Incentive Contracts and Performance Measurement
Author(s): George P. Baker
Source: *The Journal of Political Economy*, Vol. 100, No. 3 (Jun., 1992), pp. 598-614
Published by: The University of Chicago Press
Stable URL: http://www.jstor.org/stable/2138733
Accessed: 18/01/2009 18:26

Incentive Contracts and Performance Measurement

George P. Baker

Harvard University

This paper examines the characteristics of incentive contracts in which the agent's payoff is not based on the principal's objective. I show that contracts based on such performance measures will not in general provide first-best incentives, even when the agent is risk neutral. The form of the optimal contract and the efficiency of this contract depend on the relationship between the performance measure used and the principal's objective. The model provides a simple and intuitive statistical measure that serves as a metric for the efficiency of a performance measure. Applications to various incentive contracting situations, including the "gaming" of performance measures, the use of revenue-based sales commissions, and relative performance evaluation, are presented.

I. Introduction

Performance measurement, choosing which quantity or quantities to use in an incentive contract, is a central problem in agency theory. Holmström (1979) examined the use of performance measures in incentive contracts and established criteria for when a second-best contract based on the principal's objective can be improved by the inclusion of additional performance measures. However, Holm-

I would like to thank Robin Cooper, Robert Gibbons, Michael Gibbs, Bengt Holmström, Michael Jensen, Edward Lazear, Kevin J. Murphy, Birger Wernerfelt, Peter Wilson, Karen Wruck, Jerold Zimmerman, and participants in seminars at the University of Chicago Graduate School of Business, the Kennedy School of Government at Harvard, the Wharton School at the University of Pennsylvania, and the Simon School of Business at the University of Rochester for helpful comments. This research is supported by the Division of Research, Harvard Business School.

[*Journal of Political Economy*, 1992, vol. 100, no. 3]

ström's paper and most subsequent work start from the assumption that the principal's objective is always a contractible performance measure. In many organizational settings this assumption is not supportable. Many organizations (e.g., nonprofit firms or government agencies) lack a clear objective: incentive contracting in such an environment demands the use of other performance measures. Some firms lack traded residual claims, so that their objective (total value) is not a quantity that can be used in incentive contracts. In large publicly traded firms, total firm value may fluctuate so much that it is almost useless as a performance measure for risk-averse employees.

An organization's inability to use total value as the basis for incentive contracts often leads it to use a wide array of alternative performance measures: salespeople are paid commissions based on revenue, division managers receive bonuses based on divisional accounting profits, and many "merit pay" systems use attainment against a set of predetermined objectives as the basis for determining compensation.[1] Yet agency theory provides no systematic way to assess the "goodness" of any of these alternative performance measures. Under what circumstances do such contracts provide efficient outcomes, and what are the characteristics of an optimal contract based on such performance measures?

This paper examines optimal linear incentive contracts in which the principal's objective is not contractible. I find that the size of the optimal piece rate and the efficiency of the contract depend on the statistical relationship between the performance measure used and the principal's objective. Specifically, to the extent that the performance measure does not respond to the agent's actions in the same way that the principal's objective responds to these actions, the firm will reduce the sensitivity of the incentive contract to the performance measure, and the contract will not be first-best. This inefficiency occurs even if the agent is risk neutral. This result and the intuition behind it depend on an assumption that the agent is asymmetrically informed about how her actions affect outcomes. This information advantage, combined with the fact that the principal cannot contract on his true objective, forces the principal to use the performance measure to "tell the agent what to do." However, since the performance measure will not always give the agent accurate incentives, the agent will engage in actions that the principal, if he had the agent's information, would consider nonoptimal. It is these costly nonoptimal actions that lead to the inefficiency of the contract.

The model uses a quite general formulation of the performance

[1] The federal government implemented such a system for the Social Security Administration (among other agencies); see Pearce, Stevenson, and Perry (1985).

measurement problem, and the results shed light on a number of issues in incentive contracting. The model provides simple and intuitive explanations of when an employee can "game" a performance measure (i.e., take actions that increase payouts from the incentive contract without improving actual performance), when high-powered revenue-based commissions are preferred to salary-based sales compensation schemes, and when relative performance evaluation distorts an employee's incentives.

The paper proceeds as follows. In Section II, I present the information and production structures used in the model and discuss several important assumptions. In Section III, I solve for the optimal linear incentive contract under two different information regimes, one in which the agent's actions are unobservable and the other in which these actions are observable. In Section IV, I present several examples that illustrate the applicability of the model and discuss its implications.

