Individual accountability in teams

Leslie M. Marx [a,*], Francesco Squintani [b,c]

[a] *Fuqua School of Business, Box 90120, Duke University, Durham, NC 27708, United States*
[b] *Department of Economics, University of Essex, Wivenhoe Park, Colchester, Essex CO4 3SQ, United Kingdom*
[c] *Universita' degli Studi di Brescia, Dipartimento di Scienze Economiche, Via San Faustino 74/B, 25122 Brescia, Italy*

ARTICLE INFO

ABSTRACT

We consider a model of team production in which the principal observes only the team output, but agents can monitor one another (at a cost) and provide reports to the principal. We consider the problem faced by a principal who is prevented from penalizing an agent without evidence showing that the agent failed to complete his assigned actions. We show the first-best (high effort but no monitoring) can be achieved, but only if the principal assigns second-best actions. The principal requires monitoring, but agents do not monitor, and as long as output is high, the principal does not penalize agents who fail to monitor. If the principal has the responsibility for monitoring, the first-best outcome cannot be achieved, thus we identify an incentive for delegated monitoring even when agents have no informational advantage.

1. Introduction

In many business settings, supervisors ask that employees evaluate the performance of their peers, especially when the employees work as a team.[1] If taken seriously, such evaluations can be costly to produce, but our results suggest that even if they are taken lightly, they can be a valuable part of workers' incentives to exert productive effort. For example, at Intel Corporation, a new microprocessor is divided into "blocks" and a team of people is assigned to complete each block. Intel gives the team a great deal of flexibility as to how to organize work on the block and relies on team members to monitor each other. In the event that a block fails to perform as specified, Intel can turn to team member evaluations to identify unproductive employees. However, if the team performance is good, it need not overemphasize the reviews.

Although there is much discussion in the literature of the peer review process, there is remarkably little discussion of the monitoring of peer reviewers, i.e., monitoring the monitors. However, it does appear to be possible to do so. Edwards and Ewen (1996, p. 46) discuss measures for assessing the rigor of evaluations and identifying invalid reviews. For example,

---

* Corresponding author. Tel.: +1 919 660 7762; fax: +1 919 660 7971.
  E-mail addresses: marx@duke.edu (L.M. Marx), squint@essex.ac.uk (F. Squintani).

[1] In settings with team production, peer monitoring may be the only feasible option. According to May and Gueldenzoph (2006, p. 4), "Intragroup peer evaluations used for performance appraisal and compensation decisions are becoming an integral part of today's team-based organizations." They comment that "peers may be the only ones who can provide relevant information about group outcomes." (pp. 4–5) In a similar vein, Fedor et al. (1999, p. 93) state, "When it comes to team-based work situations, peers may be the only ones who can provide relevant information on an employee's contributions to group outcomes." Edwards and Ewen (1996, p. 45) state, "Units of AT&T, AlliedSignal, DuPont, Honeywell, Monsanto, Boeing, Intel, and Meridian Oil have migrated from developmental only 360-degree feedback to the use of multisource assessment to support both development and performance management (i.e., appraisal and pay)."

they say reviews can be examined for the "appropriateness of the respondent's use of the rating scale and the consistency of her or his responses with those of the other participants." As we show in this paper, a de-emphasis on following through on the monitoring of peer reviewers may actually be efficiency enhancing. Specifically, we show that in some environments the principal can induce high effort by assigning team members the task of monitoring one another but then not penalizing a failure to monitor as long as the team output is high.

In team problems, the output depends on all agents' choices, but it may be difficult for a principal to identify the actions taken by each individual agent.[2] Despite the team nature of output, it may not be possible for a principal to penalize and reward team members as a group. Joint penalties based on team rather than individual work may be illegal or ruled out by company or union policies.[3] Even when feasible, a system of joint penalties and rewards may be more expensive or less effective than a compensation system based on individual effort.[4]

In this paper, we consider an environment with team production in which the principal is unable to monitor the efforts of agents, although the principal can assign agents the task of monitoring each other. We assume that the principal is unable to penalize an agent for poor team output unless she can prove the agent shirked. We develop our main results using the assumption that peer monitoring always reveals an agent to have shirked if in fact that agent did shirk. However, we assume that if output is low and an agent fails to produce evidence that a teammate shirked, the principal cannot discern whether that is because the teammate did not shirk or because the agent failed to engage in proper monitoring.

In modeling the monitoring technology available to the agents, we attempt to connect with real-world institutions for enforcing contracts. We provide a discussion of how the monitoring technologies we consider can arise through judicial institutions, showing they can be derived from standard assumptions regarding the functioning of a judicial body. In this respect our approach is similar to that of Feess and Hege (1998), who also explicitly consider the evidence that a court might have available to it for determining liability when team production is involved.[5]

Because we assume the activity of monitoring is costly, the first-best outcome involves high effort by the agents but no monitoring. The first-best outcome cannot be achieved if the principal does not include monitoring as part of the assignment because without the threat of monitoring, agents choose low effort. The first-best outcome also cannot be achieved if the principal monitors the agents herself because in that case high effort can only be induced if the principal commits to monitor with positive probability. However, despite these negative results, we show that the first-best outcome can be achieved.

In our main result, we show that it is possible to overcome the inefficiency associated with this team problem by specifying in the agents' job assignments that each member should work hard in the joint production and monitor and report the actions of his teammates. Even though equilibrium job assignments include monitoring, this serves only to provide incentives for effort—in equilibrium agents do not engage in wasteful monitoring. The agents violate the principal's assignment in equilibrium, but in violating it, they allow the first-best to be achieved. In this environment, the first-best can only be achieved by assigning second-best actions.

Even though team-level penalties and rewards are not considered, the inclusion of the monitoring requirement in the agents' job assignments effectively makes each agent responsible for the actions of all other agents, and hence reinforces incentives to work hard. Suppose that the principal instructs agents to work hard and to monitor one another. As long as output is high, there is no evidence that any agent failed to monitor. But if output is low and an agent is unable to produce evidence of shirking by others, then it is verifiable (i.e., it can be determined by an outside authority) that the agent either shirked himself or failed to blow the whistle on the shirking agents. Because this verifies that he did not fulfill his assignment, the agent can be penalized. This makes each agent unwilling to shirk in providing productive effort. In contrast, if the principal does not include monitoring in the agents' assignments, then if output is low, the principal cannot verify whether any given agent failed to fulfill his assignment.

In our second result, we consider the alternative of (non-delegated) monitoring by the principal. If the principal assumes the responsibility of monitoring herself, then the principal must commit herself to monitor with positive probability in order to provide incentives for agents to exert effort. Because monitoring is costly, the first-best cannot be achieved when the principal has responsibility for monitoring. Thus, in addition to other reasons for delegated monitoring of teams, such as the need for technical expertise or informational advantages, we identify an efficiency enhancing incentive effect from the delegation of monitoring that complements these.[6]

---

[2] Inefficiencies may arise when the principal signs bilateral contracts with each agent separately. See for example Segal (1999, 2003) and Prat and Rustichini (2003).

[3] Joint penalties may occur, for example, when projects are cancelled for poor performance or unions accept pay cuts for companies in financial distress. However, we later detail institutional features and trade union agreements that make joint penalties infeasible. Also, employers may resist joint penalties because they are felt to be unfair by individual workers and so hurt morale and productivity. See Bewley (1999, especially Chapters 3 and 8) on the relation between pay cuts and morale.

[4] While some companies use joint bonuses (see, e.g., Che, 2002), many companies explicitly refrain from using joint bonuses based on team performance. Even in high-tech companies such as Intel and investment banks such as Goldman Sachs (see Goldman Sachs Proxy Statement, April 1, 2003, p. 32), bonuses are based on individual performance *relative to the team's average*, possibly depending also on the company's aggregate performance. Hence they can be interpreted as tournament mechanisms, possibly subject to budget constraints, but not as joint team bonuses.

[5] Feess and Hege (1998) show the efficiency of insurance-based liability rules in a model with noisy monitoring.

[6] In our model, the agents have no informational advantage over the principal in terms of identifying the efficient actions, so the strategic information transmission and incentives for delegation in Crawford and Sobel (1982); Dessein (2002) do not arise, nor does the informational rationale for delegation in Aghion and Tirole (1997).

As in Segal (1999, 2003) and Prat and Rustichini (2003), we consider an environment in which the principal signs bilateral contracts with each agent separately.[7] Since the incentives provided to an agent to complete his job assignment cannot depend directly on the choices of another agent, each agent can be held accountable only for his own actions.[8] There are a number of reasons for focusing on this environment. Joint penalties may be illegal because demotion and wage penalties are often subject to individual accountability. There are many examples in which supervisors find it difficult or impossible to sanction or dismiss an employee unless they can provide specific evidence that the individual did not fulfill his or her job assignment.[9] Union contracts commonly contain a "fire for cause agreement" that prevents the employer from firing a union member without specific cause. For example, firms employing members of the United Auto Workers are severely restricted in their ability to sanction union members.[10] Many universities and government agencies require extensive reporting and evidence before a staff member can be dismissed.[11] Italian labor law requires that in order to sanction or fire an employee, the employer must show that the employee has failed to fulfill a job assignment or violated the firm's regulations.[12] In these cases, the burden falls on the employer to prove that an individual has violated the rules before penalties can be invoked.

In contrast to joint penalties, another approach to motivating agents in a team environment is joint bonuses. However, joint bonuses can be expensive for firms because base compensation is often inflexible, and/or because joint bonuses contain an arbitrary or stochastic element, making it more expensive to compensate risk averse agents in this way. For example, union contracts and management policies often make base wages non-flexible, leaving no room for employers to motivate their employees with joint bonuses.[13] Of course, the incentive scheme that we identify is not affected by base wage inflexibility—when an agent is found shirking his individual assignment, he may be legally penalized with a salary detraction, and he may even be demoted or fired. Since commonly used instruments for bonuses, such as stock options, depend on the conduct of the entire firm, it may be difficult to tie joint bonuses to the performance of a small team, making them inefficient incentive devices for risk averse agents. To the extent that bonuses are risk free or technically equivalent to wages, one would expect them to be legally treated as wages and subject to the restrictions discussed above.

