

All-or-Nothing Monitoring

Rui R. Zhao*

October 2007
forthcoming in *AER*

*Department of Economics, University at Albany - SUNY, BA 110, 1400 Washington Avenue, Albany, NY 12222 (e-mail: rzhao@albany.edu). I am grateful to four anonymous referees as well as the editor for very helpful suggestions, and to Mike Jerison for helpful comments. Any remaining errors are solely mine.

Abstract

A principal can observe both the output and input of an agent who works at a job involving multiple tasks. We provide a simple theory that explains why it may be optimal for the principal to use only output-based incentive contract, even though the principal can monitor the agent's actions perfectly in all but one task and knows exactly which action is optimal for each task. (JEL: D82, M5, J33)

Keywords: Incentive Theory, Output-based Pay, Monitoring, Performance Measure.

A main issue in the design of incentive contracts concerns the choice of performance measures. In practice, the job of a worker typically involves several different aspects or tasks that all contribute to a final outcome valued by the principal; often both the final outcome and the agent's actions in some tasks are observable to the principal. Thus a basic choice the principal faces when designing incentive contracts is whether to reward the agent for output or for effort. In this paper we show that it can be optimal for the principal to incentivize an agent using only output-based performance pay even if the principal can monitor some of the agent's actions perfectly and knows exactly which action is optimal for each task, as long as there is one action of the agent she can not observe.

Specifically, we address the following incentive problem. A principal hires an agent for a job that involves multiple tasks. The agent faces limited liability or is wealth constrained, so to motivate effort the principal either has to pay the agent bonuses for good outcome or must monitor the agent's actions. The principal can costlessly monitor and induce the agent to take the desired actions in a subset of the tasks. In addition, there are verifiable noisy measures of outputs, which may be task specific or at the aggregate level. Then unless the principal can observe all actions of the agent she is better off using only the noisy output measures rather than combining input measures for the monitored tasks and output measures for the unmonitored tasks. Hence the principal's optimal strategy for input monitoring is *all or nothing: if you can't monitor everything, monitor nothing*.¹

What drives this result is an information synergy from using joint output-based performance measures in designing incentive contracts. Specifically, by rewarding good performance on multiple tasks, a single bonus can be used to simultaneously motivate efforts in all tasks. This efficient use of reward saves the rent the principal has to pay the agent compared with contracting on individual tasks separately. The benefit is larger when more tasks are included in the output-based contract, so much so that the principal's overall cost is lower if the monitored tasks are also included in the *output*-based contract even though the principal does not have to pay any rent by monitoring the agent's actions on those tasks. This result therefore provides a simple explanation for the use of output-based incentive contracts.

The choice between input or output based performance measures has been studied before. We want to stress that the reason offered in this paper in favor of output-based measures is not

¹Throughout the paper, by *monitoring* a task we mean the monitoring of *actions* in that task.

dependent on private information and self-selection, the reasons emphasized by most other authors. For instance, George Baker (1992) argues that often the measured outcomes are only partially related to the principal's latent objective and the agent may be asymmetrically informed about the effects of his actions; as a result the principal may choose to leave the decisions to the agent and incentivize him with an output-based contract. Edward P. Lazear (2000) points out that output-based pay may be superior to hourly wage for the purpose of selecting workers with higher abilities in a heterogeneous workforce. Canice Prendergast (2002) shows that when the environment is more uncertain, a principal may have little idea what the right kinds of actions are and hence is more likely to delegate responsibility to workers but, to constrain their discretion, base compensation on observed output. Michael Raith (2005) argues that the principal may want to use both input and output performance measures in order to encourage the agent to properly utilize his specific knowledge about the consequences of his actions.

All these studies suppose that the agent has some private information about the effects of his actions on the principal's objective; contracting on output therefore is a way to elicit the agent's information. These are no doubt important factors to consider when designing incentive contracts.² The present paper however shows that there may be another, quite basic reason for output-based pay due to the informational gain from using joint output measures in incentive contracts when the job involves multiple tasks and the agent faces limited liability. In the present model the agent has no pre-contractual private information and the principal knows perfectly which actions are optimal, so self selection plays no role in contract design.