II. Performance Measurement and Asymmetric Information

The central assumption of this model is that the principal's objective is not used in the incentive contract for the agent. This assumption is consistent with traditional economic theories of the firm and with the actual constraints faced by many organizations. In the traditional model of the firm, an owner-entrepreneur manages the enterprise. This person is assumed to have a well-specified notion of what he is trying to maximize. But how should this owner-entrepreneur contract with his subordinates, given that there is no observable quantity that captures his objective? Jensen and Meckling (1976) model such a firm; they explicitly assume that the owner-entrepreneur maximizes his utility at the expense of the financial value of the firm. The entrepreneur's objective, utility, cannot be contracted on, so if he is to write incentive contracts with his subordinates, he must use other performance measures. This same problem plagues many organizations: government agencies, nonprofit organizations, private companies, and professional partnerships. Indeed, with the exception of the publicly traded firm, few organizations have sufficiently clear or verifiable objectives that can be used directly in the incentive contracts of employees.

In the case of the publicly traded firm, the organization's objective, total firm value, is observable and contractible. However, from the perspective of most employees, total firm value contains so much random variation that it provides little information about employee actions. If employees are risk averse or wealth constrained, an optimal incentive contract will place a very small weight on total firm value. For this reason, if such firms are to provide incentives, they too must use other performance measures.

I model this problem by assuming that the principal's objective, $V(e, \epsilon)$, is not contractible. It is a function of e, the actions of the agent, and ϵ, a vector of random variables that completely characterizes the state of the world. I then assume the existence of an arbitrary, contractible performance measure, $P(e, \epsilon)$, which is also a function of the agent's actions and the state of the world. The principal uses this performance measure in a linear incentive contract of the form[2]

$$\text{agent's payoff} = S + bP(e, \epsilon),$$

where S is the fixed component of the agent's compensation, and b represents the "piece rate" paid for each unit of the performance measure delivered. The use of a linear incentive contract has several advantages. One is that it greatly simplifies the solution to the model. In addition, it simplifies the interpretation of the results and allows for straightforward comparative static analysis, since the strength of the incentives induced by the optimal contract is specified by a single parameter.

A second important set of assumptions in this model relates to the information structure. I assume that the agent is asymmetrically well informed about the state of the world and that the agent's superior knowledge affects her optimal action choice. Neither the principal nor the agent knows ϵ before signing the contract, but the realization of ϵ is known to the agent before she chooses her actions. The contract is binding: neither the principal nor the agent can renege on the contract once it is signed.[3] I also assume that at least some components of ϵ affect the marginal product of the agent's actions on both the performance measure and the value function. This set of assumptions has important implications for the intuition and interpretation of the model.

That the marginal products of the agent's actions on value and on the performance measure are functions of ϵ means that, from the perspective of the principal, these marginal products can be thought of as random variables. Since the agent's action choice depends on the marginal product of actions on the performance measure (P_e),

[2] As will be seen below, I assume that a minor adjustment is made to the performance measure (see Sec. III).

[3] The assumption that the agent cannot quit after information is revealed is justified on the grounds that the information revelation and the taking of the action are intimately related. The information asymmetry in the model is assumed to derive from the fact that the agent is the one actually there doing the job. Her action choice can thus be thought of as occurring almost simultaneously with the information revelation.

effort is also a random variable. Thus the principal does not know what the agent is going to do when faced with any particular incentive contract; nor would he know whether the agent's actions were optimal if he could observe them.[4] Forcing contracts, in which the principal tells the agent what actions to take, are not efficient under this set of assumptions (Harris and Raviv 1979).

The degrees to which the two marginal products (V_e and P_e) vary with the state of the world (ϵ) are important parameters. The standard deviation of V_e with respect to ϵ, σ_{v_e}, is a measure of the amount of valuable information that the agent possesses. When σ_{v_e} is low, the marginal product of the agent's actions on value does not vary very much in different states of the world. This means that the agent has little information that the principal does not have that would allow her to choose more valuable actions. Alternatively, when σ_{v_e} is high, the agent is able to alter her action choice significantly in response to her asymmetric information and thus produce more valuable outcomes.[5] Similarly, σ_{p_e} is a measure of the amount of valuable information the agent has about the performance measure.

III. Derivation of the Optimal Contract

The paper models the problem of designing an incentive contract for one agent. As noted above, the production function has the form, value $= V(e, \epsilon)$, where e is effort and ϵ is a vector of random variables that characterizes the state of the world. Value is interpreted as the value of all output minus the costs of all factors of production except for the compensation of the agent.