The remainder of the paper proceeds as follows: We describe the model in Section 2, and give our main results in Section 3. In Section 4 we extend our model to allow other monitoring technologies and teams of more than two agents. We review related literature in Section 5. Section 6 concludes.

2. Model

In this section, we present a model with one principal and two agents. In Section 4 we extend our results to allow teams of more than two agents. As the equilibrium concept, we use subgame perfect Nash equilibrium.

Each agent i chooses an action $a_i \in A = \{0, 1\}$ in a joint production process: either he exerts effort (action 1) or shirks (action 0). Let c_i be the cost to agent i of exerting effort and normalize the cost of shirking to zero. The output of the joint production process $\pi(a_1, a_2)$ depends on both agents' actions and may be either high ($\pi(a_1, a_2) = h$) or low ($\pi(a_1, a_2) = \ell$),

---

[7] In Segal (1999, 2003) and Prat and Rustichini (2003), even though the principals care about the choices of multiple agents, each principal enters an individual contracts with each agent, where the contract depends only on the actions of that agent.

[8] This is equivalent to assuming that an individual can be held liable for a contract violation only if he is identified as the party breaching the contract. While individual liability is usually the default rule in contract and labor law, joint and several liability may apply in some special cases. For instance, the Universal Commercial Code 3–116, regulating "negotiable instruments" (i.e. unconditional promises or orders to pay a fixed amount of money) states that "Except as otherwise provided in the instrument, two or more persons who have the same liability on an instrument [. . .] are jointly and severally liable [. . .]."

[9] In several countries of continental Europe, government-sponsored negotiations between trade unions and associations of employers (representing many corporations at the same time) prohibit individuals from being fired without elaborate due process. These centralized contracts are not limited to blue-collar jobs, but also control many white-collar industry jobs.

[10] In "Agreements between UAW® and the Ford Motor Company" (Vol. I, agreements Dated October 9, 1999 (effective Oct. 25, 1999)), Ford is severely restricted in its ability to penalize an employee who is a member of the United Auto Workers. Any action taken by Ford against a member of the UAW can be appealed through a so-called Grievance Procedure (Art. VII, Section 5). As a result, Ford cannot sanction shirking behavior unless it can establish individual guilt.

[11] The University of Rochester personnel policies and procedures on termination (Policy 136) allow termination based on poor work performance only if training and counseling fail. Moreover, the University cannot begin the procedures to terminate an employee without presenting its case to the Office of Human Resources.

[12] Paraphrased in English, the Italian Law (Statuto dei Lavoratori, Legge 20 maggio 1970 n. 300, Art. 4, 7, and 18) states that: (i) it is forbidden to use cameras, microphones, or other any devices to monitor workers' activity; (ii) the employer must make public what constitutes an infraction and the associated sanction; (iii) the employer may not sanction any worker without formally claiming an infraction, presenting the employee with the infraction, and hearing his defense; and (iv) if sanctioned, the worker may go to trial or request arbitration, in which case the sanction is suspended until the end of the trial or arbitration.

[13] Bewley (1999) finds that wages are downwardly rigid, except possibly for the secondary sector. But even for the secondary sector, minimum wage requirements may reduce wage flexibility. At Duke University, for each job category (e.g., bus driver, bookshop assistant, roofing mechanic), there is a minimum hourly wage and a maximum hourly wage. The lowest-paid Duke employee makes $6.18/h. Roofing mechanics make between $13.05 and $18.00/h. Bus drivers make between $10.97 and $16.02/h. Bookshop assistants make between $9.82/h and $13.75/h. There is not much room for variation within a pay grade, although an employee could be promoted, say from roofing mechanic to senior roofing mechanic, and then be able to earn more. Employees can be fired, but only after a detailed accounting has been made of their failings.

where we normalize $\ell = 0$.[14] We assume that

$$\ell = \pi(0, 0) = \pi(0, 1) = \pi(1, 0) < \pi(1, 1) = h,$$

so that the output is low if one or both of the agents shirk. We assume that $h > c_1 + c_2$ so that the efficient outcome is for both agents to exert effort.

While the output π is verifiable, and so can be observed by contract enforcement authorities, the agents' efforts a_1 and a_2 cannot be observed. However, at cost x_i, agent i can monitor agent j ($j \neq i$) and produce verifiable evidence R_i of j's effort. As mentioned in Section 1 , even if the principal has the ability to monitor the agents directly, she prefers to delegate because if the principal takes on the monitoring responsibility herself, she must commit to monitor with positive probability in order for the agents to have an incentive to exert effort. As we show, by delegating monitoring, the costs of monitoring can be avoided altogether.

We let $m_i = 1$ if agent i monitors agent j and $m_i = 0$ if he does not. We assume the agent's choice of monitoring is also a private action, so that it cannot be directly verified.

For any two agents i and j, agent i produces report $R_i(m_i, a_j)$ on agent j's action, which depends on i's monitoring effort and j's action. There are a variety of possibilities for the inferences that can be made from an agent's monitoring report. To identify the different possibilities, note that there are two possible types of error. In a type-one error, agent j is identified as a shirker when he expends effort. In a type-two error, agent j is identified as someone who spent effort even though he shirked. The monitoring technology may be one that does not allow either type of error. This is the case with our technology (2) described in Section 4.1 . The monitoring technology may be one that does not allow type-two errors, but that does not distinguish high effort by j from a failure to monitor by i. This is the case with technology (1) described in this section. Finally, the monitoring technology may be one that does not allow type-one errors, but that does not distinguish low effort by j from a failure to monitor by i. This is the case with technology (3) described in Section 4.1 .

As we show, a technology with type-two errors such as technology (3) cannot achieve the first-best. Thus, technology (1) is the weakest monitoring technology with the potential to support the first-best outcome. For that reason, we continue now under the assumption of technology (1), but we return to the other monitoring technologies in Section 4.1 .

We can describe technology (1) formally as follows: If agent i monitors agent j, then he is able to gather enough evidence to verify that agent j has shirked whenever, indeed, this is the case. In this case, the report $R_i(m_i = 1, a_j = 0)$ returns a value of 0, indicating shirking. If agent i's evidence does not allow the judicial body to verify that j has shirked, then it cannot be verified whether this is because agent j worked hard, $a_j = 1$, or because agent i did not monitor, $m_i = 0$. Thus, if $(m_i, a_j) \neq (1, 0)$, the report $R_i(m_i, a_j)$ is not conclusive on i's monitoring or on j's effort, but rather is a null report, which we denote as $R_i(m_i, a_j) = n$, where n is mnemonic for "null." In summary, agent i's report $R_i(m_i, a_j)$ on agent j is defined by:

$$R_i(m_i, a_j) \equiv \begin{cases} 0, & \text{if } m_i = 1 \text{ and } a_j = 0 \\ n, & \text{otherwise.} \end{cases} \tag{1}$$

Given this monitoring technology, agent i can produce verifiable evidence that j shirked, for example by proving that agent j was not present at the work site or did not spend sufficient time on the project, but agent i cannot produce verifiable evidence that agent j actually exerted effort. If i's evidence fails to show that j shirked, it can be either because j worked hard or because i did not expend enough effort in gathering evidence.

Following the motivation laid out in Section 1 , we assume that the contract between the principal and agent i cannot make agent i's compensation depend on the choices of agent j, and hence that agent i can be held accountable only for his own actions. If this were not the case, then each agent's compensation could depend directly on the joint output, regardless of whether it could be verified that the agent fulfilled his job assignment or not. In such a case, the first-best can be achieved by requiring high effort from both agents and by making compensation conditional on the joint output.

Formally, each agent i's contract specifies a job assignment, consisting of a (minimum) effort level $\hat{a}_i \in A$ and a monitoring requirement $\hat{m}_i \in M$, a transfer from the principal w_i, and a penalty p_i that is imposed if it is verified that agent i did not fulfill his job assignment. Agent i's payoff is $w_i - a_i c_i - m_i x_i$, unless it is verified that $a_i \neq \hat{a}_i$ or $m_i \neq \hat{m}_i$, in which case his payoff is $w_i - a_i c_i - m_i x_i - p_i$.

We consider a three-stage game. In the first stage, the principal simultaneously proposes contracts to the two agents, and each agent i accepts his contract if it satisfies his individual-rationality constraint. Letting $\chi_i(\mathbf{a}, \mathbf{m})$ be the indicator equal to one if agent i is penalized given actions $(\mathbf{a}, \mathbf{m})$ and equal to zero otherwise, and letting σ_i be agent i's possibly mixed strategy, agent i's individual-rationality constraint is

$$w_i \geq \sum_{\mathbf{a}, \mathbf{m}} (p_i \chi_i(\mathbf{a}, \mathbf{m}) + a_i c_i + m_i x_i) \sigma_1(a_1, m_1) \sigma_2(a_2, m_2).$$

In stage two, the agents choose their effort levels and whether to monitor the other agent. In stage three, the principal observes the output of the joint production process and the evidence provided through monitoring, and the transfers are

---

[14] As shown in an earlier version of this paper, our qualitative results extend to the case of stochastic monitoring and output.

made. Because monitoring is costly, the first-best is achieved if and only if the two agents exert effort in production without wasting resources monitoring each other.

3. Results

3.1. The optimal mechanism

We begin by observing that if the principal does not include monitoring responsibility in the job assignments, then the agents shirk. Thus, the job assignment that corresponds to the first-best choices does not produce the first-best outcome. The result follows from the observation that if the principal requires high effort but no monitoring, and if output is low, the principal cannot identify which agent shirked, and so it is not verifiable whether a given agent failed to fulfill his assignment.

