Pay For Performance and Beyond

Prize Lecture, December 8, 2016

by Bengt Holmström*
Massachusetts Institute of Technology, Cambridge, MA, USA.

In this lecture, I will talk about my work on incentive contracts, especially incentives related to moral hazard. I will provide a narrative of my intellectual journey from models narrowly focused on pay for performance to models that see the scope of the incentive problem in much broader terms featuring multitasking employees and firms making extensive use of nonfinancial instruments in designing coherent incentive systems.

I will highlight the key moments of this journey, including misunderstandings as well as new insights. The former are often precursors to the latter. In the process, I hope to convey a sense of how I work with models. There is no one right way about theorizing, but I believe it is important to develop a consistent style with which one is comfortable.

I begin with a brief account of the roundabout way in which I became an economist. It reveals the origins of my interest in incentive problems and accounts for my life long association with business practice, something that has strongly influenced my research and style of work.

---

* I want to thank George Baker, Robert Gibbons, Oliver Hart, Paul Milgrom, Canice Prendergast, John Roberts and Jean Tirole for years of fruitful discussions on the topic of this lecture and Jonathan Day, Dale Delitis, Robert Gibbons, Gary Gorton, Parag Pathak, Alp Simsek, David Warsh and especially Iván Werning for comments on various versions of the paper.

I did not plan to become an academic. After graduating from the University of Helsinki, I got a job with Ahlstrom as a corporate planner. Ahlstrom was one of the ten biggest companies in Finland at the time, a large conglomerate with 20–30 factories around the country. The company took pride in staying abreast of recent advances in management. I was hired to implement a linear programming model that had recently been built to help management with long-term strategic planning. It was a huge model with thousands of variables and hundreds of constraints, describing all the factories within the conglomerate, their production activities, investment plans and financial and technological interdependencies. The company had invested heavily in mainframe computers and had high hopes that the planning model was going to be one of the payoffs from such investments.

The most urgent task for me was to arrange the data collection process. I began visiting factories to discuss the data requirements. It did not take me very long to see that the people providing the data were deeply suspicious of a 23-year-old mathematician sent from headquarters to collect data for a planning model that would advise top management on how to spend scarce resources for investments. They wanted to know what numbers to feed into my "black box" model to ensure that their own plans for their factory would receive the appropriate amount of resources. Data were forthcoming slowly and hesitantly.

After some months, I came to the conclusion that the whole enterprise was misguided. Even assuming the best of intentions at the factory level, reasonable data were going to be very difficult to obtain. The quality of the data varied a lot and disagreements about them often surfaced, especially in cases where factories were interconnected. Also, I grew increasingly concerned about gaming. The integrity of the data therefore seemed questionable for technical as well as strategic reasons.

I suggested that we give up on the grand project and that I instead would focus on two things: (i) smaller models that could help each factory improve its own planning process and (ii) try to deal with incentive problems in investment planning at the corporate level.

My first recommendation met with some success. I worked out small linear programs for factories that seemed amenable to such models. I used these models as an economist would: I tried to replicate what the factories were doing. This involved a lot of back-and-forth. I would get the data, run the linear program and then go tell the factory what the program was proposing as the "optimal solution" given their data and, importantly, why the model was proposing the solution it did. This last step—explaining how the model was thinking—was the key. It made clear that I was not there to recommend a mechanical solution; I was there to understand what may be missing in my model specification. This lesson, on the usefulness of small models and of listening to them, would stick with me throughout my academic career.

My second recommendation to think creatively about the incentive problems surrounding investment planning was a failure. I made all the mistakes one is likely to make when one tries to design incentives for the first time. My thinking was guided by two principles: factories should pay for their borrowing and the price should be obtained at least partly through a market-like process so that funds would be allocated efficiently. In today's language, I was suggesting that we "bring the market inside the firm" in order to allocate funds to the factories.

Today, I know better. As I will try to explain, one of the main lessons from working on incentive problems for 25 years is, that within firms, high-powered financial incentives can be very dysfunctional and attempts to bring the market inside the firm are generally misguided. Typically, it is best to avoid high-powered incentives and sometimes not use pay-for-performance at all. The recent scandal at Wells Fargo explains the reason (Tayan 2016). Monetary incentives were powerful but misaligned and led some managers to sell phony accounts to enhance their bonuses. Kerr's (1975) famous article "The Folly of Hoping for A, While Paying for B" could have served as a warning, though the article is rather thin on suggestions for providing alternative ways to provide incentives. I hope to show that our understanding of incentive problems has advanced quite a bit since the days Kerr published his article. In order to appreciate the progress of thought, I will start with the early literature on principal-agent models.

I. THE PRINCIPAL-AGENT PROBLEM

A. The One-Dimensional "Effort" Model

Early contributions to the principal-agent literature on moral hazard include Wilson (1969), Spence and Zeckhauser (1971), Ross (1973), Stiglitz (1975) and Mirrlees ([1975]1999). Wilson and Ross asked under what conditions the principal's and the agent's preferences over risky lotteries will be perfectly aligned when sharing risk optimally. This is possible if the principal's and the agent's utility functions are such that linear risk sharing is optimal. Spence and Zeckhauser (1971) studied insurance contracts under varying information assumptions including the case of moral hazard as well as adverse selection.

Mirrlees was the first to study in generality the case where the agent provides a service to the principal and the issue is how to motivate the agent to work diligently. The agent's action is often referred to as "effort" though this interpretation should not be taken literally. The model applies in a wide range of situations: an employee working for a manager; a lawyer serving a client; a doctor treating a patient; or a CEO serving the board of a company, to name a few.

There are two challenges when designing an optimal incentive scheme for effort. First, the agent finds it privately costly to provide the service, at least beyond some base level, so a financial inducement based on performance is needed. But performance is imperfectly measured, so variable pay will induce risk on the agent. Since the agent is risk averse, there is a trade-off between risk and incentive. How should it be optimally solved?

One could approach this problem by studying simple incentive schemes such as a linear incentive in addition to a fixed wage, or a bonus for performance beyond some minimum standard.[1] The problem with using a particular functional form is that the analysis will not tell us why different incentives are used in different contexts. Also, fixing the form of the incentive pay may silence trade-offs that are essential for understanding the underlying incentive problem. This makes it valuable to study the problem without functional restrictions.

Let me turn to a simple, generic formulation of the principal-agent relationship. The agent chooses an unobserved level of effort e. The agent's choice of effort leads to a payoff $x = x(e,\varepsilon)$, where ε captures random external factors such as market conditions or measurement errors that the agent does not control directly. I will often work with the additive specification $x = e + \varepsilon$.

As Mirrlees ([1975] 1999) noted, it is technically convenient and more elegant to view the agent as choosing a distribution over x. For a fixed choice e, the distribution over ε induces a distribution over x, denoted $F(x|e)$. This eliminates ε and gives a simpler and, as we will see, much more informative characterization, though the explicit dependence on ε can be helpful for thinking about particular contexts.[2]

Before the agent acts, the principal offers the agent an incentive contract s, which pays the agent $s(x)$ when the realized payoff is x. The principal keeps the residual $x - s(x)$. The utilities of the agent and the principal are, respectively, $U = u(s(x)) - c(e)$ and $V = x - s(x)$, so the principal is risk neutral and the agent (in general) risk-averse. The agent's utility function is additively separable, which is restrictive but commonly used.

The principal and the agent are symmetrically informed at the time they sign the contract (this is what makes the problem one of moral hazard). In particular, they know each other's utility functions and hold the same beliefs about the distributions $F(x|e)$. The principal can therefore forecast the agent's behavior given $s(x)$ even though she cannot observe the agent's choice of effort.

The incentive scheme $s(x)$ must provide the agent an expected utility that is at least as high as the agent can get elsewhere. The agent's participation requires the principal to consider the impact $s(x)$ has on the agent's expected utility. The agent's burden from extra risk and extra effort is ultimately borne by the principal. Finding the best contract is therefore a shared interest in the model (but not necessarily in practice).