Linking the agent's pay to his overall performances across tasks is related to an idea in the repeated game literature, first pointed out by Dilip Abreu, Paul Milgrom and David G. Pearce (1991). They show that incentives in repeated partnerships may be provided more efficiently by pooling imperfect information and by "reusing" a single punishment for bad performances across multiple periods; their idea has also been applied to repeated agency models of William Fuchs (2007) and Jonathan Levin (2003). The difference is that here a bonus is used to reward good performances across tasks and direct monitoring of the agent's actions is studied. Section I.B discusses further the different implications of rewards and punishments.

²Fahad Khalil and Jacques Lawarrée (1995) consider the problem from a different angle; assuming that either input or output but not both can be monitored at the same time, they show that the choice of performance measures may depend on who the residual claimant is.

In their seminal study of multitask principal-agent problems, Bengt Holmström and Milgrom (1991) introduce the effort substitution problem. Holmström and Milgrom obtain many insights but an important finding is that if the performance measure on some task becomes less informative then the principal may want to make the agent's pay less sensitive to performance measures on *other* tasks. This is to prevent the agent from allocating too much effort in the well-measured tasks, which may be detrimental to the principal.

The question addressed in this paper is different from that in Holmström and Milgrom (1991). In the current model efforts are task specific so there is no effort substitution problem. More importantly, we do not compare precisions of performance measures across different tasks but rather compare different measures, based on either output or input, on the *same* tasks. In the present model, two performance measures are available for some tasks and the principal optimally chooses the *less* informative one.³

Glenn MacDonald and Leslie M. Marx (2001) study a multitask principal-agent problem in which efforts in different tasks are substitutes for the agent but are complements for the principal; the optimal contract in their model also involves a large bonus paid when performances are good on multiple tasks. But different from the present paper, MacDonald and Marx assume that the agent has private information about his effort costs across tasks and only output measures are available, so they do not examine the effect of direct monitoring of the agents' actions by the principal.

The present analysis is also related to efficiency wage theory (Carl Shapiro and Joseph Stiglitz, 1984), which has the implication that bonus and monitoring are substitutable means for implementing a given effort level (Erica Groshen and Alan Krueger, 1990). In contrast, the present paper considers a multitask environment and shows that a bonus is superior even if monitoring is perfect for all but one task.

³The result therefore offers an example to which the informativeness criterion of Holmström (1979) does not apply.

I The Main Result

A The Model

A principal hires an agent for a job that involves $n > 1$ tasks. All the tasks are symmetric.⁴ For each task $i = 1, \dots, n$, the agent has two effort levels $E = \{0, 1\}$, which for instance represent *shirk* and *work* respectively.⁵ Shirking costs the agent nothing, while working costs $c > 0$. A finite set of at least two verifiable output signals, Y , may be observed for each task. The tasks are independent: effort in one task does not affect performances in the other tasks. For each task i , given effort choice $e_i \in E$, let $p(y|e_i)$ be the probability that output signal $y \in Y$ will occur. Assume $p(y|e_i) \in (0, 1)$ for all $y \in Y$ and all $e_i \in E$. Assume also that there is a signal $z \in Y$ such that

$$(1) \quad \frac{p(z|1)}{p(z|0)} > \frac{p(y|1)}{p(y|0)}, \quad \forall y \in Y \text{ with } y \neq z.$$

Namely, the likelihood ratio $p(y|1)/p(y|0)$ is maximized at signal $y = z$. This assumption is satisfied generically.⁶ For convenience, let $p = p(z|1)$ and $q = p(z|0)$. It follows that $0 < q < p < 1$.

The agent maximizes expected utility and has a Bernoulli utility function $u(w, e) = w - c \sum_{i=1}^n e_i$, where w is the wage payment received from the principal and $e = (e_1, \dots, e_n)$ is the vector of effort levels in all tasks. There is a minimum wage that the agent must be paid for taking the job; this is also the agent's reservation wage. We normalize it to zero. This minimum wage may be due to limited liability or a wealth constraint of the agent or is the result of some pre-existing contractual obligations.⁷

The principal can costlessly observe the agent's effort choices in some tasks but can not prove them to a court, because effort choices are technical and subjective and hence nonverifiable. Monitoring however does allow the principal to compensate the agent according to

⁴We discuss asymmetry below.

⁵It is useful to think of 0 as the effort level the agent would choose due to intrinsic motivations.