Effort is modeled as a unidimensional variable chosen by the agent. However, it is informative and intuitive to think about e as being a vector of tasks. Under such a specification, the agent chooses not simply how hard to work but also how to allocate her effort across tasks.[6] All the results of this paper are robust to such a change in

[4] I assume that there is sufficient variability in those components of ϵ that do not affect the marginal product of effort that the principal cannot infer this marginal product from observing effort and value. A simple example may simplify the intuition on this point. Let $\epsilon = \{\mu, \gamma\}$. Assume that $V = \mu e + \gamma$. If γ has high variance, then observing V and e gives little information about the size of μ and does not help in determining the first-best level of effort.

[5] The idea that the amount of variation in an information channel is a measure of the amount of information carried by the channel is not new: Shannon pointed this out in 1949. Indeed, if the signal from an information channel follows a normal distribution, then entropy (N), Shannon's measure of information content, is equal to $N = \frac{1}{2}\log(2\pi e\sigma^2)$, where e is the base of natural logarithms, and σ^2 is the variance of the signal's distribution.

[6] Holmström and Milgrom (1990) model this situation in detail, deriving a result similar to the one in this paper that piece rates may be distorted (and incentive contracts

specification, but the analysis is more cumbersome and the results are somewhat more difficult to interpret. In the interest of parsimony, I model the single-task case only. Effort may be observable or unobservable; solutions under both assumptions will be derived.

The agent signs a binding employment contract with the principal that specifies a base salary, a piece rate, and a performance measure on which the piece rate will be based. At the time of signing, neither the principal nor the agent has information about the realization of $\boldsymbol{\epsilon}$; they share common knowledge about its distribution. After the signing, but before the choice of an effort level, all components of $\boldsymbol{\epsilon}$ are revealed to the agent.[7]

The performance measure, P, is modeled as a function of effort and the state of the world. The incentive contract pays a base salary plus a piece rate, b, per unit of the performance measure. Because the piece rate is paid for something other than what the principal cares about, the issue of how to *scale* the performance measure arises. Let $\hat{P}$ be an arbitrary, unscaled performance measure. To interpret the magnitude of b, it is useful to scale $\hat{P}$ so that the piece rate can be compared to one. Such a scaling involves multiplying the performance measure by the average dollar value of an incremental unit of this performance measure.[8] In this way, the average value of an additional unit of the scaled performance measure is \$1.00, and the first-best contract involves a piece rate of one. Consider the following example. Instead of using the number of tomatoes picked as the performance measure for an agricultural worker, use the number of tomatoes picked times the average value (to the principal) of a tomato. Then the first-best contract for a risk-neutral picker will not set b equal to some number of dollars per tomato, but will set b equal to one.

Such a transformation requires multiplying the unscaled performance measure by a constant:

$$P(e, \boldsymbol{\epsilon}) \equiv \frac{E[V_e(e, \boldsymbol{\epsilon})]}{E[\hat{P}_e(e, \boldsymbol{\epsilon})]} \times \hat{P}(e, \boldsymbol{\epsilon}),$$

inefficient), even when the agent is risk neutral. They note that the state-contingent action model used in my paper is equivalent to their vector effort model. They assume that the principal is unable to determine the precise value of each of the agent's actions, whereas my paper assumes that the principal lacks information about the value of effort in different states. Their result also relies on the fact that different types of effort are substitutes or complements in the agent's utility function. I require only that the disutility of effort function be convex.

[7] Since the agent gets perfect information about the state of the world before making her effort choice, using a menu of contracts is no more efficient than using a single contract (see Melumad and Reichelstein 1989).

[8] This average is taken across the possible states of the world.

where $\hat{P}_e$ is the marginal product of effort on the unscaled performance measure, V_e is the marginal product of effort on value, and $E[\cdot]$ is the expectation operator, taken over ϵ.

This transformation implies that the following condition holds:

$$E[P_e(e, \epsilon)] = E[V_e(e, \epsilon)]. \quad (1)$$

Equation (1) states that the performance measure is normalized so that its expected marginal product of effort equals the expected marginal product of effort on value. This normalization is a natural one and is innocuous for the theory developed in this paper.

Optimal Incentive Contract with Unobservable Effort

The model is solved in a standard way, with the principal maximizing profits (value net of compensation payments) subject to two constraints: a participation constraint and an incentive constraint. The agent is assumed to be risk neutral, so that her utility function takes the form

$$\text{utility} = S + bP - C(e),$$

where $C(e)$ is the disutility of effort, $C' > 0$, $C'' > 0$.