To determine the principal's optimal offer, it is useful to think of the principal as proposing an effort level e along with an incentive scheme $s(x)$ such that the agent is happy to choose e, that is, $s(x)$ and e are incentive compatible. This leads to the following program for finding the optimal pair $\{s(x),e\}$:

$$\text{Max } E[x - s(x)|e], \text{ subject to} \tag{1}$$

$$E[u(s(x)) - c(e)|e] \geq E[u(s(x)) - c(e')|e'] \text{ for } e' \neq e, \text{ and} \tag{2}$$

$$E[u(s(x)) - c(e)|e] \geq \underline{U}. \tag{3}$$

The first constraint assures that e is optimal for the agent. The second constraint guarantees that the agent gets at least his reservation utility $\underline{U}$ if he chooses e and therefore will accept the contract. I will assume that there exists an optimal solution to this program.[3]

B. First-best Cases

Before going on to analyze the optimal solution to the second-best program (1)–(3) it is useful to discuss some cases where the optimal solution coincides with the first-best solution that obtains when the incentive constraint (2) can be dropped, because any effort level e can be enforced at no cost. Because the principal is risk neutral, the first-best effort, denoted e^{FB}, maximizes $E(x|e) - c(e)$.

A first-best solution can be achieved in three cases:

I. There is no uncertainty.
II. The agent is risk neutral.
III. The distribution has moving support.

If (I) holds, the agent will choose e^{FB} if he is paid the fixed wage $w = u^{-1}(\underline{U} + c(e^{FB}))$ whenever $x \geq e^{FB}$ and nothing otherwise. The wage w is just sufficient to match the agent's reservation utility $\underline{U}$. If (II) holds, the principal can rent the technology to the agent by setting $s(x) = x - E(x|e^{FB}) + w$. As a risk-neutral residual claimant the agent will choose e^{FB} and will earn his reservation utility $\underline{U}$ as in case (I).

The third case is the most interesting and also hints at the way the model reasons. For concreteness, suppose $x = e + \varepsilon$ with ε uniformly distributed on [0, 1]. This corresponds to the agent choosing any uniform distribution $[e, e + 1]$ at cost $c(e)$. The density of a uniform distribution looks like a box. As e varies the box moves to the right. In first best, the agent should receive a constant payment if he chooses the first-best level of effort e^{FB}. This can be implemented by paying the agent a fixed wage if the observed outcome $x \geq e^{FB}$ and something low enough (a punishment) if $x < e^{FB}$. The scheme works, because two conditions hold: (a) the agent can be certain to avoid punishment by choosing the first-best effort level and (b) the moving support allows the principal to infer with certainty that an agent is slacking if $x < e^{FB}$ and hence punish him severely enough to make him choose first best. In the general model, inferences are always imperfect, but will still play a central role in trading off risk versus incentives.

C. Second-best with Two Actions

I proceed to characterize the optimal incentive scheme in the special case where the agent chooses between just two distributions F_L and F_H. This special case will reveal most of the insights from the basic agency model without having to deal with technical complications. Assume that F_H dominates the distribution F_L in the sense of first-order stochastic dominance: for any z, the probability that $x > z$ is higher under F_H than F_L. This is consistent with assuming that the high distribution is a more costly choice for the agent: $c_L < c_H$. As an example, for $x = e + \varepsilon$ and $e_L < e_H$, F_H first-order stochastically dominates F_L regardless of how ε is distributed.

Assume that the principal wants to implement H, the other case is uninteresting since L is optimally implemented with a fixed payment. Let μ and λ be (non-negative) Lagrangian multipliers associated with the incentive compatibility constraint (2) and the participation constraint (3) in the principal's program (1)–(3). The optimal second-best contract, denoted $s_H(x)$, is characterized by

$$u'(s_H(x))^{-1} = \lambda + \mu[1 - f_L(x)/f_H(x)], \text{ for every } x.^4 \tag{4}$$

Here $f_L(x)$ and $f_H(x)$ are the density functions of $F_L(x)$ and $F_H(x)$. It is easy to see that both constraints (2) and (3) are binding and therefore μ and λ are strictly positive.[5]

The characterization is simple but informative. First, note that the optimal incentive scheme deviates from first-best, which pays the agent a fixed wage, because the right-hand side varies with x. The reason of course is that the principal needs to provide an incentive to get the agent to put out high effort. Second, the shape of the optimal incentive scheme only depends on the ratio $f_H(x)/f_L(x)$. In statistics, this ratio is known as the likelihood ratio; denote it $l(x)$. The likelihood ratio at x tells how likely it is that the observed outcome x originated from the distribution H rather than the distribution L. A value higher than 1 speaks in favor of H and a value less than 1 speaks in favor of L.

Denote by s_λ the constant value of $s(x)$ that satisfies (4) with $\mu = 0$. It is the optimal risk sharing contract corresponding to λ.[6] The second-best contract ($\mu > 0$) deviates from optimal risk sharing (the fixed payment s_λ) in a very intuitive way. The agent is punished when $l(x)$ is less than 1, because x is evidence against high effort. The agent is paid a bonus when $l(x)$ is larger than 1 because the evidence is in favor of high effort. The deviations are bigger the stronger the evidence. So, the second-best scheme is designed *as if* the principal were making inferences about the agent's choice, as in statistics. This is quite surprising because in the model the principal knows that the agent is choosing high effort given the contract she offers to the agent *before the outcome x is observed*. So, there is nothing to infer at the time the outcome is realized.

II. THE INFORMATIVENESS PRINCIPLE

The fact that the basic agency model thinks like a statistician is very helpful for understanding its behavior and predictions. An important case is the answer that the model gives to the question: When will an additional signal y be valuable, because it allows the principal to write a better contract?

A. Additional Signals

One might think that if y is sufficiently noisy this could swamp the value of any additional information embedded in y. That intuition is wrong. This is easily seen from a minor extension of (4). Let the optimal incentive scheme that implements H using both signals be $s_H(x,y)$. The characterization of this scheme follows exactly the same steps as that for $s_H(x)$. The only change we need to make in (4) is to write $s_H(x,y)$ in place of $s_H(x)$ and the joint density $f_i(x,y)$ in place of $f_i(x)$ for $i = L,H$. Of course, the Lagrange multipliers will not have the same values if y is valuable.

Considering this variant of (4) we see that if the likelihood ratio $l(x,y) = f_H(x,y)/f_L(x,y)$ depends on x as well as y, then on the left-hand side the optimal solution $s_H(x,y)$ must depend on both x and y. In this case y is valuable.

Conversely, if $l(x,y)$ does not depend on y the optimal scheme is of the form $s_H(x)$. We can in that case write

$$f(x,y|e) = g\,(x|e)h(y|x),\ \text{for every } x,\ y \text{ and } e = L,\ H. \tag{5}$$

On the left is the density function for the joint distribution of x and y given e. On the right, the first term is the conditional density of x given e and the second term the conditional density of y given x. The key is that the conditional density of y does not depend on e and therefore y does not carry any additional information about e (given x). In words we have the following:

> *Informativeness Principle* (Holmström, 1979, Shavell, 1979). —An additional signal y is valuable if and only if it carries additional information about what the agent did given the signal x.

Stated this way the result sounds rather obvious, but it only underscores the value of using the distribution function formulation.[7] One can derive a characterization similar to (4) using the state-space formulation $x(e,\varepsilon)$, but this characterization is hard to interpret because it does not admit a statistical interpretation. The reason is that there are two ways in which y can be informative about e given x. It could be that y provides another direct signal about e ($y = e + \delta$, where δ is noise) or y provides information about ε ($y = \delta$, where δ and ε are correlated), which is indirect information about e. Both channels are captured by the single informativeness criterion.

As an illustration of the informativeness principle, consider the use of deductibles in the following insurance setting. An insured can take an action to prevent an accident (a break-in, say). But given that the accident happens, the amount of damage it causes does not depend on the preventive measure taken by the insured. In this case it is optimal to have the insured pay a deductible if the accident happens as an incentive to take precautions, but the insurance company should pay for all the damages, regardless of the amount. This is so because the occurrence of the accident contains all the relevant information about the precautions the agent took. The amount of damage does not add any information about precautions.

B. Implications of the Informativeness Principle

Several implications follow directly from the informativeness principle.

