Rochester Center for

Economic Research

Individual Accountability in Teams

Leslie M. Marx and Francesco Squintani

Working Paper No. 494
June 2002

UNIVERSITY OF
ROCHESTER

Individual Accountability in Teams

Leslie M. Marx and Francesco Squintani*
University of Rochester

May 2002

Abstract

We consider a team production problem in which the principal observes only the group output and not individual effort and in which the principal can only penalize an agent for poor performance if she has verifiable evidence that the agent in question did not fulfill his job assignment. In this environment, agents have an incentive to shirk. However, we show that by including monitoring in the agents' job assignments, the principal induces the agents to exert effort and achieves the first-best. In particular, even though equilibrium job assignments include monitoring, this serves only to provide incentives for effort, and agents do not engage in wasteful monitoring in equilibrium.

*We are grateful to Stephen Coate, John Geanakoplos, Glenn MacDonald, Marco Pagano, and Michael Riordan for valuable conversations, and we thank seminar participants at Columbia University. The authors can be reached at: Simon School of Business and Department of Economics, University of Rochester, Rochester, NY 14627, marx@simon.rochester.edu, sqnt@troi.cc.rochester.edu.

1 Introduction

Performance incentives often require that a principal be able to compensate agents differentially based on performance. To achieve high performance, an employer must be able to penalize a shirking employee; to ensure safety, a firm must be able to penalize workers who violate safety regulations. However, it is often the case that firms cannot apply punishments across groups of individuals when only one is at fault.[1] A firm may be restricted from firing an entire team of workers when it is clear that the team is under performing or that a safety rule has been violated by one of the team members. Before applying sanctions, the firm may have to establish individual guilt.

There are many examples of unionized and non-unionized firms (including universities) 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. For example, firms employing members of the United Auto Workers are severely restricted in their ability to sanction union members,[2] and many universities and government agencies require extensive reporting and evidence before a staff member can be dismissed.[3] Italian labor law requires that in order to sanction

[1] Such penalties may be considered punitive. According to Feess and Hege (1997), “In virtually all legal systems outside the United States, punitive damages are either excluded or play a very minor role. Even in the US, punitive damages are normally restricted to cases of reckless conduct, e.g. drunken driving.” See also Shavell (1987), Landes and Posner (1987), Posner (1986), and Polinsky and Che (1991).

[2] 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 Company against a member of the UAW can be appealed through a so-called Grievance Procedure (Art. VII, section 5). As a result, Ford Company cannot sanction shirking behavior unless it can establish individual guilt. Indeed, Ford Company cannot even demote employees who are chronically absent, unless it can show the “seriousness of the absenteism and its impact on the Company’s ability to compete effectively.” As well as the Grievance Procedure, an employee may appeal any sanction pertaining to poor attendance to an impartial Appeals Board (Appendix L).

[3] The University of Rochester personnel policies and procedures on termination (Policy 136) allows 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.

or fire an employee, the employer must show that the employee has failed to fulfill a job assignment or violated the firm's regulations.[4] The burden falls on the employer to prove an individual has violated the rules.

It is important to consider the effect of the legal environment on contract design because a contract, while theoretically optimal, may be useless if it will be dismissed in court. In this paper, we consider an environment in which the incentives a principal can provide for agents to follow instructions (or complete a job assignment) are restricted to penalties that can be imposed only if the principal can prove that a particular agent did not follow instructions. Thus, we assume the liability principle under which an individual may be held liable for a contract violation only if he is identified as the party breaching the contract.[5]

In this environment, strong incentives can be provided if output can be verified at an individual level. However, problems arise when only joint output can be verified.[6] Under joint liability, the principal can penalize all members of the team if the

[4]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 her 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.

[5]Research in law and economics on cases with multiple defendants considers whether legal rules that allocate costs of accidents on the basis of observed actions can induce agents to make efficient choices. Shavell (1987) provides a summary of this literature, which shows that liability rules typically exist that implement the first-best when actions are observable. See also Posner (1986, pp.151–152), Landes and Posner (1987, pp.123–131), and Kornhauser and Revesz (1989). When the court is fully informed, Kornhauser and Revesz (1989) and Arlen (1992) show that a variety of negligence rules implement the first-best. However, Feess and Hege (1998) show that an efficient liability rule may not exist when multiple tortfeasors interact in a non-separable way, there is imperfect information about their actions, and punitive damages are not feasible. See Emons (1990) and Green (1976) on the case of additive separability. See Emons and Sobel (1991) on the role of unobservable avoidance costs and environments in which agents are not identical. Feess and Hege (1997) show that basing liability on insurance policies is sufficient to obtain an efficient adjudication when there exist insurance contracts establishing proper incentives for the agents. An equal splitting rule is applied if both are insured, but an agent is fully liable for total damage if not insured.

