When Should the Principal Work?  
Managerial Interference and Motivational Monitoring

Michael T. Rauh  
Indiana University  
mtrauh@indiana.edu

April 13, 2018

Abstract

We consider a principal who can work with the agent and can commit to a mixed strategy in effort before the agent chooses his own effort. We derive conditions under which the principal strictly mixes and show that there are precisely two types of mixed-strategy equilibria. The first is what we call managerial interference, where the principal works with positive probability because it is efficient (her marginal product of effort exceeds the marginal cost) even though it reduces the effectiveness and motivation of the agent and obscures his contribution to output. The second is motivational monitoring, where the principal works with positive probability, even though it is inefficient, because it motivates the agent and generates information about his effort. We then show that if the principal observes a separate signal about the agent’s effort but only when she works with him then the principal never mixes because in this case her mixed strategy is uninformative.

JEL Classifications: D86.

Keywords: double moral hazard, help, managerial interference, micromanagement, motivational monitoring, teamwork.
1 Introduction

A sales director can help one of her sales representatives in numerous ways. She can set up contacts for him, help him tailor sales pitches to specific clients, keep him informed about recent product developments and the state of the competition, and even accompany him to meetings with clients. A central question of this paper is: when should a principal work with the agent? As we will see, the answer revolves around more than just productivity. When the principal works with the agent, her effort distorts the performance measure on which the agent is evaluated. In particular, the principal’s intervention will affect the agent’s choice of effort — positively when it is regarded as “help” and negatively when it is regarded as interference or “micromanagement.” Furthermore, working with the agent is likely to generate information that can be used to monitor him. The decision of whether or not to work with the agent is therefore inextricably linked to the design of incentives. What should the commission be when sales reflect not only the sales representative’s effort but also the sales director’s?

A novel aspect of our model is that we allow the principal to commit to a mixed strategy in effort at the start of the game, before the agent chooses his own effort. This first-mover advantage of the principal reflects an important difference between the principal and agent in an employment context: unlike the agent, the principal has complete discretion over a wide range of activities within the organization.¹ For example, suppose the principal initiates a joint venture with another firm that with probability one-half will demand her full time and attention, making it impossible for her to work with the agent, and with probability one-half will require no action on her part this period, freeing her up for team production. The principal has therefore committed herself to a strictly mixed strategy in effort where she works with probability one-half. An employee (as opposed to an independent agent) is unlikely to possess such commitment power. In addition to conferring a first-mover advantage, it also means that there is no hidden action problem for the principal, whereas the agent chooses his own effort under moral hazard. This is a key difference between our model and the literature on double moral hazard, which includes Carmichael (1983), Demski and Sappington (1991), Bhattacharyya and Lafontaine (1995), Al-Najjar (1997), Gupta and Romano (1998), and Poblete and Spulber (2017).² This raises another question: how does the

¹We assume the agent chooses his effort under moral hazard, so the principal cannot move “second” in the sense of choosing her effort after observing the agent’s effort. It is clear that the principal would rather move first, as opposed to moving simultaneously, because when she moves first, she always has the option of committing to her Nash equilibrium effort in the normal form game.

²The moral hazard of the principal is an important element in this literature. For example, Carmichael (1983), Al-Najjar (1997), and Gupta and Romano (1998) develop the idea that, with multiple agents, the optimal contract for each individual agent conditions on all the agents’ outputs, not because they are correlated or for help reasons,
agent’s moral hazard affect the principal’s optimal effort, which is not subject to moral hazard?

The principal’s optimal mixed strategy for effort depends on three factors. The first two are her marginal product of effort and the opportunity cost of her time. The third and more interesting factor is the effect of her effort on the moral hazard problem: the informativeness of output and the agent’s expected marginal product of effort. When the principal does not work, output reflects the individual performance of the agent, but when the principal works, it reflects their group performance. We often think of group performance as being less informative about an agent’s effort than individual performance, but the opposite holds when efforts are complementary. Similarly, we often assume that group performance measurement promotes free-riding, but this is only the case when efforts are substitutes. When efforts are complements, the principal’s effort increases the agent’s expected marginal product of effort, which induces him to reciprocate — the opposite of free-riding. We say that there are team economies when efforts are complements and the principal’s effort increases the informativeness of output and team diseconomies in the opposite case.

Similar factors are at work in the literature on help, including Itoh (1991, 1992, 1994), Auriol, Frieben, and Pechlivanos (2002), Chalioti (2016), and Ishihara (2017), when the principal is deciding whether or not she wants to induce multiple agents to help each other. In this literature, the principal induces more help by offering stronger incentives for group performance or for other agents’ outputs. An important difference is that in our model the principal can extend help directly, without the need to incentivize any third party, which means that she has independent control over incentives and help. This independent control allows for a broader range of endogenous relations between them. In particular, the principal may increase or decrease incentives on group performance when she works with the agent. Unlike the literature on help, the principal reduces such incentives when there are team economies, but increases them when there are team diseconomies to prevent free-riding.

We provide a necessary and sufficient condition for the optimal effort strategy of the principal to be strictly mixed. When one of the above three factors is dominant, this condition is violated, and the principal chooses a pure strategy instead. For example, the principal chooses zero effort with probability one when her opportunity cost of time is sufficiently high and in that case the optimal incentive is the same as in the textbook model where the principal does not work for exogenous reasons. The necessary and sufficient condition for a strictly mixed strategy is satisfied when the above three factors are evenly balanced. This does not mean, however, that the principal is indifferent between low and high effort. In a simultaneous-move game, an optimal mixed strategy
has the property that the player is indifferent between all the actions in the support. In contrast, in our extensive-form game where the principal moves first, she strictly prefers her optimal mixed strategy over all others, including her available pure strategies. We then show that the optimal mixed strategy can take one of two forms: managerial interference or motivational monitoring.

The principal chooses managerial interference when it is efficient for her to work in the sense that her marginal product of effort exceeds her opportunity cost of time, but this is evenly balanced by team diseconomies. In particular, when the principal works, it reduces the agent’s expected marginal product of effort because efforts are substitutes. The agent will therefore free ride unless the principal increases incentives. Managerial interference is therefore marked by a complementarity between incentives and the principal’s effort in the sense that whenever she increases one, she also increases the other. The principal’s effort also reduces the informativeness of output, so the agent collects more information rent in the form of a higher expected transfer. In these circumstances, the principal would like to promise that she will never interfere because if the agent believes her then she can offer him a lower bonus and expected transfer. Of course, the principal would then surprise the agent by working, so this cannot be an equilibrium. The principal therefore mixes instead. We call this “managerial interference” because the principal’s effort reduces the agent’s effectiveness and motivation (his expected marginal product of effort) and obscures his contribution to output (the informativeness of output). As a result, the principal has to offer a higher bonus and expected transfer. Since the principal works with only some positive probability, whereas in a first best world she works with probability one, managerial interference is characterized by an under-supply of the principal’s effort, even though she is not directly subject to moral hazard.

The principal chooses motivational monitoring when her marginal product of effort is less than her opportunity cost, so it is actually inefficient for her to work, but she sometimes does because of team economies. In this case, the principal’s help increases the agent’s expected marginal product of effort and makes output more informative about the agent’s effort and therefore serves to monitor him. But this form of monitoring is quite different from traditional monitoring, which consists of pure surveillance and is likely to signal distrust and provoke resentment. This is because the monitoring in this paper is a by-product of teamwork that increases the agent’s marginal product of effort. For example, suppose the sales director skillfully matches potential clients with the firm’s product offerings and then directs the sales representative to follow up. This increases the sales representative’s motivation because a good faith effort will result in a sale with high probability.

\footnote{We are tempted to call this “micromanagement” but this term also suggests an excessive focus on details by the principal which is not in our model. But this class of mixed strategy equilibria does capture most of the negative connotations of the term.}
It also serves to monitor him because the director will be skeptical if sales are low. In a first best world, the sales director may not do this because the opportunity cost of her time is high, but in a second best world of moral hazard, she may with positive probability because it allows her to reduce the commission. We refer to this as “motivational monitoring” because the principal works even though it is inefficient for her to do so because it motivates the agent and highlights his contribution. In this class of mixed-strategy equilibria, incentives and help are substitutes and the principal over-supplies help.

We present these results in section 2 below, where we assume that output is the only contractible performance measure. After that, we turn to the more general question as to the conditions under which the principal strictly mixes. In section 3, we show that the above results continue to hold when the principal can contract not only on output, but also on a separate signal of individual performance, so our results are not an artifact of the traditional teams setting where only group performance can be rewarded. Throughout the paper, we assume that the principal’s effort, like the agent’s, is non-verifiable to third parties such as the courts and therefore cannot be specified as part of an enforceable contract. In section 4, we consider the case where the principal can only observe the agent’s individual performance when she works with him; i.e., the principal observes only output when she does not work but she observes both output and individual performance when she does. We show that in this case the principal never mixes because her choice of mixed strategy is uninformative.