1. Randomization is suboptimal. Conditional on signal x, one does not want to randomize the payment to the agent, because by eliminating the

randomness without altering the agent's utility at x gives the principal a higher payoff. Randomization may be optimal if the utility function is not separable.[8]

2. Relative performance evaluation (Baiman and Demski, 1980, Holmström, 1982) is valuable when the performance of other agents tells something about the external factors affecting the agent's performance, since information about ε paired with measured performance x is informative about the agent's action e. The important tournament literature initiated by Lazear and Rosen (1981) studies relative performance evaluation in great depth.
3. The controllability principle in accounting states that an agent's incentive should only depend on factors that the agent can control. The use of relative performance evaluation seems to violate this principle, since the agent does not control what other agents do. The proper interpretation of the controllability principle says that an agent should be paid based on the most informative performance measure available. Since x already depends on outside factors (captured by ε) anything correlated with ε can be used to filter out external risk, making the adjusted performance measure more informative.
4. A sufficient statistic in the statistical sense is also sufficient for designing optimal incentive contracts. For instance, the sample mean drawn independently from a normal distribution with known variance is a sufficient statistic for mean effort and can be used to evaluate performance (Holmström 1982).
5. Optimal incentive pay will depend on lagged information if valuable information comes in with delay.

Bebchuk and Fried (2004) among others have argued that CEOs should not be allowed to enjoy windfall gains from favorable macroeconomic conditions. Appealing to the informativeness principle, they advocate the use of relative performance evaluation as a way to filter out luck. In many cases this is warranted, but not without qualifications. In a multitasking context relative performance evaluation may distort the agent's allocation of time and effort. When agents that work together are being compared against each other, cooperation is harmed or collusion may result (Lazear 1989). Filtering out the effects of variations in oil prices on the pay of oil executives will in general not be advisable (at least fully) because this would distort investment and other critical decisions.

Longer vesting times, causing lagged information to affect pay, are unlikely to negatively distort CEO behavior. On the contrary, too short or no vesting aggravate problems with short-termism and strategic information release. Allowing executives to sell incentive options so quickly in the 1990s was clearly unsound and the current vesting periods may still be too short.

C. Puzzles and Shortcomings

The informativeness principle captures the central logic of the basic one-dimensional effort model. In doing so it helps explain some puzzling features of the basic model and also its main shortcomings.

Surprisingly, the optimal incentive scheme need not be nondecreasing even when $x = e + \varepsilon$. The reason is that higher output need not signal higher effort despite first-order stochastic dominance. The characterization (4) shows that the optimal incentive scheme is always monotone in the likelihood ratio $l(x)$ and therefore monotone in x if and only if the likelihood ratio is monotone, in x. Suppose the density has two humps and the difference $e_H - e_L$ is small enough so that the two density functions will cross each other more than once (the humps are interlocked). This creates a likelihood ratio that is non-monotone, implying that there exist two values x such that the larger value has a likelihood ratio below one, speaking in favor of low effort, while the lower value has a likelihood ratio above one suggesting high effort. In line with inference, the agent is paid more for the lower value than the higher value. One can think of the two humps as two states of nature: bad times and good times. The higher outcome suggests that the good state obtained, but conditional on a good state the evidence suggests that the agent slacked.[9]

One can get around this empirically implausible outcome by assuming that the agent can destroy output, in which case only nondecreasing incentives are relevant (Innes 1990). Or one can assume that the likelihood ratio is monotone, a common property of many distribution functions and a frequently used assumption in statistics as well as economics (Milgrom 1981). But the characterization in (4) makes clear that the basic effort model cannot explain common incentive shapes. The universal use of piece rates for instance cannot be due to similar likelihood ratios.

The inference view also helps us understand a troubling example in Mirrlees ([1975]1999). Mirrlees studies the additive production function $x = e + \varepsilon$ where ε is normally distributed with mean zero and a constant variance. In other words, the agent chooses the mean of a normal distribution with constant variance. He analyzes the problem with a continuous choice of effort, but the main intuition can be conveyed with just two actions.

So, assume as before that the agent can choose between low (L) or high (H) effort with a (utility) cost-differential equal to $c > 0$. Note that the likelihood ratio $f_L(x)/f_H(x)$ with the normal distribution goes to infinity as x goes to negative infinity. Therefore, the right-hand side of (4), which is supposed to characterize the optimal solution, must be negative for sufficiently small values of x, because the values for λ and μ are strictly positive when high effort is implemented as discussed earlier. But that is inconsistent with the left-hand side being strictly positive for all x regardless of the function $s_H(x)$. How can that be? Mirrlees conjectured and proved that this must be because the first-best solution can be approximated arbitrarily closely, but never achieved.[10]

In particular, consider a scheme $s(x)$ which pays the agent a fixed amount above a cutoff z, say, and punishes the agent by a fixed amount if x falls below z. If the utility function is unbounded below, Mirrlees shows that one can construct a sequence of increasingly lenient standards z, each paired with an increasingly harsh punishment that maintains the agent's incentive to work hard and the agent's willingness to accept all contracts in the sequence. Furthermore, this sequence can be chosen so that the principal's expected utility approaches the first-best outcome.

One intuition for this paradoxical result is that while one cannot set $\mu = 0$ in (4), one can approach this characterization by letting μ go to zero in a sequence of schemes that assures that the right-hand side stays positive.

The inference logic suggests an intuitive explanation. Despite the way the normal distribution appears to the eye, tail events are extremely informative when one tries to judge whether the observed outcome is consistent with the agent choosing a high rather than a low level of effort. The likelihood ratio $f_L(x)/f_H(x)$ goes to infinity in the lower tail, implying that an outcome far out in the tail is overwhelming evidence against the H distribution. Statistically, the normal distribution is more akin to the moving box example discussed in connection with first-best outcomes. Punishments in the tail are effective, because one can be increasingly certain that L was the source.

III. TOWARD A MORE REALISTIC MODEL

A. Why Are Incentive Schemes Linear?

One could brush aside Mirrlees' example as extremely unrealistic (which it is) and proceed to study linear contracts because they are simple and widely used. But this would defeat the purpose of studying general incentives. An additive production function with a normal error term is a very natural example to study. The unrealistic solution prods us to look for a more fundamental reason for linearity than simplicity and to revisit the assumptions underlying the basic principal-agent model. This led us to the analysis in Holmström and Milgrom (1987).

The optimum in the basic model tends to be complex, because there is an imbalance between the agent's one-dimensional effort space and the infinite-dimensional control space available to the principal.[11] The principal is therefore able to fine-tune incentives by utilizing every minute piece of information about the agent's performance resulting in complex schemes.

The bonus scheme that approaches first-best in Mirrlees' example works for the same reason. It is finely tuned because the agent is assumed to control only the mean of the normal distribution. If we enrich the agent's choice space by allowing the agent to observe some information in advance of choosing effort, the scheme will perform poorly. Think of a salesperson that is paid a discrete bonus at the end of the month if she exceeds a sales target. Suppose she can see her progress over time and adjust her sales effort accordingly. Then, if the month is about to end and she is still far from the sales target, she may try to move potential sales into the next month and get a head start towards next month's bonus. Or, if she has already made the target, she may do the same. This sort of gaming is common (see Healy 1985 and Oyer 1998). To the extent that such gaming is dysfunctional one should avoid a bonus scheme.

Linear schemes are robust to gaming. Regardless of how much the salesperson has sold to date, the incentive to sell one more unit is the same. Following this logic, we built a discrete dynamic model where the agent sells one unit or none in each period and chooses effort in each period based on the full history up to that time. The principal pays the agent at the end of time, basing the payment on the full path of performance. Assuming that the agent has an exponential utility function and effort manifests itself as an opportunity cost (the cost is financial), it is optimal to pay the agent the same way in each period—a base pay and a bonus for a sale.[12] The agent's optimal choice of effort is also the same in each period, regardless of past history. Therefore, the optimal incentive pay at the end of the term is linear in total sales, independently of when sales occurred.[13]

The discrete model has a continuous time analog where the agent's instantaneous effort controls the drift of a Brownian motion with fixed variance over the time interval [0,1]. The Brownian model can be viewed as the limit of the discrete time model under suitable assumptions.[14] It is optimal for the principal to pay the agent a linear function of the final position of the process (the aggregate sales) at time 1 and for the agent to choose a constant drift rate (effort) at each instant. Given this, the distribution of the final position of the process at $t = 1$ is normally distributed with the agent's constant choice of effort determining the mean and the variance being a constant. In other words, we are back to the Mirrlees example $x = e + \varepsilon$ with ε normally distributed and the agent choosing e. The only difference is that we can limit the principal's choice to a linear incentive scheme with the optimal slope and constant term to be determined.

