fulfilment), recent work by Dewatripont, Jewitt and Tirole (1999) and Dewatripont *et al*. (1999) shows how the development of missions can induce better effort exertion. Finally, a first order instrument that firms can control is who to hire. In Che and Kartik (2009), Van Den Steen (2010), Prendergast (2007), (2008) and (2010) it is shown that hiring workers whose interests do not align with those of their employers can sometimes improve efficiency, despite the conflicts that this gives rise to. In all of the cases described here, there is a central issue that seems to be developing in importance, namely, understanding the role of instruments other than pay for performance for efficiency.

IX THE INTERACTION OF MONETARY INCENTIVES AND OTHER INSTRUMENTS

Another serious concern of the literature has been a tendency to see various instruments to align incentives in isolation from each other. In recent years, there has been some work attempting to understand how various instruments are likely to interact. Much of this literature has been on the interaction between pay for performance and control rights, starting from Holmstrom and Milgrom (1992) and later considered in Aghion and Tirole (1997) and Prendergast (2002).

Yet there may be other aspects that should be addressed. One issue that needs more attention is the interaction between pay for performance and oversight throught the use of boards. It is typically thought of that these are likely to act as substitutes for each other (who needs oversight if workers are working hard on pay for performance?) but the recent financial crisis might argue otherwise – namely, that the pay for performance that bankers operated under may have increased the need for other forms of oversight, whether through boards or regulation.

X BEHAVIOURAL INFLUENCES ON INCENTIVES

As in many other areas of economics, incentive provision is also likely to be affected by behavioural influences. At the theoretical level, there are a series of papers by Jean Tirole, often with Roland Benabou, which deal with aspects as wide ranging as over-confidence, intrinsic motivation, and attribution issues. This line of research appears to be becoming increasingly influential. At the empirical end of things, more research continues in the laboratory setting, dealing with many of the same issues that I have described above. This literature has already become influential, even though I am far from persuaded by it, and it is likely to become more refined over time. I do see, however, a series of problems that are likely to arise in these settings. First, in many settings the mechanism that likely leads to effort slacking off is physical tiredness, where people justify to themselves that they have worked hard enough. Such tiredness effects are hard to replicate when subjects are pressing buttons for example in 45 minute experiments. Second, many of the issues concerning pay for performance address workplaces we inhabit for tens of hours a week, and it is hard to see how these can be replicated in short laboratory settings. My concern about the arousal effects above is but one example of this. Finally, more than perhaps anything else we do, we select into jobs. These may also make the insights of randomly assigned experiments less valuable than they could be.

XI COMPARATIVE STATICS OF AGENCY MODELS

The literature in agency theory has been ongoing now for the best part of thirty years. So what do the facts say? Do they concur with the theory? The honest answer to this is that I think we do not know. The reason for this is twofold. First, most of the tests of agency theory have not been on whether the predictions of our models are correct, but rather most of the empirical evidence is on inputs to the problem rather than the outcomes. Specifically, we assume that when people are paid to do things they do more of them. Most of the empirical evidence we have accumulated has been on this issue, and despite some psychology experiments, my sense is that this question has not yet been answered. But the theories take this as a given, and address whether observed contracts reflect the trade-offs we consider. Here the evidence is weak. It is weak for two reasons. First, empirical estimates seem to be all over the map – for example, there are many empirical studies that find a negative trade-off of risk and incentives, and others that find a positive trade-off. But this leads to the second problem, which is that most comparative statistics of these models are not very robust. My 2002 paper is but one example of this, and I am currently working on other plausible settings where it seems as if predictions of these models are extremely fragile. Given the absence of a robust theoretical insight, it is not surprising that there is little consensus on whether the theories are empirically correct.