The paper concludes in section 5. All proofs are in the appendix.

2 Group Performance

2.1 Model Primitives

We extend the elementary textbook model [see Laffont and Martimort (2002, Chapter 4)] with one principal and one agent, both risk neutral, where the agent has limited liability, to the case where the principal makes an optimal effort decision. We use this framework as our starting point because it allows us to cleanly distinguish between the effects of complementarities in efforts, on the one hand, and the informativeness of the principal’s mixed strategy in effort on the other.

There are two players: a principal (she) and agent (he). Both are risk neutral. The agent makes a binary effort choice $e \in \{0, 1\}$, where the cost of $e = 0$ is zero and the cost of $e = 1$ is $\psi_A > 0$. The agent’s payoff is $t - C(e)$, where $t$ is the transfer from the principal and $C(e)$ is the cost of
effort. We assume limited liability, so the transfer $t$ must be nonnegative. Likewise, the principal also chooses a binary effort level $a \in \{0, 1\}$, where the cost of $a = 0$ is zero and the cost of $a = 1$ is $\psi_P > 0$. As discussed in the introduction, we assume the principal can commit to a mixed strategy for effort. We discuss this assumption in detail later. Let $\alpha$ be the probability that the principal works, $a = 1$. We assume throughout the paper that the principal wants to implement high effort $e = 1$ by the agent but we make no fixed assumption about whether $a = 1$ is optimal.

We now describe the stochastic production function. Output $Q$ takes two possible values: high $\overline{Q}$ or low $\underline{Q}$, where $0 \leq Q < \overline{Q}$. Let $\Delta Q = \overline{Q} - Q > 0$. Let $\pi$ be the probability that output is high $Q = \overline{Q}$ when the principal and agent both work $a = e = 1$, $\pi_P$ the probability that $Q = \overline{Q}$ when only the agent works, $a = 0$ and $e = 1$, $\pi_A$ when $a = 1$ and $e = 0$, and $\pi_{PA}$ when $a = e = 0$ (the subscript on $\pi$ denotes who shirks). We assume

$$0 < \pi_{PA} < \pi_A < \pi < 1$$

(1)

$$0 < \pi_P < \pi < 1,$$

(2)

so the probability that output is high is strictly increasing in both efforts. Let

$$\Delta \pi_P = \pi - \pi_P$$

(3)

$$\Delta \pi_A = \pi - \pi_A$$

(4)

$$\Delta \pi_{P,PA} = \pi_P - \pi_{PA}$$

(5)

$$\Delta \pi_{PA} = \pi - \pi_{PA},$$

(6)

where $\Delta \pi_P$ is the principal’s marginal product of effort when the agent works, $\Delta \pi_A$ is the agent’s marginal product of effort when the principal works, $\Delta \pi_{P,PA}$ is the agent’s marginal product of effort when the principal does not work, and $\Delta \pi_{PA}$ is the change in the probability that output is high when both players switch from low to high effort. We make no assumptions about which player has the higher marginal product of effort. For example, the principal may be more productive than the agent when she has superior knowledge and experience, while the agent may be more productive when he is highly specialized. Likewise, we make no assumptions about the relative magnitudes of the marginal costs of effort ($\psi_A, \psi_P$).

---

4 Limited liability arises when the agent has zero initial wealth, the courts will not enforce a penalty, or the agent is free to quit at any time to avoid a fine. The first reason rules out achieving the first best by selling the store.

5 We assume that when the incentive compatibility and participation constraints bind, the agent is willing to break the indifference in the principal’s favor. In particular, the agent accepts the contract and chooses high effort $e = 1$, which eliminates the need to consider mixed strategies by the agent.
2.2 First Best

A performance measure is \textit{contractible} if it is observable and verifiable to third parties such as the courts and can therefore be specified as part of an enforceable contract. We briefly consider the case where the agent’s effort \( e \) is contractible. In that case, the principal can implement \( e = 1 \) with a contract that pays the agent \( \psi_A \) when she observes \( e = 1 \) and zero when she observes \( e = 0 \). This satisfies both the incentive compatibility and participation constraints. The following result provides conditions under which it is first best for the principal and agent to both work. Later on we will consider that case, but we do not assume it in general.

\textbf{Lemma 1} If

\[
\Delta \pi_P \Delta Q - \psi_P > 0 \\
\Delta \pi_A \Delta Q - \psi_A > 0 \\
\Delta \pi_{PA} \Delta Q - \psi_P - \psi_A > 0
\]

\( \text{the first best solution is to implement } a = 1 \text{ and } e = 1, \text{ which generates expected profit} \)

\[
\bar{\pi}Q + (1 - \pi)\bar{Q} - \psi_P - \psi_A.
\]

The first condition states that the marginal net benefit \( \Delta \pi_P \Delta Q - \psi_P \) of the principal is positive when the agent works, the second that the marginal net benefit of the agent is positive when the principal works, and the third that the marginal benefit of an increase in effort by both players exceeds the marginal costs.

2.3 Moral Hazard

From now on, we assume that both efforts \( a \) and \( e \) are non-contractible.\footnote{Another possibility is that the principal receives incontrovertible evidence about the agent’s effort \( e \) which is verifiable to third parties, but only when she works with the agent. I.e., when \( a = 0 \) the only contractible performance measure is output \( Q \) but \( e \) is contractible when \( a = 1 \). This case is trivial because the principal can achieve the first best with \( a = 1 \) and a contract that pays zero when \( e = 0 \) and \( \psi_A \) when \( e = 1 \).} In this section, we assume that output \( Q \) is the only contractible performance measure. We consider further performance measures in future sections. Let \( \bar{t} \) be the transfer to the agent when output is high \( Q = \bar{Q} \) and \( t \) when output is low \( Q = \underline{Q} \).
2.4 Team Economies and Diseconomies

In classic papers by Alchian and Demsetz (1972) and Holmström (1982), the teams problem is described as a situation where an individual agent has reason to shirk because the only available performance measure is team performance. In this paper, we formalize this idea explicitly because our results crucially depend on it.

At the start of the game, when the principal chooses her mixed strategy $\alpha$ and the contract $(\bar{t}, \bar{q})$, she has to decide what the reward $\bar{t}$ in state $\bar{Q}$ will be given her mixed strategy $\alpha$ but not its realization. The appropriate statistical inference question is therefore: what is the informational content of state $\bar{Q}$ given $\alpha$? Let

$$p^\alpha = \alpha \pi + (1 - \alpha) \pi_P$$
$$p^\alpha_A = \alpha \pi_A + (1 - \alpha) \pi_{PA}$$
$$L_\alpha = \frac{p^\alpha_A}{p^\alpha}$$

where $p^\alpha$ is the probability that output is high when the agent works, $p^\alpha_A$ the probability that output is high when the agent shirks, and $L_\alpha$ the likelihood ratio. The smaller the likelihood ratio, the less likely it is that the agent shirked in state $\bar{Q}$, and the clearer is the signal provided by that state that effort was high. We therefore refer to

$$1 - L_\alpha = \frac{\Delta \pi_{P,PA} + \alpha (\Delta \pi_A - \Delta \pi_{P,PA})}{\pi_P + \alpha \Delta \pi_P}$$

as the informativeness of output when the principal chooses $\alpha$. Let $\Delta L = L_1 - L_0$. If $\Delta L > 0$ then output is more informative when $a = 0$ than when $a = 1$. Intuitively, output provides a signal of the agent’s individual performance when the principal does not work. When the principal does work, output provides a signal of group performance. If $\Delta L > 0$ then individual performance is more informative than group performance. Note that $\Delta L > 0$ is equivalent to

$$\pi_P \pi_A > \pi \pi_{PA},$$

which states that $\pi$ is strictly log-submodular in $(a, e)$ on the lattice $\{(0, 0), (1, 0), (0, 1), (1, 1)\}$. In contrast, group performance is more informative when $\Delta L < 0$, or $\pi$ is strictly log-supermodular in $(a, e)$. Intuitively, group performance is more informative when efforts are complementary, at least in the sense of strict log-supermodularity. An increase in $\Delta L$ means that individual performance
has become more informative relative to group performance, so we refer to $\Delta L$ as the relative informativeness of individual performance.