Observation 1. *If the principal does not include monitoring in the job assignment, then in equilibrium both agents shirk and output is low.*

If the principal does not include monitoring in the job assignment, then high output cannot be achieved in equilibrium. With no one monitoring him, an agent has an incentive to shirk because the principal cannot verify which agent shirked. However, if agent i monitors agent j, then agent j is forced to exert effort to avoid being penalized. This suggests that the principal may achieve high output by requiring agents to monitor each other's effort. Comparing this result with Observation 1, however, it appears that only a second-best outcome may be achieved—successful joint production takes place only if the principal requires that the agents engage in costly monitoring of each other.[15]

The intuition that only a second-best solution is achievable is overturned in the key result of this section. Proposition 1 below shows that if the principal requires both agents to work hard and to monitor each other, then the profile in which agents work hard *without* monitoring each other is an equilibrium and the first-best is achieved. Thus, job assignments that correspond to the second-best solution may induce the first-best outcome.

We begin by considering the agents' behavior following the assignment that they work hard and monitor each other. Then we examine the principal's incentives over possible assignments.

Lemma 1. *If for every i, $p_i \geq c_i$, then the profile $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$ is an equilibrium of the subgame following the assignment $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$. If in addition for every i, $p_i > c_i + x_i$, then the equilibrium is unique.*

Proof. See Appendix A. □

The intuition behind Lemma 1 is as follows: First, note that if agent i chooses not to exert effort, then it will be verifiable that agent i has not fulfilled his assignment, *even if agent j does not monitor agent i.* Since agent i does not exert effort, production is low, and so the principal can verify that at least one of the two agents did not exert effort. If it were the case that agent i exerted effort and monitored agent j, then agent j must not have exerted effort, in which case agent i would have produced evidence $R_i = 0$. Since this is not the case, the principal verifies that agent i either did not exert effort or failed to monitor the other agent. While it cannot be verified which of the two requirements he has failed to fulfill, it is verifiable that he has not fulfilled the job assignment and so can be penalized.

Second, note that when output is high, it cannot be verified whether agents monitored each other or not. Since output is high, both agents i worked hard, and this explains why $R_i = n$ for both i, i.e., neither agent's evidence can show that the other agent shirked. Hence it cannot be verified that an agent did not attempt to gather evidence that the other agent shirked. In sum, agents are not penalized when playing the profile $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$, but each has a lower payoff if he deviates from this profile (he is penalized if $a_i \neq 1$, and he incurs the cost of monitoring if $m_i \neq 0$), and so it is an equilibrium.

Depending on parameters, the equilibrium $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$ is not necessarily unique in the subgame following the assignment $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$. If $p_i \leq c_i + x_i$, there are also two pure-strategy equilibria, symmetric to each other, in which agent i chooses $(a_i, m_i) = (0, 0)$, agent j chooses $(a_j, m_j) = (0, 1)$, and output is low. But if $p_i > c_i + x_i$, then we conclude by iterated deletion of strictly dominated strategies that $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$ is the unique equilibrium. Thus, in the absence of limited liability constraints, the principal chooses large penalties and the equilibrium outcome is unique.

Having shown that the first-best outcome $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$ is achieved with the assignment $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$ if the penalties are sufficiently large, we now turn to show that the first-best outcome can be achieved only with the assignment $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$, and hence that the principal chooses the assignment $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$ in every equilibrium.[16]

Proposition 1. *In every equilibrium, the principal assigns $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$ and sets $\mathbf{w} = \mathbf{c}$ and agents choose $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$.*

Proof. See Appendix A. □

As we show in Section 3.2, the basic result of Proposition 1 does not depend on whether or not transfers are constrained by limited liability.

---

[15] A similar intuition applies to the analysis of the peer monitoring second-best solution of the moral hazard in teams problem (see, e.g., Arnott and Stiglitz, 1991).

[16] Assignments with $a_i = 0$ for some i are only optimal for the principal if she is unable to implement a positive effort level. Our maintained assumption that $h > c_1 + c_2$ is sufficient to guarantee that the principal strictly prefers to induce a subgame in which output is high.

Although the above result assumes that the agents' costs are common knowledge, it is robust to the introduction of private information. In particular, as shown in Marx and Squintani (2008), we can expand our basic model to account for the possibility of asymmetric information by assuming that each agent i's productive costs (c_i, x_i) are private information with some known distribution. In this case, the principal chooses wages that induce only sufficiently efficient agents to accept employment. Because monitoring requirements allow the principal costlessly to induce agents to exert effort, wages can be adjusted so as to screen agents efficiently. An implication is that, unlike in many adverse selection problems, informational rents are collected only by sufficiently efficient agents.

3.2. Contract restrictions

The penalties specified in Proposition 1 exceed the wage and so the principal collects a net payment from agent i if he is found shirking. This formulation may not be realistic in the environments under consideration. In many environments of interest, it is natural to assume that agents have limited liability, and that, at worst, the principal can refuse payment, i.e., $p_i \leq w_i$.

In this case, there is a continuum of Pareto-optimal equilibria in which compensation is such that $\mathbf{c} \leq \mathbf{w} \leq \mathbf{c} + \mathbf{x}$. Because agents may not coordinate on the principal's favored equilibrium in some subgames, the principal may be forced to offer compensation greater than c_i. The only effect of limited liability constraints is potentially to shift rents from the principal to the agents.

Proposition 2. *Assuming limited liability for agents, if h is sufficiently large relative to $\sum_{i=1}^{2} c_i + x_i$, then in every equilibrium the principal assigns $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$ with $w_i \in [c_i, c_i + x_i]$ for each i, and the agents choose $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$.*

Proof.   See Appendix A.   □

Proposition 2 imposes limited liability and shows the key result in Proposition 1 continues to hold—the Pareto frontier is achieved, although in contrast to Proposition 1, agents may collect rents.

Another robustness check involves noting that in the equilibrium described in Proposition 1, the assignments specify effort and monitoring, which if implemented would impose cost $c + x$ on each agent, but each agent's compensation may be only c. An interesting question is whether such assignments would be allowed in the environments under consideration. If the principal makes job assignments whose costs are disproportionate relative to the agents' compensation, it may not be possible for the principal to enforce the contracts. In particular, a judicial body may not allow an agent to be penalized even if it verifies that he did not fulfill his assignment. We refer to an assignment $(\hat{a}_i, \hat{m}_i)$ as "overburdening" if $\hat{a}_i c_i + \hat{m}_i x_i > w_i$.

We show that if agents have limited liability and overburdening assignments are illegal, then the equilibrium is unique and Pareto efficient. However, the agents capture the rent related to the monitoring costs x_i, for which they are compensated in the contract, but which they do not incur on the equilibrium path.

Proposition 3. *Assuming limited liability for agents and that overburdening assignments are illegal, if h is sufficiently large relative to $\sum_{i=1}^{2} c_i + x_i$, then in every equilibrium the principal assigns $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$ with $w_i = c_i + x_i$ for each i, and the agents choose $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$.*

Proof.   As in the proof of Proposition 2, under limited liability the principal optimally sets $p_i = w_i$ for each i. If overburdening assignments are illegal, the principal can assign $(\hat{a}_i, \hat{m}_i) = (1, 1)$ if and only if $w_i \geq c_i + x_i$. This restriction narrows the set of equilibria described in the proof of Proposition 2 to the one with $w_i = c_i + x_i$ for each i.   □

Propositions 2 and 3 show that even when we impose what we consider to be realistic restrictions on contracts between the principal and agents, our main result continues to hold—in equilibrium the principal assigns effort and monitoring, and although the agents do work hard in equilibrium, they violate the assignment by not monitoring. Thus, the efficient outcome is obtained through the imposition of a contract that calls for inefficient behavior.

3.3. Legal system interpretation

In a further attempt to link our model to real-world institutions for enforcing contracts, in this section we discuss how the monitoring technology can arise through judicial institutions.

The monitoring technology in (1), as well as others described in Section 4.1, can be derived from underlying assumptions about how a judicial body evaluates agents' evidence about their teammates and evidence presented by agents confirming their own efforts.

To see this, assume that agents can provide either hard evidence or only cheap-talk. The action of gathering hard evidence is costly. Specifically, agent i can monitor agent j's effort and produce hard evidence at cost x. We denote this choice by agent i by $m_i = 1$. We denote agent i's choice to produce cheap-talk regarding agent j's effort by $m_i = 0$. Formally, each report by agent i consists of a pair (m_i, r_i), where the first component m_i identifies the quality of the report, and the second component r_i identifies the content of the report: agent i claims that agent j has taken effort r_i. We assume that it is impossible to produce hard evidence contrary to the facts, i.e., if $m_i = 1$, then $r_i = a_j$.

One way to motivate our assumption on the monitoring technology is to assume that after the project's output is determined, the principal collects all agents' reports. If she wishes to sanction any agent i, she must bring i's case before a judicial

body. The principal may present the output and reports. In defense, each agent i may present his own self-report (m_i^s, r_i^s) on his productive effort. The burden of proof is on the principal to show that agent i did not fulfill his assignment. In order to simplify the exposition, we assume that it is never the case that $r_i^s = 0$, i.e., we assume player i never produces a self-implicating report in which he takes the blame for low output.

The judicial body allows agent i to be penalized if and only if it can conclude that $a_i \neq \hat{a}_i$ or $m_i \neq \hat{m}_i$, against the presumption that $(a_i, m_i) = (\hat{a}_i, \hat{m}_i)$. Before the judicial body takes its decision, all reports presenting contradicting claims are compared. We assume that hard-evidence reports defeat cheap-talk reports in the analysis of the judicial board, and that in the case that all reports are cheap-talk, agent i's claim that he fulfilled his assignment stands and he cannot be penalized.[17] Specifically, for any i, we introduce the concept of *aggregate evidence claim* $r^i \in \{0, 1\}$, where $r^i = r_i^s$ if $m_j = 0$, and $r^i = r_j$ if $m_j = 1$. Note that by construction all hard-evidence reports must agree.

Assuming technology (1) is equivalent to assuming that the judicial body bases its decision on the verifiable output π and the aggregate evidence claim r^i.