The agent has alternative opportunities that give her utility H and chooses effort so as to maximize utility. Since the agent signs a binding contract before knowing ϵ, but chooses effort after ϵ is revealed, the participation and incentive constraints are given by, respectively,

$$H \leq E[S + bP - C(e)] \quad (2)$$

and

$$bP_e(e^*, \epsilon) = C'(e^*), \quad (3)$$

where e^*, the agent's choice of effort, is a function of both b and ϵ.[9]

The principal's maximization problem is thus

$$\max_{b,S} E[V(e^*, \epsilon) - S - bP(e^*, \epsilon)], \quad (4)$$

subject to equations (2) and (3).

The solution to this program yields the following expression for the optimal piece rate:

$$b^* = \frac{E[V_e e_b^*]}{E[P_e e_b^*]}. \quad (5)$$

Note that if the optimal effort choice is not a function of ϵ, then e_b^*

[9] I assume that the second-order conditions are met. This implies that $C'' - bP_{ee} > 0$.

drops out of equation (5) and $b^* = E[V_e]/E[P_e] = 1$. This is the standard result in a principal-agent model with a risk-neutral agent.

To explore the characteristics of b^* when e^* is a function of ϵ, it is necessary to expand e_b^* more fully. Differentiating equation (3) with respect to b gives

$$e_b^* = \frac{P_e}{C'' - bP_{ee}}. \tag{6}$$

Substituting equation (6) into (5), using second-order Taylor approximations for C and P, yields[10]

$$b^* = \frac{E[V_e P_e]}{E[P_e^2]}. \tag{7}$$

Assume, without loss of generality, that $E[P_e] = E[V_e] = 1$ at e^*. Equation (7) can be rewritten as

$$b^* = \frac{\operatorname{cov}(V_e, P_e) + 1}{\operatorname{var}(P_e) + 1} = \frac{\rho\sigma_{v_e}\sigma_{p_e} + 1}{\sigma_{p_e}^2 + 1}, \tag{8}$$

where ρ is the correlation between P_e and V_e, σ_{v_e} is the standard deviation of V_e, and σ_{p_e} is the standard deviation of P_e.

Several important aspects of the performance measurement problem now become apparent. Note first that if P_e and V_e have the same variance and are perfectly correlated, then $b^* = 1$. Of course, these two conditions combined with equation (1) imply that $P_e = V_e$ in all states of the world; this is a necessary and sufficient condition for a first-best contract.[11] Note that this does not imply that V and P are the same. In particular, value could be much more variable than the performance measure and could be influenced by many factors that do not affect P, as long as these factors do not affect the optimal actions of the agent. More generally, any performance measure that responds to an employee's actions in exactly the same way that value responds to these actions can be used to write a first-best incentive contract for a risk-neutral agent.

Equation (8) demonstrates that the *correlation* between V_e and P_e is important to the determination of the optimal piece rate. All else equal, the higher this correlation, the higher the optimal piece rate. If the marginal product of effort on a performance measure is strongly

[10] This approximation is equivalent to assuming that the second derivatives of C and P with respect to e and ϵ are constant.

[11] The proof of sufficiency is trivial. The proof of necessity proceeds as follows. The condition for first-best outcomes is that $C' = V_e$ for all ϵ. Equation (3) implies that $C' = b^*P_e$ for all ϵ. Thus for first-best outcomes to result, $V_e = b^*P_e$ for all ϵ. However, by eq. (1), $E[P_e] = E[V_e]$. Thus $b^* = 1$, which implies that $V_e = P_e$ for all ϵ.

correlated with the marginal product of effort on value, then the agent, who chooses an effort level based on the value of P_e, will choose high levels of effort when V_e is high and low levels when V_e is low. If these marginal products are not strongly correlated, then the agent's effort choice will not match the principal's desired effort level in most states. Because the agent's disutility of effort function is convex, choosing the wrong level of effort is costly. In response to this cost, the principal reduces the piece rate and reduces incentives.