Informally, a teams problem exists when group performance is less informative than individual performance AND the agent is motivated to shirk in the teams context (i.e., when the principal works). As we will see, the agent is motivated to shirk when $\pi$ is strictly submodular in $(a,e)$

$$\pi + \pi_{PA} < \pi_A + \pi_P,$$

which is equivalent to $\Delta \pi_A < \Delta \pi_{P,PA}$. We say that efforts are substitutes in this case. In contrast, the agent will be motivated to work when $\pi$ is strictly supermodular in $(a,e)$

$$\pi + \pi_{PA} > \pi_A + \pi_P,$$

which is equivalent to $\Delta \pi_A > \Delta \pi_{P,PA}$. We say that efforts are complements in this case. The following lemma is a slight modification of a well-known result in lattice programming (see the proof in the appendix).

Lemma 2

(i) If $\Delta \pi_A < \Delta \pi_{P,PA}$ then it must be the case that $\Delta L > 0$ (team diseconomies).

(ii) If $\Delta L < 0$ then $\Delta \pi_A > \Delta \pi_{P,PA}$ (team economies).

According to (i), when efforts are substitutes $\Delta \pi_A < \Delta \pi_{P,PA}$ it must also be the case that group performance is less informative than individual performance $\Delta L > 0$ (the reverse implication does not necessarily hold). Since one implies the other, and because both aspects of the teams problem are then present, we say that there are team diseconomies when $\Delta \pi_A < \Delta \pi_{P,PA}$. On the other hand, (ii) states that if group performance is more informative $\Delta L < 0$ it must be the case that efforts are complements $\Delta \pi_A > \Delta \pi_{P,PA}$ and we say that there are team economies when $\Delta L < 0$. There is a third possibility, the mixed case where $\Delta \pi_A > \Delta \pi_{P,PA}$ and $\Delta L > 0$. In this case, individual performance is more informative than group performance but efforts are complementary, so there are neither team economies nor diseconomies.

--

7This definition supersedes the use of the word “complementary” in the previous paragraph.
2.5 Timing

The timing of the game is as follows. (i) At the start of the game, the principal chooses the probability $\alpha$ that she works and the transfers $(\bar{t}, \bar{t})$ conditional on output. We assume she can commit to both choices, so the principal is not subject to moral hazard (hidden action). (ii) If the agent rejects the principal’s offer then both players receive a payoff of zero. If the agent accepts then he chooses his effort level. (iii) The principal’s effort $a$ and output $Q$ are realized and the principal makes the appropriate transfer. We normalize the price of output to be one, so the principal’s profit is output minus the transfer and the principal’s cost of effort.

2.6 Constraints

Since the principal wants to implement $e = 1$, the incentive compatibility constraint is

$$
\alpha \left[ \pi \bar{t} + (1 - \pi)\bar{t} \right] + (1 - \alpha) \left[ \pi P \bar{t} + (1 - \pi P)\bar{t} \right] - \psi_A \geq 0
$$

(18)

The left-hand side of (18) is the expected payoff to the agent when he chooses $e = 1$ and the right-hand side is the expected payoff when he chooses $e = 0$. We can re-write (18) as

$$
\alpha \left[ \pi A \bar{t} + (1 - \pi A)\bar{t} \right] + (1 - \alpha) \left[ \pi PA \bar{t} + (1 - \pi PA)\bar{t} \right] - \psi_A \geq 0
$$

(19)

where the term in square brackets is the agent’s expected marginal product of effort and $\Delta t = \bar{t} - \bar{t}$ is the bonus when output is high. If $\Delta \pi_A < \Delta \pi_{P,PA}$ then efforts are substitutes and an increase in $\alpha$ reduces the agent’s expected marginal product of effort. In contrast, an increase in $\alpha$ increases the agent’s expected marginal product of effort when efforts are complements. The participation constraint is

$$
\alpha \left[ \pi \bar{t} + (1 - \pi)\bar{t} \right] + (1 - \alpha) \left[ \pi P \bar{t} + (1 - \pi P)\bar{t} \right] - \psi_A \geq 0
$$

(20)

This is redundant because it is implied by nonnegative transfers and the incentive compatibility constraint (18). We therefore drop it in what follows.
2.7 Optimal Contract

We now find the minimum expected cost to the principal of implementing $e = 1$ given her mixed strategy $\alpha$. The expected transfer is

$$
\alpha \left[ \pi \bar{t} + (1 - \pi) \bar{t} \right] + (1 - \alpha) \left[ \pi P \bar{t} + (1 - \pi) \bar{t} \right],
$$

which can be re-written as

$$
\bar{t} + (\pi P + \alpha \Delta \pi P) \Delta t.
$$

The principal therefore chooses the transfers $\bar{t} \geq 0$ and $\bar{t} \geq 0$ to minimize (22) subject to the incentive compatibility constraint (19).

**Lemma 3** The optimal transfers are $\bar{t} = 0$ and

$$
\bar{t} = \frac{\psi_A}{\Delta \pi_{P,PA} + \alpha (\Delta \pi_A - \Delta \pi_{P,PA})}.
$$

The bonus $\bar{t}$ is strictly decreasing (increasing) in $\alpha$ when there are team economies (diseconomies).

The proof is trivial and omitted. This result shares some common features with the elementary textbook model. In particular, the principal awards a bonus $\bar{t}$ in the most informative state $Q$ with a zero transfer in the other state $Q$. Ideally, the principal would like to impose a negative transfer when output is low but cannot because of limited liability. Since she cannot fine the agent when output is low, the optimal incentive must take the form of a bonus when output is high. The principal therefore sets $\bar{t}$ at its lowest possible level (zero) and chooses the minimum transfer $\bar{t}$ which still induces $e = 1$, so that the incentive compatibility constraint binds. Note that a zero transfer when output is low is equivalent to firing the agent because his outside option is zero.

Unlike the textbook model, the bonus $\bar{t}$ depends on the principal’s mixed strategy $\alpha$. If $\alpha = 0$ then (23) reduces to the textbook bonus

$$
\bar{t} = \frac{\psi_A}{\Delta \pi_{P,PA}},
$$

where $\Delta \pi_{P,PA}$ is the agent’s marginal product of effort when the principal does not work. More generally, the denominator of $\bar{t}$ in (23) is the agent’s expected marginal product of effort and the relationship between $\bar{t}$ and the probability $\alpha$ that the principal works depends on whether efforts are substitutes $\Delta \pi_A < \Delta \pi_{P,PA}$ or complements $\Delta \pi_{P,PA} < \Delta \pi_A$. If efforts are substitutes then an increase in $\alpha$ reduces the agent’s expected marginal product of effort. This reduces the agent’s
motivation in the sense that now the incentive compatibility constraint (19) is slack and he will shirk unless the principal increases the bonus. In this case, $\tilde{t}$ and $\alpha$ are complements in the sense that the principal increases one whenever she increases the other. When efforts are complements, an increase in $\alpha$ increases the agent’s expected marginal product of effort, which allows the principal to reduce the bonus $\tilde{t}$. In this case, the principal’s effort acts like an incentive, so $\alpha$ and $\tilde{t}$ are substitutes. The bonus $\tilde{t}$ is therefore larger than the textbook bonus when there are team diseconomies and smaller when there are team economies.

2.8 The Optimal $\alpha$

Substituting the optimal transfers into (22), we find the optimal expected transfer

$$E(t) = \frac{\pi_P + \alpha \Delta \pi_P}{\Delta \pi_{P,P,A} + \alpha (\Delta \pi_A - \Delta \pi_{P,P,A})} \psi_A = \frac{1}{1 - L}\psi_A.$$  

(25)

The expected transfer $E(t)$ is the inverse of the informativeness of output $1 - L\alpha$ scaled by the agent’s opportunity cost of effort $\psi_A$. Since $0 < L\alpha < 1$, the agent receives information rent $E(t) - \psi_A > 0$ over and above his outside option. This is because the principal can only reward and not punish due to the agent’s limited liability. The expected transfer $E(t)$ is strictly decreasing in $1 - L\alpha$, so the agent collects less information rent the more informative output is.

The marginal effect of $\alpha$ on the expected transfer is given by

$$\frac{\partial E(t)}{\partial \alpha} = \frac{\pi \pi_P \Delta L}{[\Delta \pi_{P,P,A} + \alpha (\Delta \pi_A - \Delta \pi_{P,P,A})]^2} \psi_A.$$  

(26)

If $\Delta L > 0$ then individual performance is more informative then group performance, so an increase in $\alpha$ reduces the informativeness of output $1 - L\alpha$, which increases the expected transfer as the agent collects more information rent. In this case, the marginal effect of $\alpha$ on $E(t)$ is a marginal cost to the principal. If $\Delta L < 0$ then group performance is more informative, so an increase in $\alpha$ makes output more informative, which reduces the expected transfer. In this case, the marginal effect of $\alpha$ on $E(t)$ is a marginal benefit.