Remarkably, by enriching the agent's choice judiciously, we have arrived at a very simple solution to the Mirrlees problem.[15]

B. Extensions and Alternative Approaches

There is an emerging literature that studies the robustness of linear incentive schemes, where robustness is defined as the max–min outcome in a larger environment, which may or may not be known to the principal. Hurwicz and Shapiro (1978) is the first paper in this vein showing that the linear 50–50 split widely observed in sharecropping can be rationalized along these lines. Diamond (1998), Edmans and Gabaix (2011), Chassang (2013) and Carroll (2015) are more recent variants on this theme. Carroll's paper is especially elegant and at heart simple enough to be applied in richer economic environments. The paper captures robustness in a strong sense (a guaranteed minimum payoff for the principal in a largely unknown environment), but it does not appear as tractable as the Holmström-Milgrom model, at least not yet.

Yuliy Sannikov has greatly advanced and popularized the use of continuous time models to study incentive problems that are both tractable and relevant. Sannikov (2008) solves a general, nonstationary agency problem using powerful techniques from stochastic control theory. These techniques require a keen eye for finding just the right assumptions to make the analysis go through, but the payoff can be high as illustrated in Edmans et al.'s (2012) dynamic model of CEO compensation. The model makes a compelling case for using dynamic incentive accounts that keep funds in escrow and adjust the ratio of debt and equity in response to incoming information. Intuitively, and also in practice, the solution makes a lot of sense. Another example of a continuous time model that resonates with reality is DeMarzo and Sannikov (2006).

C. The Linear Model with One Task

It is straightforward to solve the Mirrlees example with a linear incentive scheme $s(x) = \alpha x + \beta$. The agent's payoff is normally distributed with mean $\alpha e + \beta$ and variance $\alpha^2\sigma^2$. Given an exponential utility function, the agent's utility can be written in terms of his certain equivalent:

$$CE_A = \alpha e + \beta - \tfrac{1}{2}r\alpha^2\sigma^2.$$

The agent has mean-variance preferences where r is the charge for risk bearing. The agent chooses e so that the marginal cost of effort equals the marginal return: $c'(e) = \alpha$.

The principal is risk neutral so her certain equivalent is:

$$CE_P = (1 - \alpha)e - \beta.$$

Since the model is one with transferable utility one can first solve the optimal slope α by maximizing the sum of the two certain equivalents $CE_A + CE_P$. This gives the maximum total surplus which then is divided between the parties using β. Using the first-order condition for the agent's choice of e we get then the optimal value of α

$$\alpha^* = [1 + r\sigma^2 c'']^{-1}, \tag{6}$$

where the dependence of c''on e has been suppressed (c'' is constant if the cost function is quadratic).

The logic of this model is refreshingly simple. The agent works harder the stronger the incentive (higher α). According to (6) the optimal incentive strength α^* always falls between zero (no incentive) and one (first-best incentive), because of the risk that the agent has to bear. If the agent is more risk-averse or the performance measurement is less precise the financial incentive is weaker. What about c''? From the agent's first order condition we get $1/c'' = de/d\alpha$. The derivative measures the agent's responsiveness to an increased incentive. The agent is more responsive to incentives if the cost function is flatter (smaller c''), resulting in a higher commission rate as seen in (6).

The one-dimensional action of the agent and the one-dimensional control of the principal are evenly matched, making this model very well-behaved. One can extend the model in many dimensions. One can study the costs and benefits of (jointly chosen) different projects, production technologies and monitoring systems, for instance, and get simple answers. The most interesting variation in such thought experiments concerns the opportunity cost function $c(e)$. There are many ways in which the principal can vary the agent's opportunity cost function so that the cost of incentives is reduced. This insight was central in initiating my work on multitasking with Paul Milgrom (Holmström and Milgrom 1991, 1994.)

IV. MULTITASKING

Multitasking—the reality that an agent's job consists of many tasks—led to a major change in mindset and focus. Instead of studying how to get the agent to work hard enough on a single task, attention turned to how the agent allocates his effort across tasks in a manner that aligns with the principal's objectives. When tasks are interdependent, the optimal design needs to consider the agent's incentives in totality. Knowing the agent's full portfolio of activities—what his authority and responsibilities are—is essential for designing a coherent, balanced solution that takes into account the interdependencies. This is challenging when easy-to-measure and hard-to-measure activities compete for the agent's attention or if the available performance measures are poorly aligned with the principal's objectives.

A. Easy versus Hard to Measure Tasks

Consider the case of an agent with two tasks. One task can be perfectly measured—think about quantity sold. The other task is very hard to measure—think about the reputation of the firm. There may be some measures available for the latter, for instance consumer feedback. But such information is selective and often biased. People with unhappy sales experiences are likely to complain more often than people with happy experiences. Some customers may just have had a bad day. And some important customers may not have the time for feedback. All this makes it hard to assess how consumer feedback genuinely translates into valuable reputation down the road.

To be a bit more formal about this, we can use a multitask extension of the linear model (see the Appendix for a general version). Let the performance in each task be separately measured, quantity by $x_1 = e_1$ and reputation by $x_2 = e_2 + \varepsilon_2$. There is no noise term ε_1 because I assume that quantity can be measured perfectly. The variance of ε_2 is larger the noisier the consumer feedback is.

Let the principal's objective be $B(e_1,e_2) = p_1e_1 + p_2e_2$, where p_1 and p_2 measure the value of quantity and reputation and suppose for a moment that the agent's cost function is separable: $C(e_1,e_2) = c_1(e_1) + c_2(e_2)$. In this case the agent's incentives can also be analyzed separately and it is optimal to set each coefficient independently according to the formula for the single-task case. The commission rate on x_1 will be $\alpha_1= p_1$, giving first-best incentives for quantity choice, while the commission rate on x_2 will be $\alpha_2 < p_2$, trading risk against incentives for a second-best outcome.

Now, assume that the tasks are substitutes: the more time the agent spends on one task, the higher is his marginal cost of spending time on the other task (the cross partial $C_{12} > 0$.) In this case it is no longer optimal to give the agent first-best incentives for choosing quantity ($\alpha_1 = p_1$), because lowering α_1 slightly will at the margin cost nothing in lost value on the first task, but will be strictly beneficial for providing incentives on the second task. The agent can be given the incentive to supply the same amount of quality for a lower α_2, which reduces the cost of risk. Or alternatively, when α_1 is lower the marginal cost of spending time on quality has gone down, making the agent spend more time on that task. Either way, lowering the incentive on quantity is advantageous for the supply of quality, by how much depends on the precision with which quality can be measured and how substitutable the tasks are in the cost function. If the two tasks are perfect substitutes (i.e., the agent just allocates his time between the two tasks so the cost function is $C(e_1 + e_2)$) and attention to quality is essential (this will require a nonlinear benefit function B), any incentive on quantity will have to be matched by a correspondingly strong incentive on quality so that $\alpha_1 = \alpha_2$. This makes the agent indifferent between spending time on either task and he will choose whatever allocation is best for the principal. The cost is that the incentive for both tasks will have to be reduced because quality is poorly measured. If there is no quality measure at all (i.e., the variance of ε_2 is infinite) then no incentive for either task is optimal. If $C'(0) < 0$, the agent will still choose a positive level of total effort.

B. Misalignment and Manipulation

Misalignment is an important variation on the multitask theme (Baker 1992, 2002). It can also be analyzed with the general multitask model.

Suppose again that the principal's value is $B(e_1,e_2) = p_1e_1 + p_2e_2$, and there is just one performance measure $x(e_1,e_2) = g_1e_1 + g_2e_2$. Even though there is no uncertainty in the measure, there is a nontrivial incentive problem if the vectors p and g are not aligned. For example, suppose $x = e_1 + e_2$ but $B(e) = e_1$. The principal only values effort on the main task 1, but the agent can produce measured output x with e_1 as well as e_2. The second activity can be interpreted as "manipulation," which the agent may feel a moral dislike for, but if the principal pushes too strongly on measured performance x in the hope of getting the agent to work hard on the main task, the agent will engage in manipulation as well. The principal ends up compensating the agent for worthless performance, so misalignment causes waste, more so the higher α is. It will be optimal to set $\alpha < 1$, because at $\alpha = 1$ the marginal cost of reducing α is strictly negative (e_1 is first-best while e_2 is not).