⁶In fact, we make this assumption purely for expositional purposes. Clearly, there always exists some z such that Eq. (1) holds with weak inequality, which is sufficient for our analysis. Alternatively, if several signals z_1, \dots, z_m maximize the likelihood ratio, we may treat them collectively as one "signal" z and the analysis then goes through without change.

⁷We stress that the results in the present paper refer to *incentive* pay not the full wage. Many workers are only paid fixed wages, based on among other things working hours and experience.

effort level: paying c for $e_i = 1$ and zero for $e_i = 0$. We assume that the principal will follow through on such promises for reputation reasons.⁸ Thus the principal can implement effort level 1 on a monitored task without paying the agent any rent.⁹

In addition, the principal can always write enforceable incentive contracts based on verifiable output measures. The principal's objective is to make the agent choose effort level 1 in all tasks, either through monitoring or using some incentive contract based on output measures.¹⁰

B Optimal Output-Based Incentive Contract

We first analyze the principal's optimal output-based incentive contract. Let Y^n be the n -fold Cartesian product of Y and denote by $\theta = (y_1, \dots, y_n) \in Y^n$ a vector of output signals in all n tasks. For each given vector of effort levels $e = (e_1, \dots, e_n)$, let $P(\theta|e) = \prod_{i=1}^n p(y_i|e_i)$ be the probability that output vector θ will occur.

In an output-based incentive contract, the principal rewards the agent according to a wage plan $w(\theta) \geq 0, \forall \theta \in Y^n$. The principal's objective is to choose a wage plan that minimizes expected wage payments while motivating the agent to choose effort level 1 in all tasks. Denote by $H = (1, \dots, 1)$ the effort vector that the principal would like to implement. Formally, the principal's problem is defined by the following program.

Program P:

$$\begin{aligned} \max_{w(\theta)} & - \sum_{\theta \in Y^n} P(\theta|H)w(\theta) \\ \text{s.t. } w(\theta) & \geq 0, \text{ for all } \theta \in Y^n \text{ and} \\ (2) \quad \sum_{\theta} P(\theta|H)w(\theta) - nc & \geq \sum_{\theta} P(\theta|e)w(\theta) - c \sum_{i=1}^n e_i, \quad \forall e \in E^n. \end{aligned}$$

Notice that the agent has many ways to shirk. The incentive constraint (2) says choosing effort level 1 in all tasks is better than any particular way of shirking. The analysis of Program

⁸Failure of such reputation mechanisms can only make monitoring less useful and hence strengthen the result of this paper.

⁹Nonverifiability of monitoring however rules out incentive contracts that link effort choice in task i to an output-based contract for task j . This is because the agent can shirk in task i but still demand the principal to pay according to the contract for j . The court can not verify that the agent did not perform in task i .

¹⁰If the principal would rather let the agent shirk on some tasks then the analysis applies to the rest of the tasks.

(P) is much simplified by relaxing this complicated incentive constraint. Specifically, we consider the following relaxed program in which only a particular shirking strategy of the agent is allowed: shirking in all n tasks. The idea will be that shirking in all tasks gives the agent the maximum benefit from shirking, so if the optimal wage plan can deter such a deviation then it will deter any other deviation by the agent. Let $L = (0, \dots, 0)$ be the vector of minimum effort levels. Then in the relaxed program, we replace the agent's incentive constraint (2) with the following:

$$(3) \quad \sum_{\theta} P(\theta|H)w(\theta) - nc \geq \sum_{\theta} P(\theta|L)w(\theta).$$

Lemma 1. *The solution to the relaxed program is the following:*

$$(4) \quad w(\theta) = \begin{cases} \frac{nc}{p^n - q^n} & : \text{if } \theta = \theta^* \\ 0 & : \text{if } \theta \neq \theta^* \end{cases}$$

where $\theta^* \equiv (z, \dots, z)$.

Therefore the optimal incentive contract is to reward the agent a positive bonus only when output is equal to z in all tasks. The reason for this particular form of bonus contract is as follows. Notice that output vector $\theta^* = (z, \dots, z)$ is most informative for detecting the deviation from H to L in terms of likelihood ratio test: $P(\theta|H)/P(\theta|L)$ attains its maximum value at $\theta = \theta^*$. Since payoffs are linear in wages, rewarding θ^* is most cost-effective for the principal.