When the principal implements $e = 1$, expected output is given by

$$E(Q) = \alpha \left[ \pi Q + (1 - \pi)Q \right] + (1 - \alpha) \left[ \pi P Q + (1 - \pi_P)Q \right]$$

(27)

$$= Q + (\pi_P + \alpha \Delta \pi_P) \Delta Q.$$  

(28)
The expected profit of the principal is therefore

\[ \Pi = E(Q) - \alpha \psi_P - E(t). \]  

(29)

Given the optimal transfers, the principal chooses \( \alpha \) to maximize (29) subject to \( 0 \leq \alpha \leq 1 \).

**Lemma 4**  *Expected profit \( \Pi \) is strictly concave in \( \alpha \) if and only if there are team economies or team diseconomies. Otherwise, \( \Pi \) is linear or strictly convex.*

Our interest in the strict concavity of expected profit \( \Pi \) as a function of \( \alpha \) is more than just a question of the uniqueness of the optimal \( \alpha \). We are interested in situations where the principal chooses a strictly mixed strategy for \( \alpha \), and this can only happen when \( \Pi \) is strictly concave because the result will be a corner solution when \( \Pi \) is linear or strictly convex. Differentiating,

\[ \frac{\partial \Pi}{\partial \alpha} = \Delta \pi_P \Delta Q - \psi_P - \frac{\partial E(t)}{\partial \alpha}. \]  

(30)

Since the marginal net benefit \( \Delta \pi_P \Delta Q - \psi_P \) of the principal’s effort is constant, the curvature of \( \Pi \) is determined by the final term \( \partial E(t)/\partial \alpha \). In the case of team economies (\( \Delta L < 0 \) and \( \Delta \pi_A > \Delta \pi_{P,A} \)), we conclude from (26) that \( -E(t) \) increases at a decreasing rate. An increase in \( \alpha \) therefore has a diminishing marginal benefit. In the case of team diseconomies (\( \Delta \pi_A < \Delta \pi_{P,A} \) and \( \Delta L > 0 \)), an increase in \( \alpha \) corresponds to an increasing marginal cost. It follows that \( \Pi \) is strictly concave in \( \alpha \) if and only if there are team economies or diseconomies.

We now characterize the conditions under which the principal strictly mixes.

**Proposition 1**

(i) The principal uses a strictly mixed strategy \( 0 < \alpha < 1 \) if and only if one of the following two mutually exclusive conditions holds:

\[ (\Delta \pi_A)^2 < \frac{\pi \pi_P \Delta L}{\Delta \pi_P \Delta Q - \psi_P} \psi_A < (\Delta \pi_{P,A})^2 \quad \text{(managerial interference)} \]  

(31)

or

\[ \Delta L < 0 \text{ and } (\Delta \pi_{P,A})^2 < \frac{\pi \pi_P \Delta L}{\Delta \pi_P \Delta Q - \psi_P} \psi_A < (\Delta \pi_A)^2 \quad \text{(motivational monitoring)}. \]  

(32)
(ii) Let
\[ \alpha_{MI} = \frac{1}{\Delta \pi_{P,PA} - \Delta \pi_L} \left\{ \Delta \pi_{P,PA} - \left( \frac{\pi \pi_P \Delta L}{\Delta \pi_P \Delta Q - \psi_P} \right)^{1/2} \right\}. \] (33)

When (31) holds, the optimal mixed strategy for the principal is
\[ \alpha = \begin{cases} 
0 & \text{if } \alpha_{MI} \leq 0 \\
\alpha_{MI} & \text{if } 0 < \alpha_{MI} < 1 \\
1 & \text{if } \alpha_{MI} \geq 1.
\end{cases} \] (34)

(iii) Let
\[ \alpha_{MM} = \frac{1}{\Delta \pi_A - \Delta \pi_{P,PA}} \left\{ \left( \frac{\pi \pi_P \Delta L}{\Delta \pi_P \Delta Q - \psi_P} \right)^{1/2} - \Delta \pi_{P,PA} \right\}. \] (35)

When (32) holds, the optimal mixed strategy for the principal is
\[ \alpha = \begin{cases} 
0 & \text{if } \alpha_{MM} \leq 0 \\
\alpha_{MM} & \text{if } 0 < \alpha_{MM} < 1 \\
1 & \text{if } \alpha_{MM} \geq 1.
\end{cases} \] (36)

To interpret the result, we note that condition (31) requires \( \Delta \pi_{P,PA} > \Delta \pi_A \). From Lemma 2, this is the case of team diseconomies. In contrast, (32) requires \( \Delta L < 0 \), which implies team economies. The two conditions are therefore mutually exclusive. The proposition then states that the principal strictly mixes if and only if exactly one of those two conditions holds. Which condition holds is crucial because it determines the nature of the principal’s mixed strategy, which can only be one of two types.

We first discuss the class of mixed-strategy equilibria corresponding to (31), which implies team diseconomies: individual performance is more informative than group performance \( \Delta L > 0 \) and efforts are substitutes \( \Delta \pi_A < \Delta \pi_{P,PA} \). Given these two conditions, (31) dictates that the principal’s marginal net benefit \( \Delta \pi_P \Delta Q - \psi_P \) must be strictly positive. It is therefore efficient for the principal to work, but when she does her effort generates two further negative effects: it reduces the agent’s expected marginal product of effort because efforts are substitutes and it reduces the informativeness of output \( 1 - L \) because individual performance is more informative than group performance. Given the extent to which efforts are substitutes \( \Delta \pi_{P,PA} - \Delta \pi_A \), (31) states that the principal strictly mixes if and only if there is a balance between these three effects. Otherwise, the principal chooses a pure strategy, \( a = 0 \) or \( a = 1 \). For example, if the difference between \( \Delta \pi_{P,PA} \) and \( \Delta \pi_A \) is small then the principal mixes when the ratio between \( \Delta L \) and \( \Delta \pi_P \Delta Q - \psi_P \) is close
to one, so that all three effects are relatively balanced. This does not mean that the principal is indifferent. Since $\Pi$ is strictly concave in $\alpha$, the principal strictly prefers her optimal mixed strategy over all other mixed strategies in $[0, 1]$, including the pure strategies $\alpha = 0$ and $\alpha = 1$.

We refer to this class of mixed-strategy equilibria as *managerial interference*. Consider the example of the sales director and her sales representative. An important part of the sales rep’s job is doing market research, preparing sales presentations, and communicating with clients. For these activities, efforts are substitutes because a reduction in effort by one person can be compensated for by an (not necessarily equivalent) increase in effort by the other. The sales director will want to “help out” in these activities when she is efficient at doing them but this will provoke free-riding because it reduces the sales rep’s motivation and makes sales less informative about his effort. Why should he work hard on a sales presentation when the director is doing it and sales are likely to be high even when he shirks? If the director always interferes then she will have to increase the commission to control free-riding but if she never interferes then the efficiency gains from her potential contribution are wasted. The director therefore mixes, which allows her to offer a lower commission (relative to the case where she always interferes) but also to employ her considerable skills a certain percentage of the time.

The comparative statics of managerial interference are intuitive: from (33) we observe that the principal works with higher probability $\alpha_{MI}$ when the marginal net benefit $\Delta \pi_P \Delta Q - \psi_P$ of her effort is higher and the relative informativeness of individual performance $\Delta L$ and the degree to which their efforts are substitutes $\Delta \pi_P, PA - \Delta \pi_A$ are lower.

We now consider (32), which implies team economies. As before, the principal strictly mixes when there is a balance between her marginal net benefit $\Delta \pi_P \Delta Q - \psi_P$, the extent to which efforts are complements $\Delta \pi_A - \Delta \pi_P, PA$, and the relative informativeness of individual performance $\Delta L$. A major difference is that (32) implies $\Delta \pi_P \Delta Q - \psi_P < 0$ because $\Delta L < 0$. It is therefore inefficient for the principal to work but she sometimes does because it increases the agent’s expected marginal product of effort and improves the informativeness of output. We refer to this class of mixed strategies as *motivational monitoring* because the principal sometimes works even though it is inefficient because it motivates the agent and provides information about his effort.

In the previous case of managerial interference, efforts were substitutes. Now consider the case where the sales director draws upon her extensive network of contacts and knowledge about the firm’s product offerings to match specific products to potential clients, so she is confident that a good-faith sales effort will result in a sale. In that case, the sales rep will be motivated to follow up on those contacts not only because the odds of a sale are high but also because low sales will
be highly indicative of shirking. The sales director’s effort therefore motivates but also serves a monitoring function, so the sales rep will be willing to work hard at a lower commission. In a first best world where effort is perfectly observable, the director’s initiative may not be worth the opportunity cost of her time, but in a second best world of moral hazard, the cost savings may induce her to offer stochastic help. The result is a form of stochastic monitoring with positive rather than adversarial connotations because the director’s activity takes the form of productive teamwork.