It is informative to compare this result with the standard principal-agent result. In the standard model, the principal and agent have a conflict of interest over variations in outcomes: when the agent is risk averse, the principal chooses to reduce the piece rate to minimize the agent's exposure to this variation. If the agent is risk neutral in the standard model, there is no conflict of interest between the principal and the agent, and the contract yields a first-best set of outcomes. In the model in this paper, the agent is not averse to variations in income but is averse to variations in effort because of the convexity of the disutility of effort function. Since P_e may not equal V_e in every state of the world, the contract based on P gives the agent inaccurate information about how hard to work, and the agent suffers a loss of utility from this effort variation. The result is that the principal chooses to reduce the effort variation by reducing the piece rate, with an attendant cost in terms of incentives for effort.

Certain aspects of the optimal piece rate in this model seem counterintuitive. In particular, the piece rate may still be positive even when the correlation between P_e and V_e is negative. This arises because the piece rate in this model must perform two functions: one is to give the agent an incentive to exert positive effort, and the other is to get her to use her superior information in choosing her effort level. Because the marginal product of effort on value is, on average, positive, the principal wants the agent to exert effort in most states of the world.[12] Ideally, the principal would like to condition the size of the piece rate on the expected value of P given the agent's effort level and then adjust the contingent payment on the basis of the unexpected value of P. However, a contract of this sort is possible only if effort is observable.

[12] A simple example may help the intuition on this point. Suppose that there are two equally likely states of the world. In one, the marginal product of effort on the performance measure is five and that on value is 10. In the other, the marginal product of effort on P is 10 and that on V is five. The marginal products are perfectly negatively correlated, but the principal still wants the agent to exert positive effort: on average, the marginal products of both P and V are 7.5. In this example, the optimal piece rate, calculated from eq. (7), is .8.

Observable Effort Case

When effort is observable, the principal can condition the agent's payoff on her effort choice. Recall that with the information asymmetry assumed in this model, being able to observe effort does not imply that the principal can write a first-best forcing contract. However, being able to observe effort does expand the set of possible contracts that the principal can use. One contract that takes advantage of this ability is

$$\text{agent's payoff} = S + b_1 E[P|e] + b_2(P - E[P|e]).^{13}$$

This contract pays the agent two separate piece rates, one for the "expected" level of the performance measure given her level of effort and another for the difference between this expected level and the actual value realized by the performance measure. Even though the principal can observe effort, he does not tell the agent what level of effort to choose. Rather, he influences the agent's choice of effort by giving her incentives (through b_1) to exert the right level of effort on average and then providing incentives (through b_2) to adjust this effort level on the basis of her superior information.

As above, the agent signs a binding contract specifying S, b_1, and b_2 before receiving information about ϵ. Solving for the optimal contract yields the following piece rates:

$$b_1^* = \frac{E[V_e]}{E[P_e]} = 1 \tag{9}$$

and

$$b_2^* = \rho \frac{\sigma_{ve}}{\sigma_{pe}}, \tag{10}$$

where, as before, ρ is the correlation between V_e and P_e, σ_{v_e} is the standard deviation of V_e, and σ_{p_e} is the standard deviation of P_e.[14]

This contract is first-best when $\rho = 1$ (see the Appendix). Note that b_1 is essentially a piece rate on effort, scaled to reflect the marginal product of effort on value. Thus *effort* in this case serves as an addi-

[13] The form of this contract is chosen to simplify the interpretation of b_1 and b_2.

[14] Note that eq. (10) is the equation for a regression coefficient with intercept, whereas eq. (8) (without the normalization) is the equation for a regression line without intercept. This problem can be interpreted as one in which the principal chooses a linear relationship between V_e and P_e in order to maximize value net of effort costs, which is quadratic because of the Taylor approximation before eq. (8). In the unobservable effort case, there is no intercept because the principal does not know the expected value of P given effort. In the observable effort case, he does; therefore, he can "mean adjust" the regression line.

tional *performance measure,* but one whose use does not exploit the agent's information advantage.

The two-piece-rate contract is Pareto superior to a forcing contract whenever $\rho \neq 0$. The reason is that when $\rho \neq 0$, P_e contains some useful information about the state of the world and V_e. In the forcing contract, the principal would tell the agent to set effort such that $C'(e) = E[V_e]$. Note that this level of effort is the same as the effort that would be chosen by the agent if b_2 were zero. However, because b_2 will not be zero if $\rho \neq 0$, the agent has incentives to alter her effort choice to take advantage of the additional knowledge that she has about the state of the world. In particular, she will work harder when V_e is higher than expected and less hard when V_e is lower than expected. This contract will thus produce higher profits for the principal than a forcing contract.

IV. Applications and Discussion

This model illuminates a number of interesting situations in incentive contracting that have not been easily analyzed with existing models. In this section I discuss three: employee "gaming" of performance measures, the use of revenue-based sales commissions, and relative performance evaluation.