Given bonus contract (4), will the agent find deviations other than L profitable? The answer, as shown in the following proposition, is negative. Hence, the solution to the relaxed program in fact also solves the original Program (P) .

Proposition 1. *If the principal uses an output-based bonus contract to motivate effort level 1 in all tasks then the uniquely optimal contract is given by (4).*

PROOF: Recall that for each task and for effort levels 1 and 0 the respective probabilities of output signal z occurring are equal to p and q , with $0 < q < p < 1$. Let $P(k) = p^k q^{n-k}$ be the probability that signal vector $\theta^* = (z, \dots, z)$ will occur when the agent exerts effort in k out of n tasks. The key for the argument lies in the following observation:

$$(5) \quad P(k+1) - P(k) > P(k) - P(k-1) > 0, \quad \forall k = 1, \dots, n-1,$$

which says that the marginal gain per unit of effort, in terms of raising the probability of outcome θ^* , increases as total effort level increases.

It then follows that for $k = 1, \dots, n-1$,

$$\frac{P(n) - P(k)}{n - k} > \frac{P(k) - P(0)}{k}$$

and hence

$$\begin{aligned} (6) \quad \frac{P(n) - P(0)}{n} &= \frac{1}{n}(P(n) - P(k) + P(k) - P(0)) \\ &< \frac{1}{n} \left(P(n) - P(k) + \frac{k}{n-k}(P(n) - P(k)) \right) \\ &< \frac{P(n) - P(k)}{n - k}. \end{aligned}$$

Since the agent is paid a positive bonus only when θ^* is realized and the probability of this event only depends on total effort level k not the allocation of efforts among different tasks, it follows that the agent's payoff only depends on his total effort level. In particular, given bonus contract (4), for total effort level $k = 0, \dots, n$, the agent's expected payoff is given as follows:

$$U(k) \equiv \frac{p^k q^{n-k} nc}{p^n - q^n} - kc = \frac{P(k)nc}{P(n) - P(0)} - kc.$$

To prove the proposition, one need only prove

$$U(n) \geq U(k), \quad \forall k = 0, \dots, n.$$

Note that

$$\begin{aligned} U(n) - U(k) &= \frac{P(n)nc}{P(n) - P(0)} - nc - \frac{P(k)nc}{P(n) - P(0)} + kc \\ &= \frac{(P(n) - P(k))nc}{P(n) - P(0)} - (n - k)c \\ &= c(P(n) - P(k)) \left(\frac{n}{P(n) - P(0)} - \frac{n - k}{P(n) - P(k)} \right). \end{aligned}$$

It follows that $U(n) = U(0)$ and by (6) that $U(n) > U(k)$ for $k = 1, \dots, n-1$. Therefore given bonus contract (4), the agent's original incentive constraint (2) is also satisfied. Hence, (4) solves the original Program (P) as well. If Program (P) has any other solution $\tilde{w}(\cdot)$ then it must result in the same expected cost for the principal. Such a solution therefore also solves the relaxed program as it attains the same cost and satisfies constraint (3); but this is impossible given that (4) uniquely solves the relaxed program. \square

The above and the subsequent arguments hinge upon condition (5), which says that probability $P(k)$ exhibits increasing marginal returns to effort level k . The intuition for this result is as follows: the probabilities $p(z|e_i)$ of each individual signal z enter into $P(k)$ multiplicatively, which makes these individual probabilities complementary to each other in raising $P(k)$; as a result, exerting effort in an additional task generates a bigger increase in $P(k)$ when the total effort level k is higher. Since the bonus contract (4) only rewards the agent for output vector (z, \dots, z) , this probability structure makes the agent's payoff a convex function of total effort. Therefore, to implement all-working (H) we need only prevent all-shirking (L).

This analysis also helps reveal an important difference with the use of a single punishment across periods (Abreu, Milgrom and Pearce, 1991; Fuchs, 2007). In that case the agent strives to *minimize* the probability of punishment. As a result, efforts in different periods are not complements for the agent; rather, the agent's payoff displays diminishing marginal return to total effort level, so to implement all-working one need only prevent a single shirking as it is the most profitable deviation for the agent.