The comparative statics of motivational monitoring are somewhat different from managerial interference. As before, the principal works with higher probability $\alpha_{MM}$ when her marginal net benefit of effort $\Delta\pi_P\Delta Q - \psi_P$ is higher (i.e., less negative) and the relative informativeness of individual performance $\Delta L$ is lower (i.e., more negative). In the case of managerial interference, $\alpha_{MI}$ is strictly decreasing in the degree $\Delta\pi_{P,A} - \Delta\pi_A$ to which efforts are substitutes, whereas in the case of motivational monitoring, $\alpha_{MM}$ is strictly decreasing in the degree $\Delta\pi_A - \Delta\pi_{P,A}$ to which efforts are complements, which may seem counterintuitive. The difference is that in the present context the principal’s marginal net benefit of effort is negative and an increase in complementarity allows her to induce high effort with a smaller $\alpha_{MM}$.

In models where the principal does not work for exogenous reasons, the inefficiency of moral hazard is that the agent’s second best effort can be less than first best. In this paper, we have assumed that the principal always implements $e = 1$ but when implementation costs due to moral hazard become sufficiently large, the principal will implement $e = 0$ instead. In our model, there is a second source of inefficiency because moral hazard can also distort the principal’s effort even though she is not directly subject to moral hazard. In the case of managerial interference, the principal’s marginal net benefit of effort is positive and we can have a situation where the first best entails $e = 1$ and $a = 1$, so the principal works with probability one, whereas the second best entails a strictly mixed strategy $0 < \alpha < 1$ where the principal works with probability less than one. This is because the principal’s effort makes the moral hazard problem worse when there are team diseconomies, which increases the cost of implementing high effort. In the case of motivational monitoring, the principal’s marginal net benefit is negative, so we can have a situation where the first best entails $e = 1$ and $a = 0$, whereas the second best entails a strictly mixed strategy where the principal works with positive probability. This is because the principal’s effort is motivational when there are team economies, so the principal’s effort acts like an incentive which reduces the moral hazard problem. As a result, the principal’s effort is oversupplied in equilibrium.
2.9 Commitment

We now turn to the commitment issue. We continue to assume that output $Q$ is contractible, so the principal can commit to the transfers $\tilde{t}$ and $\tilde{t}$ in an enforceable contract. Suppose the principal has arranged her external activities such that with probability $\alpha$ she will be available to work with the agent and with probability $1 - \alpha$ she will be completely unavailable. After accepting the contract, the agent chooses $e = 1$. The question is: after the principal’s mixed strategy has been realized, will she actually follow through on it? With probability $1 - \alpha$, the principal’s mixed strategy calls for her to not work with the agent but in this case there is no commitment problem because she cannot work even if she wanted to. But with probability $\alpha$, the principal’s mixed strategy calls for her to work with the agent even though it may not be in her best interest to do so. To make commitment feasible, we need to add an extra condition such that the principal finds it optimal to work in that state.

We first consider the case where (31) holds, so the principal engages in managerial interference. At the moment that the principal is deciding whether or not to follow through on her mixed strategy, which calls for her to work, the agent has already chosen $e = 1$ and the transfers $\tilde{t}$ and $\tilde{t}$ are contractually fixed. The principal must now choose $\gamma$, the probability that $a = 1$, where $\gamma$ need not be the same as the initial choice $\alpha$. The expected transfer is

$$E(t) = (\pi_p + \gamma \Delta \pi_p)\tilde{t}$$  \hspace{1cm} (37)

and expected profit is

$$\Pi = E(Q) - \gamma \psi P - (\pi_p + \gamma \Delta \pi_p)\tilde{t},$$  \hspace{1cm} (38)

where $E(Q)$ is defined in (28) with $\gamma$ instead of $\alpha$. At this point in the game, the principal’s re-evaluation of her effort choice can affect expected output, the expected opportunity cost of the principal’s effort, and the probability that the agent receives the transfer $\tilde{t}$, but not the transfer $\tilde{t}$ itself. Clearly, it is in the best interest of the principal to follow through on her strategy and choose $\gamma = 1$ when

$$\Delta \pi_p \Delta Q - \psi P - \Delta \pi P \tilde{t} > 0.$$  \hspace{1cm} (39)

Since (31) implies $\Delta \pi_A < \Delta \pi_{P,A}$,

$$\frac{\psi_A}{\Delta \pi_{P,A}} < \tilde{t} < \frac{\psi_A}{\Delta \pi_A}$$  \hspace{1cm} (40)
from (23). A sufficient condition is therefore
\[ \Delta \pi_P \Delta Q - \psi_P - \frac{\Delta \pi_P}{\Delta \pi_A} \psi_A > 0, \] (41)
which requires the principal’s marginal net benefit to be sufficiently large. We conclude that if (41) holds then the principal has the power to commit to a mixed strategy for effort by arranging her external activities appropriately. With probability \( \alpha \), the realization of that mixed strategy calls upon her to work with the agent and she finds it in her interest to do so, and with probability \( 1 - \alpha \) she cannot work with the agent because of her outside responsibilities.

Commitment seems more problematic for motivational monitoring because this class of mixed strategies involves stochastic monitoring when it is inefficient for the principal to work. When the realization of the mixed strategy calls for the principal to work, she will not want to follow through. We have assumed that the principal’s effort, like the agent’s effort, is non-contractible, so the principal cannot solve the problem by committing herself contractually. One possibility is that the principal may have an incentive to follow through on her stochastic monitoring strategy for reputational purposes.

2.10 Effects on Compensation

When the principal chooses a pure strategy, the optimal bonus (23) and expected transfer (25) depend only on the agent’s cost of effort \( \psi_A \) and his marginal product: \( \Delta \pi_{P,P_A} \) when \( \alpha = 0 \) and \( \Delta \pi_A \) when \( \alpha = 1 \). But when the principal strictly mixes \( 0 < \alpha < 1 \), the optimal bonus and expected transfer depend on the principal’s marginal net benefit of effort \( \Delta \pi_P \Delta Q - \psi_P \) and the relative informativeness of individual performance \( \Delta L \).

Proposition 2 When the principal uses a strictly mixed strategy \( 0 < \alpha < 1 \),
\[ \bar{t} = \left( \frac{\Delta \pi_P \Delta Q - \psi_P}{\pi \pi_P \Delta L} \psi_A \right)^{1/2}. \] (42)

(i) If (31) holds then \( \bar{t} \) and \( E(t) \) are strictly increasing in \( \Delta \pi_P \Delta Q - \psi_P \) and strictly decreasing in \( \Delta L \).

(ii) If (32) holds then \( \bar{t} \) and \( E(t) \) are strictly decreasing in \( \Delta \pi_P \Delta Q - \psi_P \) and strictly increasing in \( \Delta L \).

To obtain (42), we substitute the optimal \( \alpha \) from Proposition 1 into the expression for the optimal \( \bar{t} \) in (23). We first consider the case where (31) holds, which is the case of managerial
interference when there are team diseconomies. As we have seen, in this case an increase in the probability \( \alpha \) that the principal works necessitates an increase in the optimal bonus \( \bar{t} \) and expected transfer \( E(t) \) because of team diseconomies. From the comparative statics of managerial interference, the principal increases \( \alpha_{MI} \) when her marginal net benefit of effort \( \Delta \pi_p \Delta Q - \psi_p \) is higher and the relative informativeness of individual performance \( \Delta L \) is lower. The first result then follows. The discussion for motivational monitoring (32) is similar, except that now \( \Delta \pi_p \Delta Q - \psi_p \) and \( \Delta L \) are negative and the expected transfer \( E(t) \) is decreasing in \( \alpha \) because of team economies.

3 Individual and Group Performance

In the previous section, the principal could only observe output \( Q \). What if the principal can also observe a separate measure \( y \) of individual performance? In this section, we show that under standard assumptions the above results do not change much. We also develop notation that will be used in the next section.

As before, output \( Q \) can be either high \( \bar{Q} \) or low \( \underline{Q} \). Likewise, individual performance \( y \) can be either high \( \bar{y} \) or low \( \underline{y} \), where \( \bar{y} > \underline{y} \). There are now four states \((i, j)\), where \( i \in \{\bar{Q}, \underline{Q}\} \) and \( j \in \{\bar{y}, \underline{y}\} \). Let \( \pi^{ij} \) be the probability of state \((i, j)\) when the principal and agent both work, \( \pi^{ij}_A \) when only the principal works, \( \pi^{ij}_P \) when only the agent works, and \( \pi^{ij}_{PA} \) when neither work. We assume all these probabilities are non-zero.