[6]The seminal paper on team incentives is Holmström (1982), who shows that an efficient sharing rule for the team's output respecting ex-post budget-balance does not exist. Extending this work, Legros and Matsushima (1991) consider asymmetry and derive conditions for the existence of efficient sharing rules, and Legros and Matthews (1993) consider approximately efficient solutions.

performance of the team is poor, allowing the first-best to be achieved. However, in a regime of individual liability, it may not be possible, or it may be too costly, for the principal to provide incentives sufficient to induce agents to follow instructions. Without verifiable evidence of individual guilt, the principal is unable to punish any individual involved in the joint production process even if her observation of the joint output indicates that one or more of the agents did not follow instructions, or she can do this only by bearing significant legal costs.

We show that in this environment, the principal can obtain the first-best outcome if it is possible for agents to monitor each other's performance.[7] In this case, the principal can include monitoring of other agents as part of her instructions, with the penalty for not following instructions extending to the failure to monitor. For example, during a team project, agents may have the capacity to monitor each other even though the principal cannot, and if an agent observes shirking by a teammate, he can report it to the principal, at which point the principal can intervene and confirm the report. If the project is completed with no reports of shirking from agent i, the principal must conclude that either agent i failed to monitor his teammates or agent i did monitor his teammates and none of them shirked.

Notice that monitoring itself is not productive, and so the cost of monitoring is a deadweight loss. Thus, the first-best requires that no monitoring be performed. Interestingly, we show that this is the case—agents do not monitor in equilibrium. In particular, we show that the principal can use a monitoring requirement to construct incentives that induce agents to exert effort even though in equilibrium they do not monitor one another. The principal instructs agents to produce high output 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 produced no report of shirking by others, then it is verifiable that the agent either shirked himself or failed to monitor

[7]Nalbantian and Schotter (1997) consider the role of monitoring on group performance in an experimental setting. They find monitoring elicits high effort from workers in some circumstances. Lizzeri, Meyer and Persico (2001) study the incentives induced by a tournament-like mechanism that includes agents' monitoring.

the other agents.

In many business settings, supervisors frequently ask that employees evaluate the performance of their colleagues, especially when the employees work as a team. If taken seriously, such evaluations can be costly to produce, but we suggest that even if they are taken lightly, they can be a valuable part of workers' incentives to exert effort. Another real-life example in which our proposed mechanism applies comes from the academic world. To deter cheating, a university must be able to penalize students who cheat. However, it is often difficult to observe directly whether a student cheats, and if two or more students turn in identical exam answers, it may be impossible to establish which student copied from which. The first-best can be achieved by instructing students not only not to cheat, but also to take steps to avoid any other student's cheating from their work. Students are penalized for failing to take such steps only if there is evidence cheating has occurred.

A remarkable theoretical feature of our analysis is that the first-best is achieved in equilibrium (high effort and no monitoring) even though it cannot be achieved if the principal announces a job assignment of high effort and no monitoring. Without the threat of monitoring, agents choose low effort. But by assigning monitoring, the principal introduces the threat that agents will be penalized for not monitoring, giving them the incentive to exert high effort. In other words, the first-best can only be achieved by assigning second-best actions. The agents violate the principal's assignment in equilibrium, but in violating it, they allow the first-best to be achieved.[8]

One interesting feature of our equilibrium is that, although team punishments are not allowed, in the off-equilibrium case in which one agent does not exert high effort, all agents are penalized for not monitoring. This occurs because the principal can independently conclude that each of the agents has failed to fulfill the assignment, either because he has not exerted effort, or because he has not monitored his team-

[8] In a general game-theoretic framework, Squintani (2002) provides general results for when the players can restrict attention to contracts that are not violated in equilibrium. It is found that the crucial issue is whether a transitivity property is satisfied by the relation between the players' choices and the contractual transfers enforceable in court.

mates. Under joint liability, it would be possible to punish the entire team if the principal could show that at least one agent violated the job assignment. In other words, even though team-level penalties are not permitted, the inclusion of the monitoring requirement in the agents' job assignments effectively restores such penalties and so allows the first-best to be achieved.

2 Model

We consider a model with one principal and two agents. In Section 4, we extend the analysis to allow more than two agents. The agents have available to them a joint production process. For simplicity, we assume that each agent i chooses an action $a_i \in \{0, 1\}$. Thus, each agent either 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 ($\pi(a_1, a_2) = \ell$). 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. To make the problem meaningful, we assume $c_1 + c_2 < \pi(1, 1)$ so that the efficient outcome is for both agents to exert effort.