Gaming a Performance Measure

The behavioral literature on compensation systems (e.g., Lawler 1971, 1990; Hamner 1975; Beer et al. 1984) argues that using incentive pay often leads to unintentional and dysfunctional consequences. Lawler (1990, p. 58) relates that

> the literature on incentive plans is full of vivid descriptions of the counterproductive behaviors that . . . incentive plans produce. One of the first books I read in compensation provided story after story about how employees were outsmarting and defeating piece-rate systems (Whyte, 1955). Indeed, as I read this classic book, I marveled at the ingenuity of the worker. . . . It was clear that the systems were motivating behavior—but unfortunately they were motivating the *wrong* behavior.

According to these writers, employees' ability to game performance measures leads firms to avoid the use of incentive pay altogether. Milgrom (1988) makes a similar argument, showing that employees' ability to engage in "influence activities" reduces the efficiency of incentive contracts. The present model provides additional insight into why basing "high-powered" incentives on performance measures may be inefficient, even when the performance measure on average gives incentives to do the right thing.

Consider a firm trying to design an incentive contract for a research and development scientist. The firm cannot observe effort, and the true value of what the scientist produces is not known and cannot be contracted on. The firm knows only that one unit of the scientist's effort is worth \$1.00. In an attempt to give the scientist incentives to work hard, the firm writes a piece-rate incentive contract that is based on the number of patents granted to the scientist. The firm knows that, on average, it takes $1/P_e$ effort units to produce a patent and that the average patent is worth \$$X$ to the firm.

Suppose that patents vary in their difficulty to produce. Some can be produced for less than $1/P_e$ units of the scientist's effort, and some require more effort. The scientist knows, but the firm does not, how difficult it will be to produce a patent on a particular project. Reference to equation (6) above shows that the firm will not use a piece rate that pays the scientist \$$X$ per patent, even though this is the average value of each patent. The optimal piece rate on patents in this example equals

$$b^* = X\left(\frac{\sigma_{pe}^2}{X^2} + 1\right)^{-1}.$$

Note that the optimal piece rate is biased downward by an amount equal to one over one plus the scaled variance of P_e. The reason is that the scientist is able to game the number of patents produced, using his superior knowledge to work too hard on the easy ones and not hard enough on the hard ones. To reduce the amount of gaming that the scientist does, the firm reduces the piece rate.

Performance Measurement "Difficulty" and Revenue-based Commissions

Alchian and Demsetz (1972) identified the cost of performance measurement (they called it "metering") as a key determinant of whether an economic system would use decentralized incentives or centralized monitoring and fiat as a way to secure effort. Within organizations, the choice of whether to use incentive contracts is often said to depend on how difficult it is to measure performance. In an interesting test of this proposition, Anderson and Schmittlein (1984) find that managers' perceived difficulty of measuring sales performance was the most important determinant of whether a sample of electronics firms used independent manufacturers' representatives (whose compensation is a multiple of sales volume) or their own captive sales force (whose compensation is a combination of salary and commission). Careful consideration of this argument, however, reveals its imprecision; it is always possible (and cheap) to measure performance in some way. In sales, for instance, revenue is a cheap and easy performance measure. The question is not whether performance is easy to measure, but rather whether the available performance measure (in this case revenue) accurately reflects the firm's objective and is thus a good measure.

Consider the following model of the sales process. Firm A sells a single standard product, one whose parameters are simple to specify and whose costs do not vary with the specifics of any one sale. Salespeople have no discretion over the price charged but can exert effort to sell a greater sales volume, Q. The sales force is asymmetrically informed about the details of how their effort affects sales volume. Sales volume and the unit cost of the product are affected by independent random shocks (represented by $\boldsymbol{\epsilon}$ and $\boldsymbol{\mu}$, respectively), but unit costs are unaffected by the efforts of the salesperson. Firm A's incremental value from a salesperson, before compensation costs, is

$$V(e, \boldsymbol{\epsilon}, \boldsymbol{\mu}) = Q(e, \boldsymbol{\epsilon})[\text{unit price} - \text{unit cost}(\boldsymbol{\mu})].$$

Firm A, in an effort to insulate the sales force from the randomness in unit costs, uses only revenue, $Q \times$ (unit price), as the performance measure for its sales force. Using revenue as a performance measure in this case is nondistorting: after revenue is scaled so that it satisfies equation (1), the marginal product of the salesperson's effort on revenue is identical to the marginal product of effort on value. Thus, under risk neutrality, the piece rate is one and the contract yields first-best outcomes.