Both variants of multitasking, disparities in measurement errors as well as misalignment between performance measures and the value created, are relevant variants of the general multitask model. It depends on the context which one is more natural to use.

C. "You Get What You Pay For"

Gross manipulation of performance measures was behind the recent Wells Fargo scandal (Tayan 2016). Wells Fargo had avoided the banking scandals associated with the financial crisis. It was known for prudence in lending, making profits by emphasizing its retail banking and customer service. Its branch managers were highly incentivized toward cross-selling: getting its regular banking customers to buy a range of products, such as credit lines. This part of their banking business had steadily grown and been highly profitable. But continued growth also required new customers and eventually, as the sales goals were tightened (the branch managers' performance was measured daily), some of the managers opened new accounts for its customers without the customers' knowledge. Shell accounts were like a second activity in the two-task model described earlier. It improved measured performance and generated bonuses, but since there was no real activity in the accounts this activity generated minuscule profits for Wells Fargo ($2.6 million according to Tayan 2016). Shell accounts caused minimal costs for customers (an estimated $2.50 per account), but they were of course illegal. Eventually the scam was discovered, causing the firing or resignation of thousands of employees and eventually also the resignation of Mr. Stumpf, the CEO. Wells Fargo had to pay $185 million in penalties, but the biggest cost by far was the enormous damage to their stellar reputation.[16]

The explanations for the BP oil spill in the Gulf of Mexico point in many directions, but the fact that BPX—the exploration arm of BP—was encouraged to be more aggressive in its exploration activity was likely one of the culprits (Garicano and Rayo 2016). Measurable results were pushed hard (implicitly or explicitly), compromising safety. This is an example where there are many activities, some easy to measure (successful exploration), and others not so easy. While safety can be monitored, the intensity of monitoring whether rules are being strictly followed is not easy. Slight delays in service or in repair of faulty parts, especially in the case of backup systems and checks, may appear to carry minimal risk and therefore be subject to trade-offs under pressure.

Several scandals are associated with the introduction of test-based incentives for teachers. One of the most notorious ones occurred in the Atlanta public school system, where the results on the Competency Tests showed irregularities in test scores (extraordinary gains and losses in a single year). An investigation revealed that 44 out of 56 schools cheated on the 2009 Competency Test, and 178 teachers and principals were found guilty of correcting student answers. In 2015, 11 teachers were convicted of racketeering and received stiff sentences, including jail terms.

The scandal was an unintended consequence of the No Child Left Behind Act of 2001 and illustrates how good intentions can go wrong. As in the case of Wells Fargo and many similar cases, those guilty of wrong-doing describe their acts as the result of feeling excessive pressure to perform.

The debate about the costs and benefits of testing and rewarding teachers and schools is ongoing (Neal 2011). The issue is less about cheating and more about other costs such as teaching to the test. Like Stephen Kerr (1975), in his well-known article "On the Folly of Rewarding A, While Hoping for B," the multitasking model warns against aggressive use of rewards for easy-to-measure subjects like mathematics at the expense of developing important social and other hard-to-measure skills. But economics goes beyond Kerr's criticism of rewarding narrow goals by studying the many substitute instruments that firms can use quite effectively. I will return to these after summarizing the main lessons from the multitask problem.[17]

D. Low-Powered Incentives

A first lesson from multitasking is that there are *two ways to provide incentives* for a hard-to-measure activity like supply of quality. The *direct way* is to pay for measured quality, however poor the measure is. The *indirect way* is to reduce the incentive on quantity, because this reduces the opportunity cost of providing quality. How effective the indirect way is depends on how poor the quality measure is and how strongly the two tasks compete with each other for the agent's time. In the pure time-allocation case the tasks compete directly with each other for the agent's attention, giving rise to a maximal substitution effect.

A second lesson is that measurement problems can explain why *even perfectly measured activities may be left without any incentive*. The fact that the agent can control a measure well, because the measure is accurate with respect to a task or a subset of tasks, does not imply that it should be used. Sometimes the exact opposite is true. Employees tend to like accounting measures that they can control such as costs, but these measures are often partial and therefore poorly aligned with true value creation, in which case low-powered incentives on these measures are called for.

The third lesson is the most general. Because the opportunity cost of providing incentives for any given task depends on all the incentives for the other tasks, one should always *consider the full portfolio of agent tasks when designing incentives*.

V. ALTERNATIVE INCENTIVE INSTRUMENTS IN FIRMS

Firms use pay-for-performance schemes rather sparingly. Most employees are on fixed pay. Williamson (1985) offers one explanation: the integrity of performance measures within the firm is weaker than in the market. Multitasking offers a complementary explanation. Firms have access to many substitutes for pay-for-performance incentives that are not easily accessible through market contracting. Foremost among them is the ability to control work through job assignments, job designs and a variety of implicit and explicit rules that the firm sets.

A. Rules and Constraints

Consider the case of working at home versus coming to the office.[18] Assume that the task that the agent is assigned—one task for simplicity—can be done just as easily at home as in the office. The only difference between the two options is that the agent has other activities that he can engage in at home, activities that provide private benefits. He can work on other projects at home, do sports, watch TV, rest, play with his kids and so on. We can think of these privately valuable activities as offsets against the cost of working for the principal. The downside is that they reduce the time the agent can spend on the principal's task.

If the agent were paid the value of the output he produces, there would be no problem working from home. As a residual claimant, the agent would do the socially correct calculation on how much time to spend on private activities and how much on the principal's task. But this contract would also force the agent to bear risk. If risk is costly and the incentive power therefore reduced, the agent's commission will not fully reflect the marginal value of his activities. The principal's task will, as a result, become less competitive against the private activities. Suppose each of the private activities can either be shut down or the agent will be free to pursue the activity as he wishes—a zero-one decision for each activity—which activities should be allowed and which ones shut down? Note that allowing private activities is a form of compensation for the agent. So, the principal will also benefit from letting the agent undertake an activity if the total surplus is increased in that way. Given a fixed commission, if a private task is shut down, the agent will spend the time freed up on the principal's task. So, whether the task should be shut down is just a question of whether the agent's chosen amount of a private activity, given the commission rate set by the principal, will be worth more when spent on the principal's task.

The lower the commission, the less time the agent will spend on the principal's task and the more on his private tasks. As a result, the marginal value of the principal's task goes up while the marginal values on the agent's private tasks go down as the commission rate goes down—the agent is spending too much time on private tasks and too little on the principal's. So, excluding private activities becomes more important, the lower the commission rate. Constraints on private activities act as substitutes for monetary incentives by increasing the amount of time the agent spends on the principal's task. In this example, the restrictions are placed in decreasing order of importance of the tasks (as measured by the average private value of tasks).

The agent's incentive will be weaker if the performance in the main task is measured with lower precision (the variance of ε is higher) and this will imply that more tasks will be excluded in order to substitute for the weaker incentives.[19] Responsibility and freedom to choose co-move as the precision of measurement varies.[20]

Beyond task restrictions, other instruments can also be important. For example, bureaucratic rules in organizations are as prevalent as the complaints about them. Bureaucracy is detested because it discourages initiative and innovation. Too often people fail to see that there is a purpose for most bureaucratic rules. In the first instance, these rules are not symptoms of an organizational disease, but rather part of the solution to complicated incentive problems. Because pay-for-performance tends to be a costly incentive instrument in the firm, due to problems with performance measurement, bureaucratic rules and restrictions will often serve as effective substitutes (Holmström 1989).

Firms also place a number of other restrictions on its employees that limit what they can do. One of the most obvious and universal rules (until recently, perhaps) is the requirement that an employee cannot work for other firms. Another common requirement is that employees have to come to the office for set hours. Naturally, these constraints are driven by other considerations, too, but their effect on employee incentives is nontrivial.

B. Job Design

Another important dimension of job design is the distribution of tasks between workers.

Go back to the two-task case with one agent. One task is perfectly measured (a routine task) and the other task is measured with considerable noise (an innovative task). The agent's cost function depends only on total time spent. I concluded that providing low-powered incentives on the routine task is appropriate as an incentive for innovation.