C All-or-Nothing Monitoring

The analysis in Section B clearly applies to any subset of tasks. Therefore, by Proposition 1, using an output-based bonus contract the principal's minimum (expected) cost for implementing effort level 1 in any $k = 1, \dots, n$ tasks is equal to

$$(7) \quad \frac{p^k \cdot kc}{p^k - q^k}.$$

When the principal can monitor the agent's effort levels in $n - k = 1, \dots, n - 1$ tasks she may choose to directly implement effort level 1 in these monitored tasks, paying c if effort level is 1 and zero otherwise, and use output-based bonus contract for the other k tasks. By Eq. (7), the principal's expected payment based on $n - k$ inputs and k outputs is given by

$$(8) \quad W(k) \equiv \frac{p^k kc}{p^k - q^k} + (n - k)c, \text{ for } k = 1, \dots, n.$$

If the principal monitors all n inputs then her payment is $W(0) = nc$. Although monitoring saves cost for the principal in motivating effort in the monitored tasks, it actually increases the principal's overall payment if not every task is monitored. This surprising result is proved in the following proposition.

Proposition 2. *If the principal can observe the agent's effort levels in $n - k$ but not in the other $k > 0$ tasks then the principal's optimal plan is to ignore the information on effort levels and reward the agent based solely on outputs in all tasks. If effort choices can be observed for all tasks then input-based compensation is optimal.*

PROOF: It is clear that $W(0) < W(k)$ for all $k = 1, \dots, n$, so input monitoring is the best if all tasks can be monitored. We shall prove the rest of the proposition by showing that $W(k) > W(k+1)$ for $k = 1, \dots, n-1$.

By (8), for $k = 1, \dots, n$,

$$W(k) - nc = \frac{q^k kc}{p^k - q^k} = \frac{kc}{\left(\frac{p}{q}\right)^k - 1} = \frac{kc}{\frac{P(k)}{P(0)} - 1}$$

where again $P(k) = p^k q^{n-k}$ for $k = 0, \dots, n$ is the probability of output vector $\theta^* = (z, \dots, z)$ when the total effort level equals k .

Therefore, with k unmonitored tasks the principal's rent payment $W(k) - nc$ is inversely related to $(P(k) - P(0))/k$. Hence, for $k = 1, \dots, n-1$, $W(k) > W(k+1)$ is equivalent to

$$(9) \quad \frac{P(k+1) - P(0)}{k+1} > \frac{P(k) - P(0)}{k}.$$

This last condition says that the average gain in the probability of θ^* per unit of effort increases with the total effort level; as such it follows from the increasing marginal return condition, Eq. (5). In conclusion, $W(k) > W(k+1) > W(0)$ for $1 \leq k < n$, completing the proof. \square

Therefore, when the principal can not monitor the agent's effort levels in all tasks it is better for her to use a bonus contract based only on output measures than also compensating the agent on the basis of effort levels in the monitored tasks. What drives this result is an informational gain from using joint output measures in incentive contracts when the principal's objective is to motivate the desired effort levels while minimizing the rent paid to the agent.¹¹ A single bonus paid only when performance is good in every task can simultaneously motivate effort in all tasks, which is cheaper for the principal than compensating the agent based

¹¹Bonding or entry fees can eliminate the rent the agent earns due to the nonnegativity constraints on wages. However, except for a few cases bonding or entry fees are rarely used in practice (Baker, Michael C. Jensen, and Kevin J. Murphy, 1988), possibly reflecting a potential moral hazard problem on the principal's part. Moreover, even if such arrangements are available so the agent does not face limited liability constraint, output-based pay is just as effective as monitoring: either way the principal need only pay the agent effort cost c per task.

on performances in individual tasks separately. More importantly, Proposition 2 shows that the principal's rent payment on k unmonitored tasks is *inversely* related to the per-task gain in the probability of the joint signal θ^* when the agent works in exactly k tasks; further, this average gain in probability increases and hence the total rent decreases in the number of unmonitored tasks. Therefore the information synergy is sufficiently strong that even though monitoring can reduce the agent's rent to zero on the monitored tasks, the higher rent paid on the unmonitored tasks actually outweigh the savings on the monitored tasks.