Lemma 2. *Assuming technology* (1) *is equivalent to assuming that the judicial body bases its conclusions on π and r^i for all i.*

Proof. By construction, $r^j = 0$ if and only if $m_i = 1$ and $a_i = 0$, hence the result is obtained by setting $R_i(m_i, a_j) = 0$ if $(m_i, a_i) = (1, 0)$ and $R_i(m_i, a_j) = n$ otherwise. □

In order to conclude the discussion, we explore the possibility that the first-best is achieved if each agent is required to work hard and produce verifiable evidence proving he has not shirked. We let $\hat{m}_i^s = 1$ denote the assignment that agent i should produce a self-report proving he worked hard.

Proposition 4. *If the principal assigns $(\hat{a}_i, \hat{m}_i^s) = (1, 1)$ and the judicial body observes m_i^s, then agent i can be penalized if he chooses $a_i = 0$. If instead the judicial body does not observe monitoring outcomes, but only observes the aggregate evidence claim, then agent i cannot be penalized if he chooses $a_i = 0$.*

Proof. If the judicial body observes m_i^s, it will allow the principal to sanction i unless $m_i^s = 1$, but in this case $r_i^s = a_i$ and hence the judicial body allows the principal to sanction i unless $a_i = 1$. If the judicial body does not observe monitoring outcomes, the observation of $\pi = \ell$ and $r^i = 1$ is insufficient to rule out the possibility that $(a_i, m_i^s) = (1, 1)$, and so the judicial body does not allow agent i to be penalized. □

When technology (1) is adopted, absent any additional evidence, the judicial body is unable to determine whether agent i's self-report is cheap-talk, and hence does not allow agent i to be penalized. The key difference in Lemma 1 is that the requirement to monitor the other agent induces a conflict between the agents that would generate conflicting evidence in case of a hearing, and help the inference of the judicial body on the agents' monitoring effort.

4. Extensions

4.1. Other monitoring technologies

As an alternative to monitoring technology (1), we consider the case in which agent i's evidence reveals whether agent i monitored agent j or not, and hence the choice of monitoring is itself verifiable, as well as revealing agent j's effort when agent i does monitor. In this case, monitoring results in evidence either of shirking or of exerting effort. Thus, when agent i monitors, we let i's report have value equal to agent j's effort choice, i.e., $R_i(1, a_j) = a_j$. If $m_i = 0$, the report $R_i(m_i, a_j)$ reveals the absence of monitoring but provides no information as to j's effort. Because the report is uninformative in this case, we let $R_i(0, a_j) = n$, where n is mnemonic for a "null" report. Formally, for any agents i and j,

$$R_i(m_i, a_j) \equiv \begin{cases} 0, & \text{if } m_i = 1 \text{ and } a_j = 0 \\ 1, & \text{if } m_i = 1 \text{ and } a_j = 1 \\ n, & \text{if } m_i = 0. \end{cases} \tag{2}$$

When monitoring technology (2) is adopted, if $R_i(m_i, a_j) = n$, then nothing can be inferred about j's action, but it can be verified that agent i did not monitor agent j. Technology (2) is appropriate to represent a specific verifiable action by agent i that generates evidence about j's behavior. For instance, suppose that the principal requests that agent i produce an evaluation or review of agent j's performance. Then, if agent i fails to produce the evaluation or review, the court verifies that he did not fulfill his assignment. As mentioned in Section 1, measures exist to ascertain the rigor of evaluations and to identify invalid evaluations, so that 'fake' evaluations can be detected (see Edwards and Ewen, 1996). Thus, assuming technology (2) is equivalent to assuming that the judicial body can directly observe m_i in any report (m_i, r_i). Thus, assuming technology (2) is equivalent to saying that the judicial body bases its conclusions on π and on (m_i, r_i) for all i.

---

[17] The theoretical properties of related mechanisms are studied, for instance, in Lipman and Seppi (1995). Müller and Wärneryd (2001) consider a model of firm ownership in which conflicts between insider and outsiders to the firm are resolved by appeal to a legal system. Prendergast (1993) considers incentives in an environment in which agents need not report their observations honestly.

When technology (2) is adopted, requiring an agent to work hard and to produce hard evidence that his own effort was high achieves high effort. But this mechanism is unlikely to be viable in the environments in which we are interested. Suppose that agent i is assigned the task $(\hat{a}_i, \hat{m}_i^s) = (1, 1)$ and that the principal accuses agent i of not having fulfilled his job assignment. Under the assignment $(\hat{a}_i, \hat{m}_i^s) = (1, 1)$, the agent is penalized unless he can gather and present evidence that he worked hard. This hearing is logically equivalent to a hearing in which agent i is accused of having played $a_i = 0$, and is penalized unless he proves that he has played $a_i = 1$, against the presumption that he has played $a_i = 0$. In such a hearing, the burden of the proof is borne by agent i, contrary to the basic principles of the environments under consideration. This can be avoided only if the job assignment $(\hat{a}_i, \hat{m}_i^s) = (1, 1)$ is ruled out. In particular, an assignment requiring an agent both to work hard and to prove that he has worked hard does not seem plausible in the economic environments of interest.

By instructing the agents to work hard and monitor each other with technology (2), the principal cannot induce the equilibrium $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$ unless transfers are restricted to *damage compensation*, by which we mean that no agent is penalized unless $\pi = 0$ and $\hat{\mathbf{a}} = \mathbf{1}$, or unless the principal can otherwise credibly precommit not to sanction agents when output is high. Of course, this precommitment cannot be included in the individual assignments $(\hat{\mathbf{a}}, \hat{\mathbf{m}})$ which depend only on each agent i's actions and not on joint output. One further possibility could be that the principal chooses not to penalize agents when output is high because of reputational concerns.

If in fact agent i works hard ($a_i = 1$) without monitoring agent j ($m_i = 0$), then the principal can verify that j did not monitor because $R_j = n$. This implies that the principal prefers to instruct agents to use technology (1) rather than technology (2), so as to save the cost of compensating the agents for monitoring each other.

When transfers are restricted to damage compensation, however, it can be shown that technologies (1) and (2) yield identical incentives to the agents any time the principal requires that they work hard.

Lemma 3. *Assuming transfers are restricted to damage compensation, if $\hat{\mathbf{a}} = \mathbf{1}$, then for any actions $(\mathbf{a}, \mathbf{m})$, agent i is penalized under technology (2) if and only if he is penalized under technology (1).*

Proof. The result trivially holds if $\hat{m}_i = 0$. Suppose $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$. If $(a_i, m_i) = (0, 0)$, then $\pi = \ell$ and agent i is penalized under either technology. If instead $(a_i, m_i) = (0, 1)$ and $a_j = 1$, then $\pi = \ell$ but only agent i is penalized under either technology. If $a_i = 1$ and $a_j = 1$, then output is high and the restriction to damage compensation prevents players from being penalized under either technology. $\square$

Lemma 3 can be used to show that if technology (2) is used and transfers are restricted to damage compensation, then Proposition 1 continues to hold—in equilibrium the principal assigns $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$ and $\mathbf{w} = \mathbf{c}$, and agents choose $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$, so that the first-best is achieved and the principal captures all gains.

As another alternative, we can consider the converse representation in which an agent can prove through monitoring that the other agent exerted the proper effort, but cannot prove through monitoring that he shirked. In this case, we let $R_i(m_i = 1, a_j = 1) = 1$. But when $(m_i, a_j) \neq (1, 1)$, the report does not reveal the level of monitoring or effort, so in that case we let the report have value $R_i(m_i, a_j) = n$. Formally, for any agents i and j,

$$R_i(m_i, a_j) \equiv \begin{cases} 1, & \text{if } m_i = 1 \text{ and } a_j = 1 \\ n, & \text{otherwise.} \end{cases} \tag{3}$$

Depending on the specific environment, there may be only one available monitoring technology, or multiple monitoring technologies may be feasible, with the principal able to instruct agents which one to adopt.

By instructing the agents to work hard and monitor each other with technology (3), the principal fails to induce high effort in equilibrium. In fact, if output is low, $\pi = \ell$, then $R_i = n$ for each agent i, under technology (3). Because $R_i = n$, the principal can conclude neither that agent i chose $m_i = 0$, nor that agent j chose $a_j = 0$. As in the case when agents are not requested to monitor each other (Observation 1), each agent i shirks because the principal cannot verify that i fails to fulfill the assignment $(\hat{a}_i, \hat{m}_i) = (1, 1)$.

Observation 2. *If the principal assigns $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$ with technology (3), then in equilibrium both agents shirk and output is low.*

The intuition for Observation 2 is that bonuses have to be feasible in order to implement a positive effort level with technology (3). It follows that monitoring under technology (3) yields the same incentives for the agents as not including monitoring requirement in the assignments at all.

4.2. More than two agents and actions

This section extends the analysis to allow for any finite number of agents and actions. For the purposes of this section, we assume agents coordinate on the principal's preferred equilibrium in each subgame. Our results hold under both unlimited and limited liability.

Let $I \geq 2$ be the set of agents, indexed by i. Each agent has utility $w_i - c_i(a_i) - x_i(\mathbf{m}_i) - p_i \chi_i(\pi, \mathbf{R})$, where $\chi_i(\pi, \mathbf{R}) = 1$ if and only if agent i can be penalized when output π and evidence $\mathbf{R}$ is presented to the judicial authority, w_i is the compensation, $\mathbf{m}_i \in \times_{j \neq i}\{0, 1\}$ is the vector of monitoring choices, and $a_i \in A$ is the productive effort by i, where A is a finite subset of $\mathbb{R}_+$ with

$\underline{a} \equiv \min A$. We allow for general cost functions: $c_i(a_i)$ is the cost of effort, increasing in a_i, and $x_i(\mathbf{m}_i)$ is the cost of monitoring, where $x_i(\mathbf{0}) = 0$ and x_i is strictly increasing in $\mathbf{m}_i$.