Let

\[
\begin{align*}
p^{ij} &= \alpha \pi^{ij}_P + (1 - \alpha) \pi^{ij}_A \\
p^{ij}_A &= \alpha \pi^{ij}_A + (1 - \alpha) \pi^{ij}_{PA} \\
L^{ij}_\alpha &= \frac{p^{ij}_A}{p^{ij}}.
\end{align*}
\]

where \( p^{ij} \) is the probability of state \((i, j)\) when the agent works, \( p^{ij}_A \) is the probability of \((i, j)\) when the agent shirks, and \( L^{ij}_\alpha \) is the likelihood ratio. Given \( \alpha \), the informativeness of state \((i, j)\) is defined by

\[
1 - L^{ij}_\alpha = \frac{\Delta \pi^{ij}_{P,PA} + \alpha \left( \Delta \pi^{ij}_A - \Delta \pi^{ij}_{P,PA} \right)}{\pi^{ij}_P + \alpha \Delta \pi^{ij}_P},
\]

where \( \pi^{ij}_P \) and \( \pi^{ij}_A \) are the probabilities when the principal and agent work, respectively. The terms \( \Delta \pi^{ij}_{P,PA} \) and \( \Delta \pi^{ij}_A \) represent the differences in performance due to team diseconomies.
where

\[ \Delta \pi_{ij}^A = \pi_{ij}^A - \pi_{ij}^A \] (47)

\[ \Delta \pi_{ij}^P = \pi_{ij}^P - \pi_{ij}^P \] (48)

\[ \Delta \pi_{ij}^{P,P,A} = \pi_{ij}^{P} - \pi_{ij}^{P,A} \] (49)

Note that the denominator in (46) is positive but the numerator can be positive, negative, or zero depending on the effect of \(a\) and \(e\) on the probability of state \((i, j)\). We assume the probability of state \((Q, y)\) is increasing in \(a\) and \(e\), so that \(\Delta \pi_{Q}^C, \Delta \pi_{Q}^C, \text{ and } \Delta \pi_{P}^{C, C} \) are all positive. We do not need to make any specific assumptions for the other states.

We continue to assume that the principal wants to implement \(e = 1\). Let \(t_{ij} \geq 0\) be the transfer in state \((i, j)\). The incentive compatibility constraint is

\[
\alpha \left( \sum_{ij} \pi_{ij}^t t_{ij} \right) + (1 - \alpha) \left( \sum_{ij} \pi_{ij}^P t_{ij} \right) - \psi_A \geq 0
\] (50)

\[
\alpha \left( \sum_{ij} \pi_{ij}^A t_{ij} \right) + (1 - \alpha) \left( \sum_{ij} \pi_{ij}^{P,A} t_{ij} \right)
\] (51)

where \(\sum_{ij}\) represents the sum over all four states. The above can be re-written as

\[
\sum_{ij} \left[ \alpha \Delta \pi_{ij}^A + (1 - \alpha) \Delta \pi_{ij}^{P,P,A} \right] t_{ij} - \psi_A \geq 0,
\] (52)

where the term in square brackets is the expected change in the probability of state \((i, j)\) when the agent works. As before, the participation constraint is redundant.

Given \(e = 1\), the expected transfer is

\[
E(t) = \sum_{ij} \left[ \alpha \pi_{ij} + (1 - \alpha) \pi_{ij}^P \right] t_{ij}.
\] (53)

We first find the transfers \(\{t_{ij}\}\) that minimize \(E(t)\) subject to \(t_{ij} \geq 0\) for all states and the incentive compatibility constraint (52).

**Proposition 3** Assume that \((Q, y)\) is the most informative state for all \(0 \leq \alpha \leq 1\) in the sense
that $1 - L^ij_\alpha$ is highest in that state. In that case, the optimal transfers are

$$t_{ij} = \frac{\psi_A}{\Delta \pi_{P,P,A} + \alpha (\Delta \pi_A - \Delta \pi_{P,P,A})}$$

and $t_{ij} = 0$ for all other states.

A standard assumption which ensures that $(Q, y)$ is the most informative state is the \textit{monotone likelihood ratio property}, which states that $L^ij_\alpha$ is strictly decreasing in $i$ and $j$, but this assumption is stronger than we need. As in the previous section, the principal only rewards the most informative state. The optimal incentive $t_{ij}$ is then found from the binding incentive compatibility constraint.

The expected transfer is

$$E(t) = \frac{\pi_{P,P,A} + \alpha \Delta \pi_{P,A}}{\Delta \pi_{P,P,A} + \alpha (\Delta \pi_A - \Delta \pi_{P,P,A})} \psi_A = \frac{1}{1 - L^ij_\alpha} \psi_A$$

This is the same as $E(t)$ in the previous section except that here we have the superscript $(Q, y)$ rather than the implicit superscript $Q$. Since the two expressions are essentially the same,

$$\frac{\partial E(t)}{\partial \alpha} = \frac{\pi_{P,P,A} \Delta L}{[\Delta \pi_{P,P,A} + \alpha (\Delta \pi_A - \Delta \pi_{P,P,A})]^2} \psi_A,$$

where

$$\Delta L = L^ij_1 - L^ij_0.$$  

If $\Delta L > 0$ then the state $(Q, y)$ is more informative when the principal does not work $a = 0$ than when she does $a = 1$.

Expected output is given by

$$E(Q) = \alpha \left[ \pi_{P} \bar{Q} + \left(1 - \pi_{P} \right) \bar{Q}\right] + (1 - \alpha) \left[ \pi_{P} \bar{Q} + \left(1 - \pi_{P} \right) \bar{Q}\right]$$

$$= Q + \left( \pi_{P} + \alpha \Delta \pi_{P} \right) \Delta Q,$$
where

\[
\begin{align*}
\pi Q &= \pi Q_P + \pi Q_y \\
\pi P &= \pi P_P + \pi P_y \\
\Delta \pi P &= \pi Q - \pi Q_P.
\end{align*}
\] (60)

The expected profit of the principal is

\[\Pi = E(Q) - \alpha \psi P - E(t) \] (63)

with derivative

\[
\frac{\partial \Pi}{\partial \alpha} = \Delta \pi Q \Delta Q - \psi P - \frac{\partial E(t)}{\partial \alpha}.
\] (64)

If we compare (56) and (64) with their counterparts (26) and (30) in the previous section, we see that the structure of the two models is exactly the same. The only difference is that in the previous section, the principal was concerned about the realization of output \(Q\) while the agent was concerned about the realization of performance \(Q\), which were one and the same. In the present context, the principal is still concerned about the realization of output \(Q\), whereas the agent is concerned about the realization of performance \((Q, y)\). The probabilities in (56) refer to \((Q, y)\) because that is the state where the agent is rewarded. All the results are therefore exactly the same as before except for this difference in interpretation.

### 4 Conditional Observation of Individual Performance

We now consider the case where the principal can only observe individual performance \(y\) when she works with the agent. When \(a = 0\), the principal is not involved in the agent’s activities and can therefore only observe the outcome \(Q\), which is either high \(Q\) or low \(Q\) with the probabilities given in section 2. When \(a = 1\), the fact that the principal actively participates in the production process alongside the agent (e.g., working on a sales presentation) allows her to observe both the outcome \(Q\) and individual performance \(y\). As before, \(y\) can be high \(\bar{y}\) or low \(\underline{y}\) and the joint signal \((Q, y)\) is governed by the probabilities given in the previous section.
Since the principal wants to implement $e = 1$, the incentive compatibility constraint is

$$\alpha \left( \sum_{ij} \pi_{ij}^j t_{ij} \right) + (1 - \alpha) \left[ \pi_P \bar{t} + (1 - \pi_P) \underline{t} \right] - \psi_A \geq 0 \quad (65)$$

$$\alpha \left( \sum_{ij} \pi_{ij}^t t_{ij} \right) + (1 - \alpha) \left[ \pi_{PA} \bar{t} + (1 - \pi_{PA}) \underline{t} \right]. \quad (66)$$

With probability $\alpha$, the principal works and there are four possible states $(i, j)$, where $i \in \{\overline{Q}, Q\}$ and $j \in \{\overline{y}, y\}$. With probability $1 - \alpha$, the principal does not work and there are just two states \{\overline{Q}, Q\}. We can re-write the incentive compatibility constraint as

$$\alpha \sum_{ij} \Delta \pi_{ij}^j t_{ij} + (1 - \alpha) \Delta \pi_{P,PA}^t - \psi_A \geq 0 \quad (67)$$

using notation defined previously. The expected transfer is

$$E(t) = \alpha \left( \sum_{ij} \pi_{ij}^j t_{ij} \right) + (1 - \alpha) (\bar{t} + \pi_P \Delta t). \quad (68)$$

Given $\alpha$, the principal chooses the transfers \{\(t_{ij}\), \(\bar{t}\), and \(\underline{t}\)\} to minimize the expected transfer $E(t)$ subject to the incentive compatibility constraint (67). The proof of the following result is similar to previous ones and is omitted.