We should remark that in this model total output does not exhibit increasing returns to scale as the tasks are all identical; the benefit to the principal from using joint output measures is purely due to the optimal use of performance information in the design of incentive contracts. It is noteworthy that the principal's cost monotonically *increases* in the number of monitored tasks until there is perfect monitoring for all tasks, at which point monitoring becomes the better choice. Therefore, the effect of input monitoring is not monotone and the principal's best monitoring strategy is all or nothing.

D Extension and Discussion

In the preceding analysis we assumed that separate output measures are available for each task. This assumption made the analysis simple and intuitive and the result easy to interpret, but it is not essential. Suppose that there is only an aggregate output signal drawn from a finite set Θ ; given an effort vector $e = (e_1, \dots, e_n)$, each signal $\theta \in \Theta$ occurs with probability $P(\theta|e) \in (0, 1)$. The principal again would like to implement effort vector $H = (1, \dots, 1)$ with minimum cost. Assume that there is a signal θ^* such that for all $e \neq H$,

$$\frac{P(\theta^*|H)}{P(\theta^*|e)} \geq \frac{P(\theta|H)}{P(\theta|e)}, \quad \forall \theta$$

and the inequality is strict for some θ .¹² For simplicity, assume that probabilities $P(\theta^*|e)$ only depend on total effort level, so we can write them as $P(\sum_{i=1}^n e_i)$ instead. One can then show, along the lines of the previous analysis, that the conclusions of Proposition 2 remain

¹²This assumption is slightly more general than the standard *monotone likelihood ratio condition*, which in this case would require θ^* to be the highest output level.

valid as long as $P(\cdot)$ satisfies (5) and the following condition:¹³

$$(10) \quad \frac{\frac{P(n)}{P(n-k-1)} - 1}{k+1} > \frac{\frac{P(n)}{P(n-k)} - 1}{k}, \forall k = 1, \dots, n-1.$$

To see the intuition of the above condition, note that when the agent has exerted effort in $n-k$ tasks, $r(k) \equiv \left(\frac{P(n)}{P(n-k)} - 1 \right) / k$ measures the simple (as opposed to compounded) per task growth rate of the probability of θ^* for the k *remaining* tasks. Therefore, while (5) requires the probability of θ^* to grow by a larger amount for each additional unit of effort, (10) requires the forward growth *rate* of the probability to diminish when there are fewer tasks left. Both conditions are essential: the former guarantees the optimality of a bonus contract; the latter ensures all-or-nothing monitoring is profitable as one can show (as in Proposition 2) that after working on $n-k$ monitored tasks the agent's rent on k unmonitored tasks is inversely related to $r(k)$.

To see the connection with the previous case, note that the probabilities $P(k) = p^k q^{n-k}$ satisfy a type of scalability: $P(n)/P(n-k) = P(k)/P(0)$, which says that working on k additional tasks always increases the probability of θ^* by the same proportion regardless of the actions in the previous tasks; as a result, (10) is equivalent to (9), which as shown is implied by (5); thus (5) alone is sufficient for the result. This property is a consequence of symmetric independent distributions and may not hold in general when only aggregate signals are available. Eqs. (5) and (10) are the general sufficient conditions for the result.

We have assumed that the tasks are symmetric. Large asymmetries may make monitoring a better choice for the principal when the monitored tasks are more costly to implement using an output-based contract.¹⁴ Nevertheless, the insight of Proposition 2 can still apply to the other tasks where asymmetry is small. Also, the bonus schemes are clearly risky, so if the agent has a strong aversion to income risk it may not pay for the principal to use bonuses when monitoring is available. The model might still be applicable if the agent has decreasing risk aversion and the reservation wage is large enough to guarantee an acceptable standard of living; then the agent might be more receptive to bonus contracts.¹⁵

¹³Details are available upon request.

¹⁴This could also apply to some aspects of a job that the principal deems particularly important: for instance it may be desirable to monitor certain actions of the agent just to make sure the actions are legal and ethical.

¹⁵The implication is that asymmetry and risk aversion are important factors that determine the choice between input and output based pay.