What if one could split the two tasks between two agents? One agent takes care of the routine task while the other specializes on the innovative task. This avoids the tension between routine and innovative tasks. The routine task can be provided with high-powered incentives while the innovative task has low-powered incentives. More generally, one can show (Holmström and Milgrom 1991) that if there is a continuum of tasks (to avoid integer problems) with varying measurement costs, then one should give all the hard-to-measure tasks (defined by a cutoff) to one agent and the easy-to-measure tasks to the other agent. The hard-to-measure tasks are provided with equal low-powered incentives, while the easy-to-measure tasks are provided with equal high-powered incentives. The more risk-averse agent takes on the low-risk tasks.

More specialization may be one way to deal with the teaching-to-the-test dilemma that is so hotly debated. The idea has been promoted by Neal (2011) based on the multitasking model. There is of course already a significant degree of specialization in teaching, but one may want to increase as well as shift the responsibility of soft skills to teachers that specialize in these. These teachers should not be held responsible for narrow test results.

There is an important additional lesson hidden in the logic of specialization. In the model discussed, no task is shared because that would require incurring risk for both agents in the given task. If one agent is already incentivized to do half of the project, incurring the risk cost associated with adequate incentives, there is no additional risk cost to have him do the whole task. While this result depends partly on our additive production function and the continuum of tasks, the logic is clear. There is a fixed cost to having a second agent join the project, because both have to be incentivized. This feature is of more general validity and may be one reason for individual responsibility (unity of command).

The same principle is relevant for job rotation. One benefit of job rotation is that it offers a form of relative performance evaluation: having two agents work in the same task gives information about the difficulty of the task and this is valuable for evaluating each agent's performance in line with the informativeness principle. But there is also a downside from an incentive point of view. If performance information comes in with a lag, one wants to make the incentive of the first agent depend on the results coming in during the second agent's shift. This is similar to two agents sharing the same task as discussed above. The first agent is partly incentivized for the later periods because of the delayed information, and from this point of view it would be efficient to let him continue.

C. Career Incentives

People want to be appreciated for their work. Career concern models (Holmström 1999a; Dewatripont, Jewitt, and Tirole 1999) can capture the desire for appreciation through a dynamic signaling mechanism. Employers are learning about the value of employees from past performance. The value is initially unknown (also for the employee) but will be revealed at least partly over time. Employees try to make a favorable impression because better performance can lead to promotions, higher pay, more status, and other rewards in the future.

Career concerns can be expected to be more powerful in firms than in markets, because in firms the employees typically know whom they should impress. Firms can exploit this motivation by indicating what is desired from an employee. In employee surveys, one of the most common complaints is that the employees would like a clearer understanding of what is expected of them. If they only knew what the boss wants, they could work on making a good impression in those dimensions. The craving for appreciation and the desire to impress superiors explains why a mere change in the accounting system can have a big impact on the behavior of employees. Key performance indicators are important because they signal what management wants. Alone or combined with relatively modest bonuses, they can have a surprisingly strong effect. Putting money behind a measure conveys a stronger message of what is expected. Employees also respond to mission statements as in Dewatripont, Jewitt and Tirole (1999).

Career incentives can be so strong as to obviate any need for financial incentives. In fact, they can be too strong and lead to wasteful pandering and influence activities (Milgrom and Roberts 1988; Holmström and Ricart i Costa 1986; Prendergast 1993). Just like financial incentives, career incentives can be imbalanced, because some indicators of performance are more visible and therefore more salient than others. The firm can regulate visibility through the access the employee has to superiors and they can institute pay and promotion practices that do not respond to performance as strongly. Basing salary increases and promotions on seniority mutes career incentives that are perceived to be too strong or lead to short-sighted behavior. In general, the firm's overall hierarchy and its promotion policies are powerful incentive drivers without direct counterparts in the market.[21]

The problems of multitasking tend to manifest themselves in much the same way when career concerns are the driver of employee behavior in place of explicit bonuses. This is why the earlier discussion of multitasking assuming commissions as incentive instruments provides useful guidance even though commissions are relatively less used by firms.

VI. TWO INCENTIVE SYSTEMS: EMPLOYMENT VERSUS CONTRACTING

The preceding discussion makes clear the point that firms have access to a variety of incentive instruments that can compensate for low-powered pay-for-performance incentives. In this section I want to highlight two implications for the design and use of incentive systems.

The first is that structuring efficient incentives within the firm requires a concerted use of all available incentive instruments. This parallels the earlier lesson from multitasking that commissions for all tasks need to be considered in concert. The second point is that incentives in firms and in markets form two logically coherent systems, each with its distinct comparative advantage.

In Holmström and Milgrom (1994) we illustrate both points by applying our multitask model to the choice between employment and contracting. We want to explain the findings of Anderson and Schmittlein (1984) and Anderson (1985), who studied how industrial selling is organized in the electronics industry. Some sales agents work as independent representatives, others as employees. A firm often uses both forms of sales organization. What determines the choice between the two forms of sales and how are the incentives structured within each alternative?

Anderson and Schmittlein find that the most important measures determining the choice between employment and independent contracting were "the difficulty of evaluating performance" and "the importance of non-selling activities." Performance evaluation is difficult for complex sales, while non-selling activities are important when cooperation with other sales agents is desired. Increases in both of these measures made employment more likely. Moreover, independent representatives were compensated entirely by commission and they were allowed to represent other manufacturers, while sales employees were on fixed wages and could not sell the products of other manufacturers.

These findings can be explained in our standard multi-tasking setting where tasks compete for the sales agent's attention. The fact that independent agents are free to sell the products of other manufacturers and have strong incentives parallels the earlier argument about giving an agent the freedom to pursue private tasks if the commission rate is high, but not if the commission rate is low. The difficulty of evaluating performance drives the commission rate down, as does the desire to have the agent allocate time to non-selling (cooperative) activities.

The model includes two types of selling activities: direct selling, which affects short-term returns and is influenced by commissions, and indirect selling which increases future sales. We assume that the latter returns can only be transferred through ownership. This is a reduced-form way to introduce a role for ownership (Grossman and Hart 1986). An independent sales agent owns the long-term returns; a sales employee does not. The remedy for the loss of long-term incentives is to pay the sales employee a fixed wage. This makes the sales employee willing to allocate attention to tasks in whatever way the firm desires, at the cost of weaker overall initiatives.

We have, then, two coherent incentive systems as found by Anderson and Schmittlein. An independent contractor (sales representative) owns the long-term returns, is paid high commissions for direct sales and is free to represent other firms. A sales employee does not own the long-term returns, is paid a fixed wage, and cannot represent other firms. And consistent with the evidence, employment is favored if direct selling is hard to measure or if non-selling activities are valuable.

Another way to introduce a distinction between employment and contracting is to assume that a firm can restrict the tasks of a sales employee more easily than it can for an independent representative (Holmström 1999b). Both alternatives show that the theory of incentive contracts and the theory of property rights are complementary. Together they can provide a richer perspective on organization.

VII. CONCLUSION

Let me close with a summary of the intellectual journey I have described. When I started to study moral hazard, the main paradigm was the basic one-dimensional effort model. Despite the seeming simplicity of this model, it behaved in ways that were perplexing at times. The Informativeness Principle revealed the basic logic of the model, providing useful insights about the value of information at the same time as it made clear that the basic model cannot explain the shape of common incentive schemes.[22] This led Paul Milgrom and me to ask why incentive schemes are linear. Our willingness to listen to the basic agency model introduced us, rather serendipitously, into the world of multitasking, which opened up a rich set of issues and opportunities. The value of low-powered incentives in the context of multitasking explains why firms make so little use of explicit bonuses and instead use alternatives like job design and bureaucratic rules to construct coherent incentive systems that are very distinct from the way incentives are designed in the market.

The firm's comparative advantage relative to markets rests partly with its unique ability to use low-powered incentives combined with constraints. This explains why bringing the market inside the firm is such a misguided idea, something I failed to understand at Ahlström and advocates of market-like incentives in firms seem to miss today.