**Lemma 5**

(i) If $\alpha = 0$ the optimal transfers are $\underline{t} = 0$ and $\bar{t} = \frac{\psi_A}{\Delta \pi_{P,PA}}$.

(ii) If $\alpha = 1$ and $(\overline{Q}, \overline{y})$ is the most informative of the four possible states then the optimal transfers are $t_{\overline{Q} \overline{y}} = \frac{\psi_A}{\Delta \pi_{\overline{Q} \overline{y}}} \Delta t$ and $t_{ij} = 0$ for the other three states.

(iii) If $0 < \alpha < 1$ the optimal transfers depend on which of the six possible states is the most informative:

(a) If $(\overline{Q}, \overline{y})$ is the most informative state then the optimal transfers are $t_{\overline{Q} \overline{y}} = \frac{\psi_A}{\alpha \Delta \pi_{\overline{Q} \overline{y}}} \Delta t$ and zero in the other five possible states.

(b) If $\overline{Q}$ is the most informative state then the optimal transfers are $\bar{t} = \frac{\psi_A}{(1 - \alpha) \Delta \pi_{P,PA}}$ and zero in the other five states.
As before, the principal only rewards the most informative state and the optimal bonus paid in that state comes from the binding incentive compatibility constraint (67) with all other transfers equal to zero. What is different is that in this section, the set of possible states depends on the principal’s choice of \( \alpha \). In (i), the principal does not work for sure \( \alpha = 0 \) and the set of possible states is \( \{Q, \bar{Q}\} \). The optimal bonus \( t \) in state \( \bar{Q} \) is the same as the textbook bonus (24) in section 2. In (ii), the principal works for sure \( \alpha = 1 \) and there are four possible states \( \{(i, j)\} \), where \( i \in \{Q, \bar{Q}\} \) and \( j \in \{y, \bar{y}\} \). We continue to assume that \( (Q, \bar{y}) \) is the most informative state of these four and the optimal bonus \( t_{Q\bar{y}} \) for that state is the same as (54) in the previous section with \( \alpha = 1 \). In (iii), the principal chooses a strictly mixed strategy \( 0 < \alpha < 1 \) and all six states are possible. If the state is \( \bar{Q} \), it means that output was high when the agent worked alone. If the state is \( (Q, \bar{y}) \), it means that both individual and group performance were high when the principal worked. Either of these two states can be the most informative, depending on the parameters, and in (iii) we allow for both cases. The principal rewards exactly one of these two states, whichever is the most informative, with a zero transfer in all other states.

Notice that when the principal strictly mixes \( 0 < \alpha < 1 \) and \( (Q, \bar{y}) \) is the most informative state, the agent’s expected marginal product of effort is \( \alpha \Delta \pi_A \bar{Q}\bar{y} \), so efforts are complements in the incentive compatibility constraint (67). In this case, the optimal bonus \( t_{Q\bar{y}} \) is strictly decreasing in \( \alpha \). On the other hand, when \( \bar{Q} \) is the most informative state, the agent’s expected marginal product of effort is \( (1 - \alpha) \Delta \pi_{P,P_A} \) and efforts are substitutes. In this case, the optimal bonus \( \bar{t} \) is strictly increasing in \( \alpha \).

We now find the expected transfer in each case. If \( \alpha = 0 \), the agent receives \( \bar{t} \) in (i) with probability \( \pi_P \), so the expected transfer is

\[
E(t) = \frac{\pi_P}{\Delta \pi_{P,P_A}} \psi_A. \tag{69}
\]

If the principal strictly mixes \( 0 < \alpha < 1 \) and \( \bar{Q} \) is the most informative state, the agent receives \( \bar{t} \) in (iiib) with probability \( (1 - \alpha)\pi_P \), so the expected transfer is again (69). If \( \alpha = 1 \), the agent receives \( t_{Q\bar{y}} \) in (ii) with probability \( \pi_{Q\bar{y}} \), so the expected transfer is

\[
E(t) = \frac{\pi_{Q\bar{y}}}{\Delta \pi_A} \psi_A. \tag{70}
\]

This is also the expected transfer when \( 0 < \alpha < 1 \) and \( (Q, \bar{y}) \) is the most informative state because the agent receives \( t_{Q\bar{y}} \) in (iiia) with probability \( \alpha \pi_{Q\bar{y}} \).
The crucial point is that the expected transfer \( E(t) \) does not depend on \( \alpha \) for all \( 0 \leq \alpha \leq 1 \). This is because the principal’s choice of \( \alpha \) is no longer informative. To illustrate, assume that \( \overline{Q} \) is the most informative among the six possible states when the principal chooses a strictly mixed strategy \( 0 < \alpha < 1 \). Recall that this state \( \overline{Q} \) only occurs when the principal does not work. The probability of \( \overline{Q} \) when the agent works is \( p_{\overline{Q}} = (1 - \alpha)\pi_P \) and the probability of \( \overline{Q} \) when the agent does not work is \( p_{A\overline{Q}} = (1 - \alpha)\pi_{PA} \), so the likelihood ratio is

\[
L_{\alpha} = \frac{p_{A\overline{Q}}}{p_{\overline{Q}}} = \frac{\pi_{PA}}{\pi_P}.
\]

(71)

The principal’s choice of \( \alpha \) therefore has no informational consequences. In contrast, in section 2 the state \( \overline{Q} \) occurred with probability \( \pi_P \) when \( \alpha = 0 \) and probability \( \pi \) when \( \alpha = 1 \). In other words, the state \( \overline{Q} \) always occurred with positive probability, albeit with different probabilities, so changes in \( \alpha \) could be potentially informative. In fact, an increase in \( \alpha \) moved the likelihood ratio away from \( \pi_{PA}/\pi_P \) towards \( \pi_A/\pi \), which improves information when there are team economies and and garbles it when there are team diseconomies.

Expected output is given by

\[
E(Q) = \alpha \left[ \pi_{\overline{Q}} \overline{Q} + (1 - \pi_{\overline{Q}}) Q \right] + (1 - \alpha) \left[ \pi_P \overline{Q} + (1 - \pi_P) Q \right]
\]

(72)

using notation from previous sections. Expected profit is given by

\[
\Pi = E(Q) - \alpha \psi_P - E(t).
\]

(73)

The derivative of \( E(Q) - \alpha \psi_P \) with respect to \( \alpha \) is

\[
\left( \pi_{\overline{Q}} - \pi_{P} \right) \Delta Q - \psi_P,
\]

(74)

which is constant.

**Proposition 4** Generically, the principal never mixes: the optimal \( \alpha \) is either zero, one, or does not exist.

To explain the result, consider the case where \( \overline{Q} \) is more informative than \( (\overline{Q}, \overline{y}) \) when the
principal strictly mixes. When $0 \leq \alpha < 1$ the expected profit of the principal is

$$\Pi = E(Q) - \alpha \psi_P - \frac{\pi_P}{\Delta \pi_{P,P,A}} \psi_A. \quad (75)$$

When $\alpha = 1$,

$$\Pi = E(Q) - \alpha \psi_P - \frac{\pi_{Q,Y}}{\Delta \pi_{Q,Y}} \psi_A, \quad (76)$$

where the expected transfer in (69) is lower than the expected transfer in (70). This is because $Q$ is more informative than $(Q, y)$, so the agent receives less information rent when $\alpha = 0$ and the state is $Q$ than when $\alpha = 1$ and the state is $(Q, y)$. If the derivative with respect to $\alpha$ in (74) is negative, then expected profit $\Pi$ declines linearly, with a discontinuous drop down at $\alpha = 1$, so the solution is $\alpha = 0$. If (74) is positive then expected profit increases linearly in $\alpha$ with a discontinuous jump down at $\alpha = 1$, so there is no optimal $\alpha$. In the non-generic case where (74) is zero, the principal is indifferent between all values of $\alpha$ such that $0 \leq \alpha < 1$. The discussion when $(Q, y)$ is more informative than $Q$ when $0 \leq \alpha < 1$ is similar.