II Conclusion

To conclude, a principal generally faces the choice of whether to reward an agent for outcome or for effort (input). We have shown that a multitask job and limited liability of the agent can make output-based incentive systems valuable for the principal: rewarding the overall outcome is more efficient than paying for piecemeal actions. This is true even under conditions that seem rather favorable for input-based systems: the principal knows exactly which action the agent should take in each task and can perfectly monitor the agent's actions in all but one task. To the extent that not every important aspect of a worker's job can be easily monitored, the result provides a strong rationale for the use of outcome-based rather than input-based incentive contracts. The result may shed light on the common use of bonuses and other output-based performance pay schemes in organizations even when many specific actions of a worker can be monitored at low cost.

III Appendix

PROOF OF LEMMA 1.

Let $\lambda \geq 0$ and $\mu(\theta) \geq 0$ be the Lagrange multipliers associated with constraint (3) and the nonnegativity constraints $w(\theta) \geq 0$, respectively. Then the first-order necessary condition for $w(\theta)$ is given by the following equation:

$$(11) \quad -P(\theta|H) + \lambda(P(\theta|H) - P(\theta|L)) + \mu(\theta) = 0.$$

If $\lambda = 0$ then $\mu(\theta) = P(\theta|H) > 0, \forall \theta$, and hence by complementary slackness $w(\theta) = 0, \forall \theta$, which clearly violates constraint (3). Thus, $\lambda > 0$.

Suppose for some θ , $w(\theta) > 0$ and hence $\mu(\theta) = 0$. If there is some $\phi \in Y^n$ with

$$\frac{P(\phi|L)}{P(\phi|H)} < \frac{P(\theta|L)}{P(\theta|H)}$$

then

$$-1 + \lambda \left(1 - \frac{P(\phi|L)}{P(\phi|H)} \right) > -1 + \lambda \left(1 - \frac{P(\theta|L)}{P(\theta|H)} \right) = -\frac{\mu(\theta)}{P(\theta|H)} = 0.$$

But then by (11), $\mu(\phi) < 0$, a contradiction. This implies that $w(\theta) > 0$ only if

$$\frac{P(\theta|L)}{P(\theta|H)} \leq \frac{P(\phi|L)}{P(\phi|H)}, \quad \forall \phi \in Y^n.$$

It follows that $w(\theta) > 0$ only for $\theta = (z, \dots, z)$. Namely, the agent should receive a bonus $w^* > 0$ only when output is good in all tasks. Then from constraint (3) with equality, one has

$$p^n w^* - nc = q^n w^*$$

which gives rise to the optimal wage contract in (4). □

References

- Abreu, Dilip, Paul Milgrom, and David G. Pearce.** 1991. “Information and Timing in Repeated Partnerships.” *Econometrica*, 59(6): 1713–33.
- Baker, George P.** 1991. “Incentive Contracts and Performance Measurement.” *Journal of Political Economy*, 100 (3): 598–614.
- Baker, George P., Michael C. Jensen, and Kevin J. Murphy.** 1988. “Compensation and Incentives: Practice vs. Theory.” *Journal of Finance*, 43(3): 593–616.
- Fuchs, William.** 2007. “Contracting with Repeated Moral Hazard and Private Evaluations.” *American Economic Review*, 97(4): 1432–48.
- Groshen, Erica and Alan Krueger.** 1990. “The Structure of Supervision and Pay in Hospitals.” *Industrial and Labor Relations Review*, 43(3): 134S–146S.
- Holmström, Bengt.** 1979. “Moral Hazard and Observability.” *Bell Journal of Economics*, 10 (1): 74–91.
- Holmström, Bengt and Paul Milgrom.** 1991. “Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design.” *Journal of Law, Economics and Organization*, Special Issue 7: 24–52.
- Khalil, Fahad and Jacques Lawarrée.** 1995. “Input versus Output Monitoring: Who Is the Residual Claimant?” *Journal of Economic Theory*, 66(1): 139–157.
- Lazear, Edward P.** 2000. “The Power of Incentives.” *American Economic Review*, 90 (2): 410–414.

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

MacDonald, Glenn and Leslie M. Marx. 2001. “Adverse Specialization.” *Journal of Political Economy*, 109 (4): 864–99.

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

Raith, Michael. 2005. “Specific Knowledge and Performance Measurement.” <http://www.simon.rochester.edu/fac/raith>.

Shapiro, Carl and Joseph E. Stiglitz. 1984. “Equilibrium Unemployment as a Worker Discipline Device.” *American Economic Review*, 74(3): 433–44.