So to conclude, despite the psychology literature I started out with, I think the main conclusion from the literature is that if you pay people to do things, with the exception of some examples like charities, people will do more of it. We may sometimes want to think that there are nobler objectives for what people do, but most of the time money still seems to win. Given this, what are the relevant constraints? First, do you want them to do more of what you can contract on? There are two reasons why maybe not: you cannot trust the information that you get (such as the LAPD) example; or you are just rewarding on a subset of what they do (as in the teachers' example). Second, might there be some cheaper or better way to do this? Given that only 20 per cent of people get pay for performance, the answer would seem to be yes, where plausible alternatives are likely opportunities to be promoted, subjective ratings by superiors, and the prospect of a better job. How all these interact with each other and other aspects of the employment relationship will be the challenge of the agency literature over the next decade.

REFERENCES

AGHION, P. and J. TIROLE, 1997. "Formal and Real Authority in Organizations", *Journal of Political Economy*, Vol. 105, pp. 1-29.

ARIELY, D., U. GNEEZY, G. LOEWENSTEIN, and N. MAZAR, 2009. "Large Stakes and Big Mistakes", *Review of Economic Studies*, Vol. 76, pp. 451-469.

BROWN, J., 2010. "*Quitters Never Win: The (Adverse) Incentive Effects of Competing with Superstars*." Mimeo, Northwestern.

CHE, Y.-K. and N. KARTHIK, 2009. "Opinions as Incentives", *Journal of Political Economy*, Vol. 117, Issue 5, pp. 815-860.

DEWATRIPONT, M., I. JEWITT and J. TIROLE, 1997. "Information Structures and Career Concerns", Mimeo, ECARE.

DEWATRIPONT, M., I. JEWITT and J. TIROLE, 1999. "Missions and the Accountability of Government Agencies", *Review of Economic Studies*, Vol. 66, Issue 1, pp. 199-217.

GIBBONS, R. and K. J. MURPHY, 1992. "Optimal Incentive Contracts in the Presence of Career Concerns", *Journal of Political Economy*, Vol. 100, pp. 468-506.

HOLMSTROM, B., 1979. "Moral Hazard and Observability", *Bell Journal of Economics*, Vol. 10, pp. 74-91.

HOLMSTROM, B., 1982. "Managerial Incentive Problems: A Dynamic Perspective", Mimeo, Kellogg School of Management.

HOLMSTROM, B. and P. MILGROM, 1992. "Multi-Task Principal Agent Analyses: Linear Contracts, Asset Ownership and Job Design", *Journal of Law, Economics, and Organization*, Vol. 7, pp. 24-52.

LAZEAR, E. and S. ROSEN, 1981. "Rank Order Tournaments as Optimal Labor Contracts", *Journal of Political Economy*, Vol. 89, pp. 841-864.

MILGROM, P. and J. ROBERTS, 1988. "An Economic Approach to Influence in Organizations", *American Journal of Sociology*, Vol. 94, pp. S154-179.

PODURSKY, M. and M. SPRINGER, 2007. "Teacher Performance Pay: A Review", *Journal of Policy Analysis and Management*, Vol. 26, pp. 909-949.

PRENDERGAST, C., 1999. "The Provision of Incentives in Firms", *Journal of Economic Literature*, Vol. 37, No. 1, pp. 7-63.

PRENDERGAST, C., 2002. "The Tenuous Trade-Off Between Risk and Incentives", *Journal of Political Economy*.

PRENDERGAST, C., 2007. "The Motivation and Bias of Bureaucrats", *American Economic Review*, Vol. 97, No. 1, pp. 180-196.

PRENDERGAST, C., 2008. "Work Incentives, Motivation, and Identity – Intrinsic Motivation and Incentives", *American Economic Review: Papers and Proceedings*, Vol. 98, No. 2, pp. 201-205.

PRENDERGAST, C., 2010. "Contracts and Conflict in Organizations", www.faculty.chicagobooth.edu/appliedtheory/papers/2010-08.pdf

PRENDERGAST, C. and R. TOPEL, 1996. "Favoritism in Organizations", *Journal of Political Economy*, Vol. 104, No. 5, pp. 958-978.

TIROLE, J., 1986. "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations", *Journal of Law, Economics and Organizations*, Vol. 2, No. 2, pp. 181-214.

VAN DEN STEEN, E., 2005. "Organizational Beliefs and Managerial Vision," *Journal of Law, Economics and Organization*, Vol. 21, No. 1, pp. 256-283.