We extend the monitoring representations (1) and (2) in the following natural way. We extend (1) so that agent i's report on agent j is

$$R_{ij}(m_{ij}, a_j) \equiv \begin{cases} 0, & \text{if } m_{ij} = 1 \text{ and } a_j \neq \hat{a}_j \\ n, & \text{otherwise.} \end{cases} \tag{1'}$$

We extend (2) so that agent i's report on agent j is

$$R_{ij}(m_{ij}, a_j) \equiv \begin{cases} a_j, & \text{if } m_{ij} = 1 \\ n, & \text{if } m_{ij} = 0. \end{cases} \tag{2'}$$

In equilibrium, each agent i joins the team if and only if his participation constraint is satisfied:

$$w_i \geq \sum_{\mathbf{a}, \mathbf{m}} (p_i \chi_i(\pi, \mathbf{R}) + c_i(a_i) + x_i(\mathbf{m}_i)) \Pi_{j \in I} \sigma_j(a_j, \mathbf{m}_j).$$

The output is given by the production function $\pi(\mathbf{a})$, where we impose anonymity, i.e., we assume that π is invariant to permutations of the agents' actions. Let $\mathbf{a}^*$ be the (assumed unique, for expositional simplicity) socially optimal effort, i.e.,

$$\mathbf{a}^* \in arg\max_{\mathbf{a} \in \times_{i \in I} A} \pi(\mathbf{a}) - \sum_{i \in I} c(a_i).$$

To rule out trivial solutions, we assume $a_i^* > \underline{a}$ for any i. For future reference, we let $V(n) \equiv \pi(\mathbf{a}^*) - \pi(\underline{\mathbf{a}}) - (\sum_{i=1}^{n} c_i(a_i^*) - \sum_{i=1}^{n} c_i(\underline{a}))$ denote the incremental social value associated with effort $\mathbf{a}^*$ versus $\underline{\mathbf{a}}$ as a function of the number of players n.

As in the case of two agents, the principal cannot achieve high effort by means of job assignments that do not require monitoring. But unlike the two-agent case, the extent of the monitoring requirements necessary to achieve the first-best depends on the monitoring technology. Specifically, we show that the monitoring requirements inducing high output are more demanding when technology (1') is adopted—output $\pi(\mathbf{a}^*)$ is achieved only if the principal requires every agent to monitor every other agent. As in the previous section, in equilibrium the agents exert effort, but do not monitor each other, so that the first-best is achieved.

Proposition 5. *Assuming technology* (1'), *in every equilibrium the principal assigns* $(\hat{a}_i, \hat{\mathbf{m}}_i) = (a_i^*, \mathbf{1})$ *with* $w_i = c_i(a_i^*)$ *for each i, and the agents choose* $(\mathbf{a}, \mathbf{m}) = (\mathbf{a}^*, \mathbf{0})$.

Proof. See Appendix A. □

While it may seem infeasible to require each agent to monitor every other agent, this requirement is not carried out on the equilibrium path. In fact, this equilibrium always exists, regardless of the value of the monitoring cost functions x_i. When overburdening assignments are illegal, however, we obtain an impossibility result for the case of large teams, because in a team of n members, each agent is required to monitor all remaining $n - 1$ team members in order to achieve the first-best, and the principal is required to compensate the agents also for these monitoring efforts. Hence the burden laid out in the assignments grows quickly in the size of the team, while the principal's net benefit from achieving the first-best need not grow as fast. For simplicity, the following proposition focuses on linear monitoring costs.

Proposition 6. *Assuming overburdening assignments are illegal and assuming technology* (1'), *if for every i, $x_i(\mathbf{m}_i) = x \sum_{j \neq i} m_{ij}$, and $V(n)$ grows slower than quadratically in n, then for n sufficiently large, there is no equilibrium with actions $\mathbf{a}^*$.*

Proof. See Appendix A. □

The above impossibility result suggests the possibility that random monitoring schemes may alleviate the legal burden associated with monitoring. Suppose that the parties may use an external randomizing device. The principal may now choose an assignment $\hat{\alpha}$, a (possibly correlated) distribution over the possible deterministic assignments $(\hat{\mathbf{a}}, \hat{\mathbf{m}})$. It is understood by the parties that the external randomizing device is operated only *after* the agents have made their choices, so as to extract an outcome $(\hat{\mathbf{a}}, \hat{\mathbf{m}})$ out of the distribution $\hat{\alpha}$. Suppose that any such random assignment $\hat{\alpha}$ is deemed overburdening only when the expected assignment $\mathbb{E}(\hat{\mathbf{a}}_i c_i + x(\hat{\mathbf{m}}_i)) > w_i$ for some agent i. Proposition 5 implies that each agent i will play $(a_i, \mathbf{m}_i) = (a_i^*, \mathbf{0})$, as long as $\hat{\alpha}$ places enough probability on the outcome $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{a}^*, \mathbf{1})$. Under unlimited liability, this makes it possible to find an efficient random assignment $\hat{\alpha}$ for all parameter constellations.

Proposition 7. *Assuming overburdening assignments are illegal and assuming technology* (1'), *if parties may use an external randomizing device, there exist equilibria with actions $\mathbf{a}^*$.*

Proof. See Appendix A. □

Although Proposition 7 offers a positive result, it seems unlikely that overburdening assignments would be illegal and yet agents unlimitedly liable. Under limited liability, a negative result analogous to Proposition 6 holds.

Proposition 8. *Assuming overburdening assignments are illegal and agents are limited liable, and assuming technology* $(1')$*, if (i) for every i, $x_i(\mathbf{m}_i) = x\sum_{j \neq i} m_{ij}$, (ii) $c_i(a_i)$ is constant across i, and (iii) $V(n)$ grows in n at a rate slower than $n^{3/2}$, then for n sufficiently large, there is no equilibrium with actions $\mathbf{a}^*$.*

Proof. See Appendix A. □

When the monitoring technology $(2')$ is adopted, the principal may achieve the first-best only if transfers are restricted to damage compensation. In this case, it is not necessary to request that agents monitor every other agent. Agents can be induced to exert effort by requiring them to monitor sufficiently many other agents that the monitoring assignment is more costly to fulfill than the effort assignment. In order to avoid the penalty, each agent prefers to exert effort rather than monitoring.

Proposition 9. *Assuming transfers are restricted to damage compensation and assuming technology* $(2')$*, the first-best is achieved in every equilibrium if and only if $\hat{a}_i = a_i^*$ and either $x_i(\hat{\mathbf{m}}_i) \geq c_i(a_i^*)$ or $\hat{\mathbf{m}}_i = \mathbf{1}$.*

Proof. See Appendix A. □

While in the previous sections, we have identified reasons why technology $(1')$ is preferable to technology $(2')$, the corollary of Proposition 9 stated below identifies one special case in which the reverse is true. If transfers are restricted to damage compensation and overburdening assignments are illegal, then technology $(2')$ may be preferable to technology $(1')$ when the team size is large enough. Unlike technology $(1')$, technology $(2')$ allows the principal profitably to give the agents incentive to work hard, regardless of team size. For simplicity, the corollary focuses on linear symmetric costs and linear output.

Corollary 1. *Assuming overburdening assignments are illegal and transfers are restricted to damage compensation, and assuming technology* $(2')$*, if there is a profile $(\hat{\mathbf{m}}_i)_{i \in I}$ such that for every i, $x_i(\hat{\mathbf{m}}_i) \geq c_i(a_i^*)$, and $\sum_{i \in I} x_i(\hat{\mathbf{m}}_i) \leq V(n)$, then the profile $\mathbf{a}^*$ can be achieved in equilibrium.*

One can show that the conditions of Corollary 1 hold, for example, when c_i and x_i are constant in i, x_i is linear in $\sum_{j \neq i} m_{ij}$, and $V(n)$ grows linearly in n.

5. Literature review

Following the seminal work of Holmström (1982), which shows that an efficient sharing rule for the team's output respecting ex-post budget balance does not exist, a substantial literature has studied environments of moral hazard in teams. Our paper considers the principal multi-agent setting studied by Ma (1988), which we discuss further below. We focus on peer monitoring and provide the first study, to our knowledge, to show that monitoring mechanisms play a role in achieving the first-best, even when monitoring itself is wasteful.

The previous literature, such as Arnott and Stiglitz (1991), focuses on second-best solutions. Baliga and Sjostrom (2001) study optimal peer assessments mechanisms and find a tendency for overly negative valuations. Miller (1997) considers an asymmetric joint monitoring mechanism. In experimental work, Nalbantian and Schotter (1997) consider the role of monitoring on group performance and find monitoring elicits high effort from workers only if agents perceive a high monitoring frequency. In contrast, our results suggest a mechanism by which the requirement of monitoring can be valuable even if monitoring does not occur.

Our paper is related to Ma (1988), who also considers a model in which a principal hires multiple agents and is unable to monitor their actions. Our paper is similar to Ma's in that agents have the ability to monitor each other, but our papers differ in that Ma's regularity conditions rule out a team production technology such as ours where symmetric agents must all exert high effort to achieve high output. Ma shows that the first-best is achieved in the unique equilibrium if the principal first asks one agent to reveal both agents' actions and then allows the other agent the opportunity to either agree or challenge the first agent's report, where with a challenge the agent reports a different action pair. Ma shows that payoffs can be designed so that truthful revelation for the first agent and no challenge by the second agent is the unique best response. A key difference between Ma (1988) and our paper is that in Ma's paper monitoring is costless so the first-best involves monitoring. In our environment, monitoring is costly so the first-best outcome requires the first-best productive actions but no monitoring. In the equilibrium of our model the first-best is achieved because although the principal includes monitoring in the agents' job assignments, they do not complete that assignment in equilibrium. Another key difference is that our model is motivated by environments with individual liability. In contrast, the transfers that follow a challenge in Ma's model depend on both players' actions.[18]

We derive a first-best incentive system in a multi-agent setting using bilateral contracts. In contrast, Segal (1999) shows that when bilateral contracts are public, inefficiencies arise due to the externalities on agents' reservation utilities; whereas when the principal's offers are privately observed, inefficiencies are due to the externalities at efficient outcomes. Inefficiencies can be eliminated if the principal can adopt a multilateral contract scheme, whereby each agent's contract depends on

---

[18] There are also other differences in modeling choices. In Ma (1988), productive actions and monitoring reports are made sequentially, but in our model agents simultaneously choose their productive action and whether to monitor. In Ma, monitoring reports are made on the equilibrium path and determine payoffs, but in the equilibrium of our model, monitoring reports are not relevant when observed output is high.

other agents' messages. We show that inefficiencies can be eliminated without conditioning on other agents' messages by incorporating monitoring requirements into the contracts.