APPENDIX

A1. The General Multitask Lab

In Holmström and Milgrom (1987) we study a multi-dimensional Brownian process in which the agent can choose drift rates independently, at a cost that depends on the rates chosen. This model features an optimal solution that is linear in the different dimensions of the Brownian process, and just as in the one-dimensional case the optimal coefficients can be solved through a static model where the agent has several tasks to perform, each one corresponding to one of the dimensions of the Brownian process. In its most general form the static multitask model has the following elements:

- The agent chooses "inputs" $e = (e_1,\ldots,e_n)$.
- There are m measures of performance $x_i = k_i(e) + \varepsilon_i$, $i = 1,\ldots,m$. The measurement errors follow a joint normal distribution with mean vector 0 and variance-covariance matrix Σ. The "production" functions k_i determine how the agent's choices map into the mean of the corresponding performance measure.
- The principal's benefit function is $B(e)$ and the agent's cost function is $C(e)$.
- The optimal incentive scheme is linear: $s(x) = \Sigma_i \alpha_i x_i + \beta$, with commission rates α_i and salary β.

The agent's behavior is characterized by the first order conditions

$$\sum_i \alpha_i \partial k_i(e)/\partial e_j = \partial C(e)/\partial e_j \text{ for every } j.$$

It is readily seen that the misalignment model is a special case of this model. The model also encompasses cases where a risk neutral agent observes a random signal θ before acting as in Baker (1992).

See Holmström and Milgrom (1991) for further details on the solution and variations.

REFERENCES

Anderson, Erin. 1985. "The Salesperson as Outside Agent or Employee: A Transactions Cost Analysis," *Marketing Science* **4** (3): 234–54.

Anderson, Erin and David C. Schmittlein. 1984. "Integration of the Sales Force: An Empirical Examination," *RAND Journal of Economics* **15** (3): 385–95.

Baiman, Stanley, and Joel S. Demski. 1980. “Economically Optimal Performance Evaluation and Control Systems,” *Journal of Accounting Research* **18** (Supplement): 184–220.

Baker, George P., 1992. “Incentive Contracts and Performance Measurement,” *Journal of Political Economy* **100** (3): 598–614.

Baker, George. 2002. “Distortion and Risk in Optimal Incentive Contracts,” *Journal of Human Resources* **37** (4): 728–51.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 1994. “Subjective Performance Measures in Optimal Incentive Contracts,” *Quarterly Journal of Economics* **109** (4): 1125–56.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 1999. “Informal Authority in Organizations,” *Journal of Law, Economics, and Organization* **15** (1): 56–73.

Baker, George, and Bjorn Jorgensen. 2003. “Volatility, Noise and Incentives.” http://www.bristol.ac.uk/media-library/sites/cmpo/migrated/documents/baker.pdf.

Bebchuk, Lucian A., and Jesse M. Fried. 2004. *Pay without Performance: The Unfulfilled Promise of Executive Compensation*. Cambridge, MA: Harvard University Press.

Carroll, Gabriel. 2015. “Robustness and Linear Contracts,” *American Economic Review* **105** (2): 536–63.

Chaigneau, Pierre, Alex Edmans, and Daniel Gottlieb. 2014. “The Informativeness Principle Under Limited Liability,” National Bureau of Economic Research Working Paper 20456.

Chassang, Sylvain. 2013. “Calibrated Incentive Contracts,” *Econometrica* **81** (5): 1935–71.

DeMarzo, Peter M., and Yuliy Sannikov. 2006. “Optimal Security Design and Dynamic Capital Structure in a Continuous-Time Agency Model,” *Journal of Finance* **61** (6): 2681–2724.

Dewatripont, Mathias, Ian Jewitt, and Jean Tirole. 1999. “The Economics of Career Concerns, Part II: Application to Missions and Accountability of Government Agencies,” *Review of Economic Studies* **66** (1): 199–217.

Diamond, Peter. 1998. “Managerial Incentives: On the Near Linearity of Optimal Compensation,” *Journal of Political Economy* **106** (5): 931–57.

Edmans, Alex, and Xavier Gabaix. 2011. “Tractability in Incentive Contracting,” *Review of Financial Studies* **24** (9): 2865–94.

Edmans, Alex, Xavier Gabaix, Tomasz Sadzik, and Yuliy Sannikov. 2012. “Dynamic CEO Compensation,” *Journal of Finance* **67** (5):1603–47.

Garicano, Luis, and Luis Rayo. 2016. “Why Organizations Fail: Models and Cases,” *Journal of Economic Literature* **54** (1): 137–92.

Gjesdal, Frøystein. 1982. “Information and Incentives: The Agency Information Problem,” *Review of Economic Studies* **49** (3): 373–90.

Grossman, Sanford J., and Oliver D. Hart. 1983. “An Analysis of the Principal-Agent Problem,” *Econometrica* **51** (1): 7–45.

Grossman, Sanford J., and Oliver D. Hart. 1986. “The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration,” *Journal of Political Economy* **94** (4): 691–719.

Healy, Paul M. 1985. “The Effect of Bonus Schemes on Accounting Decisions,” *Journal of Accounting and Economics* **7** (1–3): 85–107.

Hellwig, Martin F., and Klaus M. Schmidt. 2002. "Discrete-Time Approximations of the Holmström-Milgrom Brownian-Motion Model of Intertemporal Incentive Provision," *Econometrica* **70** (6): 2225–64.

Holmström, Bengt. 1977. "On Incentives and Control in Organizations," PhD diss., Stanford University.

Holmström, Bengt. 1979. "Moral Hazard and Observability," *Bell Journal of Economics* **10** (1): 74–91.

Holmström, Bengt. 1982. "Moral Hazard in Teams," *Bell Journal of Economics* **13** (2): 324–40.

Holmström, Bengt. 1989. "Agency Costs and Innovation," *Journal of Economic Behavior and Organization* **12** (3): 305–27.

Holmström, Bengt. 1999a. "Managerial Incentive Problems: A Dynamic Perspective," *Review of Economic Studies* **66** (1): 169–82.

Holmström, Bengt. 1999b. "The Firm as a Subeconomy," *Journal of Law, Economics, and Organization* **15** (1): 74–102.

Holmström, Bengt, and Paul Milgrom. 1987. "Aggregation and Linearity in the Provision of Incentives," *Econometrica* **55** (2): 303–28.

Holmström, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design," *Journal of Law, Economics, and Organization* **7**: 24–52.

Holmström, Bengt, and Paul Milgrom. 1994. "The Firm as an Incentive System," *American Economic Review* **84** (4): 972–91.

Holmström, Bengt, and Joan Ricart i Costa. 1986. "Managerial Incentives and Capital Management," *Quarterly Journal of Economics* **101** (4): 835–60.

Hurwicz, Leonid, and Leonard Shapiro. 1978. "Incentive Structures Maximizing Residual Gain under Incomplete Information," *Bell Journal of Economics* **9** (1): 180–91.

Innes, Robert D. 1990. "Limited Liability and Incentive Contracting with Ex-ante Action Choices," *Journal of Economic Theory* **52** (1): 45–67.

Kerr, Steven. 1975. "On the Folly of Rewarding A, While Hoping for B," *Academy of Management Journal* **18** (4): 769–83.

Kim, Son Ku. 1995. "Efficiency of an Information System in an Agency Model," *Econometrica* **63** (1): 89–102.

Lazear, Edward P.. 1989. "Pay Inequality and Industrial Politics," *Journal of Political Economy* **97** (3): 561–80.

Lazear, Edward P.. 2000. "Performance Pay and Productivity," *American Economic Review* **90** (5): 1346–61.

Lazear, Edward P., and Sherwin Rosen. 1981. "Rank Order Tournaments as Optimum Labor Contracts," *Journal of Political Economy* **89** (5): 841–64.

Levin, Jonathan. 2003. "Relational Incentive Contracts," *American Economic Review* **93** (3): 835–57.

MacLeod, W. Bentley, and James M. Malcomson. 1988. "Reputation and Hierarchy in Dynamic Models of Employment," *Journal of Political Economy* **96** (4): 832–54.

Milgrom, Paul R. 1981. "Good News and Bad News: Representation Theorems and Applications," *Bell Journal of Economics* **12** (2): 380–91.

Milgrom, Paul R., and John Roberts. 1988. "An Economic Approach to Influence Activities in Organizations," *American Journal of Sociology* **94**: S154–S179.

Milgrom, Paul R., and John Roberts. 1990. "The Economics of Modern Manufacturing: Technology, Strategy, and Organization." *American Economic Review* **80** (3): 511–28.