We draw two conclusions. The first is the practical conclusion that the principal never mixes when she observes different signals when she works. The second is that the principal never mixes unless mixing is informative. In particular, the fact that efforts are complements or substitutes is not a sufficient reason for mixing.

## 5 Conclusion

In this paper, we extended the elementary textbook model where the principal and agent are both risk neutral and the agent has limited liability to the case where the principal makes an optimal effort decision. We assumed that the principal can commit to a mixed strategy in effort before the agent chooses his own effort, so the principal’s effort is not directly subject to moral hazard. The principal’s effort decision depends on three factors: the marginal net benefit of her effort, its effect on the informativeness of output, and whether efforts are complements or substitutes. We defined team diseconomies as the case where efforts are substitutes and individual performance is more informative than group performance and team economies as the case where efforts are complements and group performance is more informative.

We showed that when there are team diseconomies and the above three factors are evenly balanced, the optimal mixed strategy is managerial interference, where the principal works with positive probability because it is efficient for her to work in a first best world, despite the fact that
it reduces the agent’s expected marginal product of effort and the informativeness of output in the second best world. As a result, the principal offers stronger incentives, a higher expected transfer, and can undersupply effort relative to the first best. When there are team economies and the above three factors are evenly balanced, the optimal mixed strategy is *motivational monitoring*, where the principal works with positive probability, even though it is inefficient, because it increases the agent’s expected marginal product of effort and the informativeness of output. In this case, the principal offers weaker incentives, a lower expected transfer, and can oversupply effort relative to the first best. Almost the same results hold when the principal also observes a separate signal of individual performance. We then showed that the principal never mixes when she only receives a separate signal of individual performance when she works because in that case her mixed strategy is uninformative. This shows that complementarities in efforts are not a sufficient condition for mixing, which instead hinges on informational considerations.

6 Appendix

**Proof of Lemma 1.** We allow both the principal and the agent to use mixed strategies. Let \( \alpha \) be the probability that \( a = 1 \) and \( \beta \) the probability that \( e = 1 \). Let

\[
q(\alpha, \beta) = \alpha \beta \pi + \alpha (1 - \beta) \pi_A + (1 - \alpha) \beta \pi_P + (1 - \alpha) (1 - \beta) \pi_{PA}.
\]  

(77)

The principal chooses \( \alpha \) and \( \beta \) on the square with vertices \{ (0, 0), (1, 0), (0, 1), (1, 1) \} to maximize expected surplus

\[
q(\alpha, \beta) \bar{Q} + [1 - q(\alpha, \beta)] \underline{Q} - \alpha \psi_P - \beta \psi_A.
\]  

(78)

We now compare

\[
\pi \bar{Q} + (1 - \pi) \underline{Q} - \psi_P - \psi_A
\]  

(79)

\[
\pi_P \bar{Q} + (1 - \pi_P) \underline{Q} - \psi_A
\]  

(80)

\[
\pi_A \bar{Q} + (1 - \pi_A) \underline{Q} - \psi_P
\]  

(81)

\[
\pi_{PA} \bar{Q} + (1 - \pi_{PA}) \underline{Q}
\]  

(82)

where (79) is expected profit when \( a = e = 1 \), (80) when \( a = 0 \) and \( e = 1 \), (81) when \( a = 1 \) and \( e = 0 \), and (82) when \( a = e = 0 \). Under the stated conditions, (79) is the largest so the principal will not mix between (79) and the others. ■
Proof of Lemma 2. We prove (i). If $\Delta \pi < \Delta \pi P,PA$ then $\pi$ is strictly submodular on the lattice

$$
\{(0,0),(1,0),(0,1),(1,1)\}.
$$

Lemma 2.6.6 in Topkis (1998) states that if $f(x)$ is positive, increasing, and submodular then it is log-submodular. Trivial adjustments to the proof show that when $f(x)$ is positive, strictly increasing, and strictly submodular then it is strictly log-submodular. Since $\pi > 0$ and strictly increasing in $(a,e)$, we conclude that $\pi$ is strictly log-submodular, which is equivalent to $\Delta L > 0$. The proof of (ii) is similar. ■

Proof of Lemma 4. Assume $\Delta \pi A < \Delta \pi P,PA$. It follows that $\Delta L > 0$ and (26) is positive and decreasing in $\alpha$, so $\Pi$ is strictly concave in $\alpha$. Now assume $\Delta L < 0$. It follows that $\Delta \pi A > \Delta \pi P,PA$ and (26) is negative and increasing in $\alpha$, so $\Pi$ is strictly concave in $\alpha$. If $\Delta \pi A > \Delta \pi P,PA$ and $\Delta L > 0$ then (26) is positive and decreasing in $\alpha$, so $\Pi$ is strictly convex in $\alpha$. If $\Delta L = 0$ and/or $\Delta \pi A = \Delta \pi P,PA$ then $\Pi$ is linear. ■

Proof of Proposition 1. Assume $0 < \alpha < 1$. According to lemma 4, $\Pi$ is either linear, strictly convex, or strictly concave. The first two cases imply a corner solution, so $\Pi$ must be strictly concave. It follows that either $\Delta \pi A < \Delta \pi P,PA$ or $\Delta L < 0$. Assume the first case, so $\Delta \pi A < \Delta \pi P,PA$ and $\Delta L > 0$. Since $\Pi$ is strictly concave, the first-order conditions for $\alpha$ are necessary and sufficient. Since the solution is interior, we must have

$$
\frac{\partial \Pi}{\partial \alpha} \bigg|_{\alpha=0} = \Delta \pi P\Delta Q - \psi_P - \frac{\pi \pi P \Delta L}{(\Delta \pi P,PA)^2} \psi_A > 0
$$

and

$$
\frac{\partial \Pi}{\partial \alpha} \bigg|_{\alpha=1} = \Delta \pi P\Delta Q - \psi_P - \frac{\pi \pi P \Delta L}{(\Delta \pi A)^2} \psi_A < 0.
$$

Combining these two inequalities gives (31). In the second case, $\Delta L < 0$ and $\Delta \pi A > \Delta \pi P,PA$. Since $\Delta L < 0$, the same calculations produce (32) instead of (31). We now prove the reverse implication. Assume that either (31) or $\Delta L < 0$ and (32) hold. In the first case, (31) implies $\Delta \pi A < \Delta \pi P,PA$, so $\Pi$ is strictly concave. The first-order conditions are therefore necessary and sufficient and (31) rules out both corner solutions. The solution must therefore be interior which implies a strictly mixed strategy. The proof for the second case is similar. ■
Proof of Proposition 3. The Lagrangean is

$$\mathcal{L} = -E(t) + \lambda \left\{ \sum_{ij} \left[ \alpha \Delta \pi_{ij}^{A} + (1 - \alpha) \Delta \pi_{ij}^{P,P_A} \right] t_{ij} - \psi_A \right\}.$$  \hspace{1cm} (86)

The first-order conditions are

$$\mathcal{L}_{t_{ij}} = - \left[ \alpha \pi_{ij}^{A} + (1 - \alpha) \pi_{ij}^{P} \right] + \lambda \left[ \alpha \Delta \pi_{ij}^{A} + (1 - \alpha) \Delta \pi_{ij}^{P,P_A} \right] \leq 0,$$  \hspace{1cm} (87)

along with the incentive compatibility constraint and the complementary slackness conditions. Since the objective function and constraint are both linear in the transfers, the first-order conditions are necessary and sufficient. If $\lambda = 0$ then $\mathcal{L}_{t_{ij}} < 0$ for all states and therefore $t_{ij} = 0$ for all states by complementary slackness. This violates incentive compatibility, so $\lambda > 0$ and the incentive compatibility constraint binds. Next, we observe that $t_{ij} = 0$ for any state $(i, j)$ such that

$$\alpha \Delta \pi_{ij}^{A} + (1 - \alpha) \Delta \pi_{ij}^{P,P_A} \leq 0$$  \hspace{1cm} (88)

by complementary slackness. We now consider states $(i, j)$ such that $1 - L_{ij}^A > 0$. We can re-write (87) as

$$\lambda \leq \frac{1}{1 - L_{ij}^A}. $$  \hspace{1cm} (89)

Since all transfers equal to zero violates incentive compatibility, at least one of these inequalities must hold as an equality. Since the state $(\bar{Q}, \bar{y})$ is the most informative, it follows that

$$\lambda = \frac{1}{1 - L_{\bar{Q}\bar{y}}}$$  \hspace{1cm} (90)

and all the transfers are zero except the transfer in state $(\bar{Q}, \bar{y})$. Finally, we obtain (52) from the binding incentive compatibility constraint. \hfill \blacksquare