Segal (2003) extends the model of Segal (1999) to examine the effects of prohibiting the principal from (i) coordinating agents on her preferred equilibrium, and (ii) making different contracts available to different agents. The inefficiencies under different contracting regimes are shown to be typically reduced by both prohibitions. In our model, all inefficiencies can often be eliminated by incorporating monitoring requirements into the contracts. The only effect of prohibiting the principal from coordinating agents on her preferred contract is to generate equilibrium multiplicity and potentially shift rents from the principal to the agents.

6. Conclusion

We consider a team production problem in which the principal observes only the group output and not individual effort, but in which agents are only individually accountable, so that the principal's compensation of each agent cannot depend on other agents' actions. In this environment, agents have incentives to shirk, but we show that by including monitoring in the agents' job assignments, the principal induces the agents to exert effort. Moreover, even though equilibrium job assignments include monitoring, this serves only to provide incentives for effort. Agents do not engage in wasteful monitoring in equilibrium, and the first-best is achieved.

The first-best outcome cannot be achieved if the principal takes the monitoring responsibility on herself because she would have to commit herself to monitor with positive probability in order to induce effort and monitoring is costly. Thus, we identify a benefit of delegated monitoring not addressed in the existing literature.

Our analysis shows that the main results hold with different representations of the monitoring technology and under various restrictions of the transfers between the principal and agents. When large teams are considered, we show that it may become more difficult to provide incentives for effort because the required monitoring may become too costly and, even if not undertaken on the equilibrium path, assignments may be legally constrained not to be disproportionate relative to compensation.

At the methodological level, our analysis describes an environment in which the standard approach to solving contract and implementation problems fail—while the contract prescribing the first-best fails to induce the desired outcome, the first-best can be achieved with a contract prescribing second-best actions, where the agents do not comply in equilibrium. We emphasize that, in order to achieve desired actions, negotiating parties should not restrict attention to contracts under which certain players commit to choose these actions. Instead, they may need to consider "roundabout," apparently contradictory mechanisms that specify different actions, with the understanding that the contracts will be violated so as to achieve the optimal outcome.

Acknowledgements

Appendix A. Proofs

Proof of Lemma 1. We begin by proving the following claim: if the principal observes low output and $R_i \neq 0$, then it is verifiable that $(a_i, m_i) \neq (1, 1)$. To see this, suppose output is low and $R_i \neq 0$, and let $j \neq i$. Since $R_i \neq 0$, then the principal can verify that $(m_i, a_j) \neq (1, 0)$. At the same time, since output is low, the principal can verify that $(a_i, a_j) \neq (1, 1)$. If $a_i = 1$, then the second conclusion implies that $a_j = 0$, but in such a case, the first conclusion implies that $m_i = 0$. Thus, the principal can verify that $(a_i, m_i) \neq (1, 1)$.

Continuing the proof, assume $(\hat{a}_1, \hat{m}_1) = (\hat{a}_2, \hat{m}_2) = (1, 1)$ and $(a_j, m_j) = (1, 0)$. If agent $i \neq j$ chooses $(a_i, m_i) = (1, 0)$, then output is high, no agent is penalized, and agent i's payoff is $w_i - c_i$. A deviation by agent i to $(a_i, m_i) = (1, 1)$ gives payoff $w_i - c_i - x_i$, and so is not profitable. If agent i deviates and chooses $(a_i, m_i) = (0, 0)$, then output is low and $R_i = n$, so using the claim, the principal can verify that $(a_i, m_i) \neq (1, 1)$ and agent i's payoff is $w_i - p_i$. This makes the deviation unprofitable, under the assumption that $p_i > c_i$. If agent i deviates and chooses $(a_i, m_i) = (0, 1)$, then output is low and $R_i = n$, so once again the claim implies that the principal can verify that $(a_i, m_i) \neq (1, 1)$ and can penalize agent i, making the deviation unprofitable.

To prove the claim of uniqueness, suppose for all i, $p_i > c_i + x_i$. We proceed by iterated deletion of strictly dominated strategies. First note that $(a_i, m_i) = (0, 0)$ yields a payoff $w_i - p_i$ to agent i, because output is low and $R_i = n$ regardless of agent j's strategy. Hence, $(a_i, m_i) = (0, 0)$ is strictly dominated by $(a_i, m_i) = (1, 1)$, which gives the payoff $w_i - c_i - x_i$ regardless of the agent j's action. Once strategy $(a_i, m_i) = (0, 0)$ is deleted for both players, then $(a_i, m_i) = (0, 1)$ is strictly dominated by $(a_i, m_i) = (1, 1)$. In fact, if agent i plays $(a_i, m_i) = (0, 1)$, then his payoff is $w_i - x_i - p_i$ as long as the opponent plays either $a_j = 1$ or $m_j = 1$ or both. Once strategy $(a_i, m_i) = (0, 1)$ is deleted for both players, then the strategy $(a_i, m_i) = (1, 0)$ strictly dominates the strategy $(a_i, m_i) = (1, 1)$. Because both players choose high effort in the strategies that have not been deleted,

output is high regardless of the monitoring choice of either player. So, $(a_i, m_i) = (1, 1)$ yields payoff $w_i - c_i - x_i$ which is smaller than $w_i - c_i$, the payoff given by $(a_i, m_i) = (1, 0)$. □

Proof of Proposition 1. It follows from Lemma 1 and the principal's interest in minimizing wages subject to the individual rationality constraint that the proposed equilibrium is, in fact, and equilibrium. It remains to show that it is unique. If the principal assigns $(\hat{a}_i, \hat{m}_i) = (1, 0)$ to each agent i, we have already established that each agent i chooses $a_i = 0$, and thus output is low. It is immediate that assigning $\hat{a}_i = 0$ to either agent induces $a_i = 0$, and thus output is low. Thus, we are left to consider only the subgame associated with the asymmetric assignments $(\hat{a}_j, \hat{m}_j) = (1, 1)$ and $(\hat{a}_i, \hat{m}_i) = (1, 0)$. We show that in any equilibrium of this subgame, it must be the case that $\Pr(\pi(\mathbf{a}) = h) < 1$ (in this subgame there are no pure-strategy equilibria).

Suppose $(\hat{a}_j, \hat{m}_j) = (1, 1)$ and $(\hat{a}_i, \hat{m}_i) = (1, 0)$. Consider the strategy $(a_i, m_i) = (1, 0)$. Since $\hat{m}_i = 0$, agent i cannot be penalized, regardless of the strategy of agent j. Since $x_i > 0$, this implies that the strategy $(a_i, m_i) = (1, 0)$ strictly dominates the strategy $(a_i, m_i) = (1, 1)$. This implies that an equilibrium such that $\Pr(\pi(\mathbf{a}) = h) = 1$ can exist only if agent i chooses $(a_i, m_i) = (1, 0)$. Following the Proof of Lemma 1, agent j's best response to $(a_i, m_i) = (1, 0)$ is $(a_j, m_j) = (1, 0)$. The best response of agent i to $(a_j, m_j) = (1, 0)$ is $(a_i, m_i) = (0, 0)$. Since $m_j = 0$, it follows that $R_i = n$, and thus the principal cannot rule out $a_i = 1$. Since $\hat{m}_i = 0$, agent i cannot be penalized for choosing $m_i = 0$. This concludes the proof that when $(\hat{a}_j, \hat{m}_j) = (1, 1)$ and $(\hat{a}_i, \hat{m}_i) = (1, 0)$, there cannot be an equilibrium in which $a_i = a_j = 1$ with probability one.

To complete the proof, we show that the principal prefers $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$ over $(\hat{a}_j, \hat{m}_j) = (1, 1)$ and $(\hat{a}_i, \hat{m}_i) = (1, 0)$. As a matter of notation, recall that we define $\chi_i(\mathbf{a}, \mathbf{m})$ to be the indicator equal to one if agent i is penalized given actions $(\mathbf{a}, \mathbf{m})$ and equal to zero otherwise. Let σ be the possibly mixed strategy profile played in the continuation game following the assignment $(\hat{a}_j, \hat{m}_j) = (1, 1)$ and $(\hat{a}_i, \hat{m}_i) = (1, 0)$, i.e., $\sigma_i(a_i, m_i)$ is the probability with which player i chooses (a_i, m_i) following this assignment. The principal's expected payoff from the assignment is

$$\mathbb{E}\left(\pi + \sum_{i=1,2} p_i\right) - \sum_{i=1,2} w_i = h\Pr(\pi(\mathbf{a}) = h) + \sum_{\mathbf{a},\mathbf{m}}(p_1\chi_1(\mathbf{a}, \mathbf{m}) + p_2\chi_2(\mathbf{a}, \mathbf{m}))\sigma_1(a_1, m_1)\sigma_2(a_2, m_2) - w_1 - w_2$$

$$\leq h\Pr(\pi(\mathbf{a}) = h) - \sum_{\mathbf{a},\mathbf{m}}(a_1 c_1 + m_1 x_1 + a_2 c_2 + m_2 x_2)\sigma_1(a_1, m_1)\sigma_2(a_2, m_2)$$

$$= h\Pr(\pi(\mathbf{a}) = h) - \sum_{\mathbf{a} \neq \mathbf{1},\mathbf{m}}(a_1 c_1 + m_1 x_1 + a_2 c_2 + m_2 x_2)\sigma_1(a_1, m_1)\sigma_2(a_2, m_2)$$

$$- \sum_{\mathbf{a}=\mathbf{1},\mathbf{m}}(c_1 + m_1 x_1 + c_2 + m_2 x_2)\sigma_1(a_1, m_1)\sigma_2(a_2, m_2) \leq h\Pr(\pi(\mathbf{a}) = h)$$

$$- \sum_{\mathbf{a}=\mathbf{1},\mathbf{m}}(c_1 + c_2)\sigma_1(a_1, m_1)\sigma_2(a_2, m_2) = (h - c_1 - c_2)\Pr(\pi(\mathbf{a}) = h) < h - c_1 - c_2,$$

where the first inequality uses the agents' individual rationality constraints, and the last inequality uses the maintained assumption that $h > c_1 + c_2$ and the result that $\Pr(\pi(\mathbf{a}) = h) < 1$. Since the principal's payoff from $\mathbf{w} = \mathbf{c}$ and $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$ is $h - c_1 - c_2$, and since individual rationality is satisfied at $\mathbf{w} = \mathbf{c}$, the proof is complete. □

Proof of Proposition 2. Under limited liability the principal optimally sets $p_i = w_i$ for each i. In equilibrium, the principal can achieve $\pi = h$ with positive probability only by assigning $\hat{a}_1 = \hat{a}_2 = 1$ and $\hat{m}_i = 1$ for at least one agent i.