Firm B sells a more complex product, whose specifications are not standard. In this firm, the efforts of the salesperson affect unit costs in addition to the number of units sold. Once again, the salesperson is asymmetrically informed about how this effort affects volume and unit cost. In this firm, the incremental value from a salesperson (before compensation) is

$$V(e, \boldsymbol{\epsilon}, \boldsymbol{\mu}) = Q(e, \boldsymbol{\epsilon})[\text{unit price} - \text{unit cost}(e, \boldsymbol{\mu})].$$

Although this might be referred to as a situation in which "performance is harder to measure," in fact exactly the same performance measure (revenue) is available. However, this performance measure does not have the same properties that it had for firm A. In particular, the marginal product of effort on revenue is imperfectly correlated with the marginal product of effort on value. This leads to distorted incentives and a contract that does not induce first-best outcomes, even under risk neutrality. Only by basing the incentive contract on value and forcing the sales force to bear all the randomness in product costs can firm B's contract yield undistorted incentives.

Relative Performance Evaluation

The insight on performance measurement developed in this paper encompasses and refines the analysis of the use of relative performance evaluation for employees. Lazear (1989) argues that the possibility that employees will sabotage the output of others leads firms to use less relative performance evaluation than they otherwise would. Gibbons and Murphy (1990, pp. 34–35) argue that "paying workers based on relative performance, instead of absolute performance, distorts the worker's incentives whenever the worker can take actions that affect the average output of the reference group." The results developed above allow an examination of exactly when relative performance evaluation distorts incentives and when it does not.

Define a worker's absolute performance as V and the average output of the reference group against whom he will be compared as R. Then P, the relative performance measure, is

$$P(e, \boldsymbol{\epsilon}) \equiv V(e, \boldsymbol{\epsilon}) - R(e, \boldsymbol{\epsilon}).$$

When will such a relative performance measure provide incentives that are nondistorting? In the case of unobservable effort, P provides nondistorting incentives only when $P_e = V_e$ in all states of the world. This condition will be satisfied only when $R_e \equiv 0$, that is, when the employee's effort has no effect on the reference group. This is the condition that Gibbons and Murphy argue is required for a nondistorting relative performance measure. If the effect of R is to reduce the variance of P relative to the variance of V and R_e is zero, then it will allow the firm to write a more efficient contract with a wealth-constrained or risk-averse employee. Note that if R_e is small but not zero and R can substantially reduce the variance of P relative to that of V, then it might be worth using the relative performance measure and suffering a small amount of incentive distortion in exchange for the risk-reducing benefits that come from using the less variable performance measure in the incentive contract.

In the case of observable effort, a relative performance measure can be nondistorting even if the employee can take actions to affect the reference group. If P_e and V_e are perfectly correlated, then an incentive contract based on P induces first-best effort. This correlation is perfect when either the *variance* of R_e is zero or R_e is perfectly correlated with V_e. Thus if the employee's actions affect the reference group identically in all states of the world or if they affect its output in exact proportion to how they affect individual output, then a relative performance evaluation will provide distortion-free incentives.

V. Conclusions: Optimal Compensation Systems

This paper derives the optimal linear incentive contract under the assumption that a principal cannot use his own objective in the contract. The result is a contract that distorts and dampens the agent's incentives for effort, even when the agent is risk neutral. The model provides an intuitive and simple metric for the "goodness" of a performance measure: when the marginal product of the agent's actions on the performance measure is highly correlated with the marginal product of these actions on the principal's objective, then the performance measure is a good one and the resulting contract will be efficient. If not, the resulting contract will induce outcomes that significantly diverge from the first-best.

The results of the model depend crucially on the assumption that the agent is asymmetrically informed with knowledge that is useful in determining optimal actions. This assumption is a valid description of the situation in many agency relationships; indeed it is a primary reason for the existence of many of these relationships. When such an information asymmetry exists, the performance measurement problem becomes one of trying to induce the agent to use her information productively, while at the same time avoiding incentives to engage in dysfunctional actions. Many organizationally relevant issues in incentive contracting can be understood by focusing on this tradeoff. For instance, bonus- and commission-based compensation systems will tend to dominate when the agent possesses valuable information and good performance measures are available. Straight salary compensation systems will tend to dominate either when the agent is not asymmetrically informed about the job being performed or when no good performance measures exist. In these situations, expending resources to monitor effort or to mitigate the information asymmetry will be efficient. Incentive contracts based on the total value of the organization, such as partnerships and stock ownership, will dominate when information asymmetries are great and no good performance measures exist.