Milgrom, Paul R. and John Roberts. 1995. "Complementarities and Fit: Strategy, Structure and Organizational Change in Manufacturing," *Journal of Accounting and Economics* **19** (2–3): 179–208.

Mirrlees, J. A. (1975) 1999. "The Theory of Moral Hazard and Unobservable Behaviour: Part I," *Review of Economic Studies* **66** (1): 3–21. Paper completed 1975, published 1999.

Murphy, Kevin J. 1999. "Executive Compensation." In *Handbook Of Labor Economics*, Vol.3B, edited by Orly Ashenfelter and David Card, 2485–2563. New York: Elsevier.

Neal, Derek. 2011. "Design of Performance Pay in Education." In *Handbook of Economics of Education*, Vol. 4, edited by Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, 495–550. Amsterdam: Elsevier.

Oyer, Paul. 1998. "Fiscal Year Ends and Nonlinear Incentive Contracts: The Effect on Business Seasonality," *Quarterly Journal of Economics* **113** (1): 149–85.

Prendergast, Canice. 1993. "A Theory of 'Yes Men'," *American Economic Review* **83** (4): 757–70.

Prendergast, Canice. 2002. "The Tenuous Trade-Off between Risk and Incentives," *Journal of Political Economy* **110** (5): 1071–1102.

Rogerson, William P. 1985. "The First-Order Approach to Principal–Agent Problems," *Econometrica* **53** (6): 1357–67.

Ross, Stephen A. 1973. "The Economic Theory of Agency: The Principal's Problem," *American Economic Review* **63** (2): 134–39.

Sannikov, Yuliy. 2008. "A Continuous-Time Version of the Principal-Agent Problem," *Review of Economic Studies* **75** (3): 957–84.

Shavell, Steven. 1979. "Risk Sharing and Incentives in the Principal Agent Relationship," *Bell Journal of Economics* **10** (1): 55–73.

Spence, Michael, and Richard Zeckhauser. 1971. "Insurance, Information, and Individual Action," *American Economic Review* **61** (2): 380–87.

Stiglitz, Joseph E. 1975. "Incentives, Risk, and Information: Notes Towards a Theory of Hierarchy," *Bell Journal of Economics* **6** (2): 552–79.

Tayan, Brian. 2016. "The Wells Fargo Cross-Selling Scandal," https://www.gsb.stanford.edu/facultyresearch/publications/wells-fargo-cross-selling-scandal.

Williamson, Oliver E. 1985. *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. New York: Free Press.

Wilson, Robert. 1969. "The Structure of Incentives for Decentralization under Uncertainty," In *La Decision: Agregation et Dynamique des Ordres de Preference*, edited by Georges-Théodule Guilbaud, 287–307. Paris: Centre National de la Recherche Scientifique.

NOTES

1. It is common to see companies pay executives a bonus that is linear within a performance interval, but capped both at the top and the bottom; see Murphy (1999). On

the other hand, real estate agents and sales people are paid commissions without an upper bound. Stiglitz's (1975) paper on sharecropping was the first to study linear incentives.

2. Mirrlees ([1975]1999) was the first to use this formulation. It avoids taking derivatives of the endogenous incentive scheme $s(x)$, which may well be non-differentiable a priori.
3. Holmström (1977) proves existence by assuming that $s(x)$ has to be chosen from a finite interval. A more instructive existence proof is provided by Grossman and Hart (1983) when the number of outcomes is finite. The key assumption is that probabilities have to be strictly bounded away from zero for all action choices.
4. This equation is the first-order condition of the Lagrangian with respect to $s(x)$ for each x.
5. If $\mu = 0$, the formula implies a constant $s_H(x)$ in which case the agent would choose L, violating the incentive constraint; so $\mu > 0$. If $\lambda = 0$ constraint (2) is slack and the principal can do better by reducing all utility levels by a constant; so $\lambda > 0$.
6. Typically s_λ does not correspond to optimal risk sharing for the reservation value of $\underline{U}$ in the agent's participation constraint (3), because that problem will have a different λ value. I thank Jörgen Weibull for pointing this out.
7. The informativeness principle as stated applies when there is just one binding incentive compatability constraint (2). If the agent is indifferent among several actions, there will be a Lagrange multiplier μ_j for each binding incentive constraint j. One can extend the inference interpretation to this situation by evaluating the outcome against the likelihood of the agent randomizing across actions of indifference with the Lagrange multipliers providing the (relative) weights of this randomized strategy. I am grateful to Paul Milgrom for suggesting this extension. Gjesdal (1982) and Grossman and Hart (1983) provide a weaker ordering using Blackwell's Theorem, which applies regardless of the number of binding incentive constraints. For other variations, see Kim (1995) and Chaigneau, Edmans and Gottlieb (2014).
8. See Gjesdal (1982).
9. Grossman and Hart (1983) show that the only thing one can say in a general model with discrete outcomes x is that the optimal incentive scheme cannot be such that whenever the agent is paid more, the principal is paid less. There must be one increment in x such that payments co-move.
10. This nonexistence result violates Grossman and Hart's (1983) assumption that probabilities are bounded strictly away from zero.
11. For every x, there is a separate control $s(x)$. The space of all possible functions s is infinite-dimensional.
12. The agent's utility function is $u(m, e) = 1 - \exp[-r(m - c(e))]$, where r is the agent's absolute risk aversion, $c(e)$ is the opportunity cost of effort and m is money. This utility function is multiplicatively rather than additively separable. The reason we chose this function is that income levels do not affect the agent's choice.
13. This seems to violate the informativeness principle: the timing of sales is not used. The reason is that it is optimal for the principal to implement a constant action and for this implementation the timing of sales is irrelevant.

14. See Holmström and Milgrom (1987) for an intuitive derivation and Hellwig and Schmidt (2002) for a rigorous analysis.
15. When the agent can observe his progress and make his choice of effort contingent on the current state, the agent can generate essentially any distribution over the final position of the process at time 1 using a Brownian bridge (which takes the process from its starting point at $t = 0$, to an arbitrary point at time $t = 1$). I am grateful to Michael Harrison for showing me this.
16. So the second activity e_2 actually cost the principal a lot and should therefore enter $B(e)$, but with a big negative sign, reflecting the costs of the reputation loss.
17. There are also examples where firms implement aggressive piece rate plans successfully. Safelite, the dominant firm in the windshield replacement market, is one case in point (Lazear 2000). Lincoln Electric, a welding equipment manufacturer, is another (Milgrom and Roberts 1995). In each case, the introduction of piece rates required a number of matching changes in the organizational structure to avoid the kind of problems described above.
18. A formal treatment of this example can be found in Holmström and Milgrom (1991).
19. Comparative statics results like this one, involving several endogenous variables, will in general require a more complex analysis using monotone methods (see Milgrom and Roberts 1990 and Holmström and Milgrom 1994). Because the exclusion of private tasks only depends on the commission rate, but not directly on the measurement error, this case is straightforward. If instead the value of the principal's task increases, there is a negative direct effect on exclusion, and it is therefore possible that fewer private tasks will be allowed even though the commission rate goes up. There are two ways of increasing the agent's attention to the principal's task: raise the commission rate or exclude private tasks. Both instruments may be used together.
20. In an important paper, Prendergast (2002) observes that empirically we often see that higher risk go together with stronger, not weaker, incentives as just described. The reason for changing conclusion is that risk in his model concerns productivity, not measurement error, and there is therefore value in having the agent respond to changing circumstances. This case can be incorporated in the multitask model by replacing the additive output function with a multiplicative one. The agent's action will then be a contingent strategy and the single measure will be a weighted average of the possible states of nature. Prendergast's model as well as the multitask model can explain the positive co-movement (see Baker and Jorgensen 2003).
21. The related literature on relational contracting studies the scope of informal contracts that can support rules and practices; see e.g. Levin (2003); MacLeod and Malcomson (1988); and Baker, Gibbons and Murphy (1994, 1999).
22. Much attention was paid to problems with the First-Order Approach—the fact that one cannot in general replace the agent's incentive compatibility constraint with a first-order condition. Rogerson (1985) provided a sufficient condition for the first-order approach and efforts continue to refine his work. But it is evident by now that the one-dimensional effort model as such has serious shortcomings.