Supposing that $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$, the agents' subgame is:

(a, m)	$0, 0$	$0, 1$	$1, 0$	$1, 1$
0,0	$0, 0$	$0, w_2 - x_2$	$0, -c_2$	$0, w_2 - c_2 - x_2$
0,1	$w_1 - x_1, 0$	$-x_1, -x_2$	$-x_1, -c_2$	$-x_1, w_2 - c_2 - x_2$
1,0	$-c_1, 0$	$-c_1, -x_2$	$w_1 - c_1, w_2 - c_2$	$w_1 - c_1, w_2 - c_2 - x_2$
1,1	$w_1 - c_1 - x_1, 0$	$w_1 - c_1 - x_1, -x_2$	$w_1 - c_1 - x_1, w_2 - c_2$	$w_1 - c_1 - x_1, w_2 - c_2 - x_2$.

It is immediate to see that when $w_i > c_i + x_i$, then $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$ is the unique equilibrium; whereas if $w_i \leq c_i + x_i$ and $w_j \geq x_j$, there is also an equilibrium in which $(a_i, m_i) = (0, 0)$ and $(a_j, m_j) = (0, 1)$.

Supposing that $(\hat{a}_1, \hat{m}_1) = (1, 1)$ and $(\hat{a}_2, \hat{m}_2) = (1, 0)$, the agents' subgame is

(a, m)	$0, 0$	$0, 1$	$1, 0$	$1, 1$
0,0	$0, w_2$	$0, w_2 - x_2$	$0, w_2 - c_2$	$0, w_2 - c_2 - x_2$
0,1	$w_1 - x_1, 0$	$-x_1, -x_2$	$-x_1, w_2 - c_2$	$-x_1, w_2 - c_2 - x_2$
1,0	$-c_1, w_2$	$-c_1, w_2 - x_2$	$w_1 - c_1, w_2 - c_2$	$w_1 - c_1, w_2 - c_2 - x_2$
1,1	$w_1 - c_1 - x_1, 0$	$w_1 - c_1 - x_1, -x_2$	$w_1 - c_1 - x_1, w_2 - c_2$	$w_1 - c_1 - x_1, w_2 - c_2 - x_2$.

Case (i). $w_1 \in (c_1, c_1 + x_1)$. The strategies $(a_1, m_1) = (1, 1)$, $(a_2, m_2) = (1, 1)$, and $(a_2, m_2) = (0, 1)$ are strictly dominated. To define mixed strategies, let σ_s^i be the probability with which agent i plays pure strategy s. In the reduced game, the unique equilibrium is $\sigma_{01}^1 = \frac{c_2}{w_2}$, $\sigma_{00}^1 = 1 - \sigma_{01}^1$, and $\sigma_{00}^2 = \frac{x_1}{w_1}$, and hence $\Pr(\pi = h) = 0$.

Case (ii). $w_1 > c_1 + x_1$. The strategies $(a_1, m_1) = (0, 0)$, $(a_2, m_2) = (1, 1)$, and $(a_2, m_2) = (0, 1)$ are strictly dominated. In the reduced game, the unique equilibrium is $\sigma_{11}^1 = \frac{c_2}{w_2}$, $\sigma_{10}^1 = 1 - \sigma_{11}^1$, and $\sigma_{00}^2 = \frac{x_1}{w_1}$, which yields the expected payoff to the principal of

$$\sigma_{10}^2(h - w_1 - w_2) + (1 - \sigma_{10}^2)(-\sigma_{10}^1 w_2 - (1 - \sigma_{10}^1)w_1) = \frac{w_1 - x_1}{w_1}(h - w_1 - w_2) + \frac{x_1}{w_1}\left(c_2\left(1 - \frac{w_1}{w_2}\right) - w_2\right). \tag{A.1}$$

Maximizing this expression yields the optimal wages $w_1^* = \max\{c_1 + x_1, \sqrt{x_1(h - c_2)}\}$, $w_2^* = \max\{c_2, \sqrt{c_2 x_1}\}$.

As long as h is sufficiently large relative to the agents' costs $\sum_{i=1}^2 c_i + x_i$, the expression

$$h\left(\frac{x_1}{w_1^*}\right) - \sum_{i=1}^2 (c_i + x_i) + \frac{w_1^* - x_1}{w_1^*}(w_1^* + w_2^*) + \frac{x_1}{w_1^*}\left(w_2^* - c_2\left(1 - \frac{w_1^*}{w_2^*}\right)\right)$$

is positive. Hence, for any w_1 and w_2, (A.1) is less than $h - \sum_{i=1}^2 (c_i + x_i)$, which is the payoff achieved by assigning $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{1}, \mathbf{1})$ with $w_1 = x_1 + c_1$ and $w_2 = x_2 + c_2$.

Thus, in equilibrium w_1 and w_2 are set to the lowest values subject to continuation play of $(\mathbf{a}, \mathbf{m}) = (\mathbf{1}, \mathbf{0})$, which means $w_i \in [c_i, c_i + x_i]$. $\square$

Proof of Proposition 5. Regardless of the assignment $(\hat{\mathbf{a}}, \hat{m})$, there cannot be any equilibrium in which $\mathbf{a} = \mathbf{a}^*$ and some agent i chooses $m_{ij} = 1$ for some $j \neq i$. Since $\mathbf{a} = \mathbf{a}^*$, no agent is penalized, and the choice $\mathbf{m}_i \neq \mathbf{0}$ cannot be part of a best response because $x_i(\mathbf{m}_i) > 0$. So we now show that the profile $(a_i, \mathbf{m}_i) = (a_i^*, \mathbf{0})$ for all i cannot be an equilibrium in the subgame induced by any assignment where for some agent i, $\hat{a}_i \neq a_i^*$, or $\hat{m}_{ij} = 0$ for some $j \neq i$.

Suppose first that for some agent i, $\hat{a}_i \neq a_i^*$. Then, regardless of $\hat{\mathbf{m}}$, the profile $(a_k, \mathbf{m}_k) = (a_k^*, \mathbf{0})$ for all k cannot be an equilibrium. If $\hat{a}_i < a_i^*$, then agent i prefers to play $(a_i, \mathbf{m}_i) = (\hat{a}_i, \mathbf{0})$ instead of $(a_i, \mathbf{m}_i) = (a_i^*, \mathbf{0})$. Since for any $k \neq i$, the principal cannot rule out $a_k = a_k^*$ and $m_{ik} = 1$, it follows that the principal cannot rule out $(a_i, \mathbf{m}_i) = (\hat{a}_i, \hat{\mathbf{m}}_i)$, and that agent i is not penalized. The same is true if $\hat{a}_i > a_i^*$ and agent i is penalized when choosing $a_i = a_i^*$. If $\hat{a}_i > a_i^*$ and agent i is not penalized when choosing $a_i = a_i^*$, then he is not penalized when playing $a_i = \underline{a}$ either, therefore he plays $\underline{a}$, which is less than a_i^*.

Second, consider any assignment such that $\hat{a}_i = a_i^*$, but $\hat{m}_{ji} = 0$ for some $j \neq i$. The profile $(a_k, \mathbf{m}_k) = (a_k^*, \mathbf{0})$ for all k cannot be an equilibrium, because agent i can profitably deviate by choosing $(a_i, \mathbf{m}_i) = (\underline{a}, \mathbf{0})$. In such a case, since $R_{ki} = n$ for all k, the principal cannot rule out $a_i = a_i^*$, and $\mathbf{m}_i = \hat{\mathbf{m}}_i$. Therefore, i cannot be penalized.