Appendix

This Appendix proves the proposition that the two-piece-rate contract with observable effort is first-best whenever V_e and P_e are perfectly correlated. Perfect correlation implies that V_e is a linear function of P_e:

$$V_e = \alpha + \beta P_e. \qquad \text{(A1)}$$

This implies that $\sigma_{v_e} = \beta\sigma_{p_e}$, which implies

$$\beta = \frac{\sigma_{v_e}}{\sigma_{p_e}}. \tag{A2}$$

Because $E[V_e] = E[P_e]$ by equation (1),

$$\alpha = E[P_e]\left(1 - \frac{\sigma_{v_e}}{\sigma_{p_e}}\right). \tag{A3}$$

Combining (A1), (A2), and (A3) yields

$$V_e = E[P_e]\left(1 - \frac{\sigma_{v_e}}{\sigma_{p_e}}\right) + \left(\frac{\sigma_{v_e}}{\sigma_{p_e}}\right)P_e. \tag{A4}$$

First-best outcomes are achieved when $C' = V_e$. From equations (9) and (10),

$$b_1^* = 1 \tag{A5}$$

and

$$b_2^* = \frac{\sigma_{v_e}}{\sigma_{p_e}}. \tag{A6}$$

The first-order conditions for the agent's utility maximization imply

$$b_1^*E[P_e] + b_2^*(P_e - E[P_e]) = C'. \tag{A7}$$

Substituting (A5) and (A6) into (A7), rearranging terms, yields

$$C' = E[P_e]\left(1 - \frac{\sigma_{v_e}}{\sigma_{p_e}}\right) + \left(\frac{\sigma_{v_e}}{\sigma_{p_e}}\right)P_e,$$

which equals the right-hand side of equation (A4).

References

Alchian, Armen A., and Demsetz, Harold. "Production, Information Costs, and Economic Organization." *A.E.R.* 62 (December 1972): 777–95.

Anderson, Erin, and Schmittlein, David C. "Integration of the Sales Force: An Empirical Examination." *Rand J. Econ.* 15 (Autumn 1984): 385–95.

Beer, Michael; Spector, Bert; Lawrence, Paul R.; Mills, D. Quinn; and Walton, Richard E. *Managing Human Assets.* New York: Free Press, 1984.

Gibbons, Robert, and Murphy, Kevin J. "Relative Performance Evaluation for Chief Executive Officers." *Indus. and Labor Relations Rev.* 43 (suppl; February 1990): 30S–51S.

Hamner, W. Clay. "How to Ruin Motivation with Pay." *Compensation Rev.* 7, no. 3 (1975): 17–27.

Harris, Milton, and Raviv, Artur. "Optimal Incentive Contracts with Imperfect Information." *J. Econ. Theory* 20 (April 1979): 231–59.

Holmström, Bengt. "Moral Hazard and Observability." *Bell J. Econ.* 10 (Spring 1979): 74–91.

Holmström, Bengt, and Milgrom, Paul R. "Multi-Task Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design." Working Paper ser. D, no. 45. New Haven, Conn.: Yale School Organization and Management, May 1990.

Jensen, Michael C., and Meckling, William H. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *J. Financial Econ.* 3 (October 1976): 305–60.

Lawler, Edward E., III. *Pay and Organizational Effectiveness: A Psychological View.* New York: McGraw-Hill, 1971.

———. *Strategic Pay: Aligning Organizational Strategies and Pay Systems.* San Francisco: Jossey-Bass, 1990.

Lazear, Edward P. "Pay Equality and Industrial Politics." *J.P.E.* 97 (June 1989): 561–80.

Melumad, Nahum D., and Reichelstein, Stefan. "Value of Communication in Agencies." *J. Econ. Theory* 47 (April 1989): 334–68.

Milgrom, Paul R. "Employment Contracts, Influence Activities, and Efficient Organization Design." *J.P.E.* 96 (February 1988): 42–60.

Pearce, Jone L.; Stevenson, William B.; and Perry, James L. "Managerial Compensation Based on Organizational Performance: A Time Series Analysis of the Effects of Merit Pay." *Acad. Management J.* 28 (June 1985): 261–78.

Shannon, Claude E. *The Mathematical Theory of Communication.* Urbana: Univ. Illinois Press, 1949.