Finally we show that the profile $(a_i, \mathbf{m}_i) = (a_i^*, \mathbf{0})$ for all i is an equilibrium in the subgame induced by the assignment $\hat{a}_i = a_i^*$ and $\hat{m}_{ij} = 1$ for all i and $j \neq i$, with $p_i > c_i(a_i^*)$. Suppose that each agent $j \neq i$ chooses $(a_j, \mathbf{m}_j) = (a_j^*, \mathbf{0})$. Any deviation such that $a_i = a_i^*$ and $\mathbf{m}_i \neq \mathbf{0}$ is not profitable because $x(\mathbf{m}_i) > 0$. Suppose that agent i deviates and chooses $a_i \neq a_i^*$. As a result $\pi \neq \pi(\mathbf{a}^*)$, and so it is verifiable that $\mathbf{a} \neq \mathbf{a}^*$. For any $j \neq i$, since $a_j = a_j^*$, with technology (1) it must be that $R_{ij} = n$. Hence it is verified that there is no $a_j \neq a_j^*$ such that $(a_j, m_{ij}) = (a_j, 1)$, and thus that either $m_{ij} = 0$ for some $j \neq i$, or $a_j = a_j^*$ for all $j \neq i$. It follows that either $a_i \neq a_i^*$ or $m_{ij} = 0$ for some $j \neq i$, and hence agent i is penalized, making the deviation unprofitable. The proof is concluded by noting that, by assumption, the agents coordinate on the principal's preferred equilibrium in every subgame. $\square$

Proof of Proposition 6. If no assignment is provided, each agent i plays $a_i = \underline{a}$ in the production subgame yielding profit $\pi(\underline{\mathbf{a}}) - \sum_{i=1}^n c_i(\underline{a})$. In order to induce him to play any $a_i' \neq \underline{a}$, it is necessary to assign $(\hat{a}_i, \hat{\mathbf{m}}_i) = (a_i', \mathbf{1})$. If overburdening assignments are illegal, it is required that $w_i = c_i(a_i') + x\sum_{j \neq i}\hat{m}_{ij}$, so that the principal's profit is:

$$\pi(\mathbf{a}') - \sum_{i=1}^n c_i(a_i') - x\sum_{i=1}^n \sum_{j \neq i} \hat{m}_{ij} \leq \pi(\mathbf{a}^*) - \sum_{i=1}^n c_i(a_i^*) - n(n-1)x,$$

and the result follows. $\square$

Proof of Proposition 7. The Proof of Proposition 5 concludes that $\mathbf{a} = \mathbf{1}$ in equilibrium if and only if $(\hat{\mathbf{a}}, \hat{\mathbf{m}}) = (\mathbf{a}^*, \mathbf{1})$ and every agent i prefers to play $(a_i, \mathbf{m}_i) = (a_i^*, \mathbf{0})$ over $(a_i, \mathbf{m}_i) = (\underline{a}, \mathbf{0})$ when all the other agents j play $(a_j, \mathbf{m}_j) = (a_j^*, \mathbf{0})$. With a random assignment $\hat{\boldsymbol{\alpha}}$, this condition can be written as

$$u_i(\mathbf{a}^*, \mathbf{0}) = w_i - c_i(a_i^*) \geq w_i - c_i(\underline{a}) - p_i\hat{\boldsymbol{\alpha}}(\mathbf{a}^*, \mathbf{1}) = u_i((\underline{a}, \mathbf{a}_{-i}^*), (0, \mathbf{0})),$$

whereas individual rationality implies $w_i \geq c_i(a_i^*)$. Setting $w_i = c_i(a_i^*)$ for all i, the principal maximizes her profit: $\pi(\mathbf{a}^*) - \sum_i w_i$. The principal sets $\hat{\boldsymbol{\alpha}}(\underline{\mathbf{a}}, \mathbf{0}) = 1 - \hat{\boldsymbol{\alpha}}(\mathbf{a}^*, \mathbf{1})$ to minimize legal burden of the assignment $\hat{\boldsymbol{\alpha}}$; for every agent i, the assignment $\hat{\boldsymbol{\alpha}}$ must satisfy:

$$\mathbb{E}(\hat{\mathbf{a}}_i c_i + x(\hat{\mathbf{m}}_i)) = \hat{\boldsymbol{\alpha}}(\mathbf{a}^*, \mathbf{1})(c_i(a_i^*) + x_i(\mathbf{1})) + (1 - \hat{\boldsymbol{\alpha}}(\mathbf{a}^*, \mathbf{1}))c_i(\underline{a}) \leq w_i. \tag{A.2}$$

Hence, in equilibrium, the principal selects

$$\hat{\alpha}(\mathbf{a}^*, \mathbf{1}) = \max_i \frac{c_i(a_i^*) - c_i(\underline{a})}{c_i(a_i^*) - c_i(\underline{a}) + x_i(\mathbf{1})},$$

$$\text{and for any } i, p_i \geq \frac{c_i(a_i^*) - c_i(\underline{a})}{\hat{\alpha}(\mathbf{a}^*, \mathbf{1})}. \tag{A.3}$$

This induces the agents play $(\mathbf{a}, \mathbf{m}) = (\mathbf{a}^*, \mathbf{0})$. □

Proof of Proposition 8. Under limited liability, it must be that $p_i \leq w_i$, and the principal optimally sets $p_i = w_i$. The constraints (A.2) and (A.3) imply:

$$w_i^2 \geq (c_i(a_i^*) - c_i(\underline{a}))^2 + (c_i(a_i^*) - c_i(\underline{a}))x_i(\mathbf{1}) + c_i(\underline{a})$$

If no assignment is provided, each agent i plays $a_i = \underline{a}$ in the production subgame yielding profit $\pi(\underline{\mathbf{a}}) - \sum_{i=1}^n c(\underline{a})$. In order to induce him to play any $a_i' \neq \underline{a}$, it is necessary to assign $(\hat{a}_i, \hat{\mathbf{m}}_i) = (a_i', \mathbf{1})$. The principal's profit is thus

$$\pi(\mathbf{a}^*) - \sum_{i=1}^n w_i \leq \pi(\mathbf{a}^*) - \sum_{i=1}^n c_i(a_i^*) - \sum_{i=1}^n \left(\sqrt{(c_i(a_i^*) - c_i(\underline{a}))^2 + (c_i(a_i^*) - c_i(\underline{a}))x(n-1) + c_i(\underline{a})} - c_i(a_i^*) \right),$$

because $c_i(a_i)$ is constant across i for all actions a_i, the last term grows in n at rate $n^{3/2}$. This delivers the result. □

Proof of Proposition 9. When transfers are restricted to damage compensation, if the profit $\pi(\mathbf{a}^*)$ is achieved, then the principal does not penalize agent i regardless of the monitoring choices, and hence there is no equilibrium in which $a_i = a_i^*$ for all i, and $m_{ij} = 1$ for some i and j. Take the strategy profile where $a_i = a_i^*$ for all i, and $m_{ij} = 0$ for all i and $j \neq i$. Each agent i has payoff $w_i - c_i(a_i^*)$. By the same argument as in Proposition 5, if $\hat{m}_{ij} = 1$ for all $j \neq i$, then agent i will not take any deviation such that $a_i \neq a_i^*$. Suppose that $\hat{\mathbf{m}}_i \neq \mathbf{1}$. With technology (2), it is verifiable that agent i did not fulfill his job assignment, and hence can be penalized, if there is j such that $m_{ij} \neq \hat{m}_{ij}$. Since $w_i \geq c_i(a_i^*)$, this deters any deviation such that $a_i \neq a_i^*$, $m_{ij} = 0$, and $\hat{m}_{ij} = 1$ for some j. Any deviation $(a_i, \mathbf{m}_i)$ such that $a_i = 0$ and $m_{ij} = 1$ for all j with $\hat{m}_{ij} = 1$ yields payoff to player i that is less than or equal to $w_i - x_i(\hat{\mathbf{m}}_i)$. If $x_i(\hat{\mathbf{m}}_i) \geq c_i(a_i^*)$, then any such deviation is deterred. If $x_i(\hat{\mathbf{m}}_i) < c_i(a_i^*)$, the deviation such that $a_i \neq a_i^*$ and $m_{ij} = 1$ is not deterred even if $\hat{m}_{ij} = 1$. □

References

Aghion, P., Tirole, J., 1997. Formal and real authority in organizations. Journal of Political Economy 105, 1–29.

Arnott, R., Stiglitz, J.E., 1991. Moral hazard and nonmarket institutions: dysfunctional crowding out or peer monitoring? American Economic Review 81, 179–190.

Baliga, S., Sjostrom, T., 2001. Optimal design of peer review and self-assessment schemes. RAND Journal of Economics 32, 27–51.

Bewley, T., 1999. Why Wages don't Fall during a Recession. Harvard University Press, Cambridge, MA.

Che, Y.-K., 2002. Joint Liability and Peer Monitoring under Group Lending. Contributions to Theoretical Economics 2, Article 3.

Crawford, V., Sobel, J., 1982. Strategic information transmission. Econometrica 50, 1431–1452.

Dessein, W., 2002. Authority and communication in organizations. Review of Economic Studies 69, 811–838.

Edwards, M.R., Ewen, A.J., 1996. How to manage performance and pay with 360-degree feedback. Compensation Benefits Review 28, 41–46.

Fedor, D.B., Bettenhausen, K.L., Davis, W., 1999. Peer reviews: employees' dual roles as raters and recipients. Group & Organization Management 24, 92–120.

Feess, E., Hege, U., 1998. Efficient liability rules for multi-party accidents with moral hazard. Journal of Institutional and Theoretical Economics 154, 422–450.

The Goldman Sachs Group, Inc., 2003. Proxy Statement – Annual Meeting of Shareholders, April 1, 2003. 85 Broad Street, New York.

Holmström, B., 1982. Moral hazard in teams. Bell Journal of Economics 13, 324–340.

Lipman, B., Seppi, D.J., 1995. Robust inference in communication games with partial provability. Journal of Economic Theory 66, 370–405.

Ma, C.T., 1988. Unique implementation of incentive contracts with many agents. Review of Economic Studies 55, 555–572.

Marx, L.M., Squintani, F., 2008. Individual accountability in teams with adverse selection. Working Paper. Available at http://faculty.fuqua.duke.edu/~marx/bio/papers.html.

May, G.L., Gueldenzoph, L.E., 2006. The effect of social style on peer evaluation ratings in project teams. Journal of Business Communication 43, 4–20.

Miller, N.H., 1997. Efficiency in partnerships with joint monitoring. Journal of Economic Theory 77, 285–299.

Müller, H.M., Wärneryd, K., 2001. Inside versus outside ownership: a political theory of the firm. RAND Journal of Economics 32, 527–541.

Nalbantian, H.R., Schotter, A., 1997. Productivity under group incentives: an experimental study. The American Economic Review 87, 314–341.

Prat, A., Rustichini, A., 2003. Games played through agents. Econometrica 71, 989–1026.

Prendergast, C., 1993. A theory of yes men. American Economic Review 83, 757–770.

Segal, I., 1999. Contracting with externalities. Quarterly Journal of Economics 114, 337–388.

Segal, I., 2003. Coordination and discrimination in contracting with externalities: divide and conquer? Journal of Economic Theory 113, 147–327.