While the output π is observed by the principal and verifiable in court, 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 a verifiable report R_i of j's effort. Letting $m_i = 1$ if agent i monitors agent j and $m_i = 0$ if he does not, we consider two alternative assumptions regarding the relation among m_i, a_j, and R_i.[9] We view an agent's monitoring decision as affecting the type of evaluation or report that he will be able to submit to the

[9] When there are only two agents, as in this section and the next, all results hold under both of the assumptions we consider, but in Section 4, where we allow more than two agents, the distinction matters.

principal. Our two alternative assumptions allow for differences in what can be inferred about an agent's monitoring choice from the inspection of his report.[10]

We first consider the case in which agent i's report reveals whether agent i monitored agent j or not, as well as revealing agent j's effort when agent i does monitor.

Assumption 1 *For any agents i and j,*

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

Under Assumption 1, if $R_i(m_i, a_j) = n$, then the principal can infer nothing about j's action, but it can verify that agent i did not monitor agent j. In any case, agent i's report reveals whether agent i monitored or not.

We also consider the alternative assumption that the principal cannot directly observe whether agent i monitored or not. Despite this, agent i's report allows the principal to verify that agent j has shirked if agent j shirks and agent i monitors agent j. The key difference relative to Assumption 1 is that if agent i's report does not allow the principal to verify that j has shirked, the principal cannot verify whether this is because agent j worked hard or because agent i did not monitor.

This assumption may be appropriate to describe, for instance, the inferences that follow from a hearing or trial. Suppose that the output is low and agent j is accused of shirking. The court will conclude that agent j is guilty of shirking, against the presumption that j has not shirked, and against j's own claims on his own effort, if and only if agent i has monitored j and his report contains hard evidence showing that j shirked. If the court cannot conclude that agent j is guilty of shirking, it is not revealed whether this is because agent j exerted effort or because agent i did not monitor him, and hence was not able to produce hard evidence showing that j shirked.

[10] It may also be the case that the principal can always observe an agent's monitoring choice, but that a court is unable to verify it from the inspection of the report.

Assumption 2 *For any agents i and j,*

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

Under Assumption 2, if $R_i(m_i, a_j) = n$, then it cannot be verified whether agent i chose $m_i = 0$ or whether agent i chose $m_i = 1$ and agent j chose $a_j = 1$.

The principal's payoff in the production game is $\pi(a_1, a_2)$. The principal makes job assignments for each agent i, consisting of an effort level $\hat{a}_i$ and a monitoring requirement $\hat{m}_i$. For example, agent i might be assigned the job of exerting effort on the joint project as well as monitoring agent j. Formally, a requirement is a list $(\hat{a}_1, \hat{m}_1; \hat{a}_2, \hat{m}_2)$, where for each i, $\hat{a}_i \in \{0, 1\}$ and $\hat{m}_i \in \{0, 1\}$.[11]

The principal imposes sanction p_i on agent i if and only if output is low and it is verifiable that agent i did not fulfill his job assignment.[12] Agent i can be held accountable only for his own actions.[13] If agent i is not penalized, his payoff is $-a_i c_i - m_i x_i$; if agent i is penalized, his payoff is $-a_i c_i - m_i x_i - p_i$. To make the punishment meaningful, we assume that $p_i > x_i + c_i$. We also assume that effort is more costly than monitoring, i.e., $c_i > x_i$.[14]

We consider a three-stage game. In the first stage, the principal simultaneously issues job assignments to the two agents. In stage two, the agents observe the principal's job assignments and then choose their effort levels and whether to monitor the other agent. In stage three, the principal observes the output of the joint production

[11] In our notation, $\hat{a}_i = 1$ $(\hat{m}_i = 1)$ stands for the assignment to choose $a_i = 1$ $(m_i = 1)$, and $\hat{a}_i = 0$ $(\hat{m}_i = 0)$ denotes the principal's choice not to restrict the choice a_i (m_i).

[12] In order to simplify the exposition, we do not model the principal's incentives in the game after production has taken place. One can assume a utility function for the principal in which she values her reputation. Then reputational concerns prevent the principal from renegotiating the agents' penalties. Even if side payments from agents were legally binding, then agents would prefer to accept penalties rather than make the side payments necessary to dissuade the principal from following through on the penalty.

[13] If this were not the case, then a punishment could be assigned to any agent when output is low, regardless of whether it can be verified in court that he has not fulfilled a job assignment. In such a case, the first-best is achieved by requiring high effort from both agents and committing to punish both in case of low output.

[14] This assumption is relaxed in the Section 4.

process and the reports and, if output is low, penalizes agents for whom there is verifiable evidence that they did not fulfill their job assignments. 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

We begin by establishing 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 output is low, then it is not verifiable whether an individual agent failed to fulfill the assignment. The results of this section hold under both Assumption 1 and Assumption 2.

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

Proof. Suppose the principal assigns $\hat{m}_1 = \hat{m}_2 = 0$. Pick any arbitrary i. If $\hat{a}_i = 0$, in addition to $\hat{m}_i = 0$, then agent i cannot be penalized, regardless of her choices. Since effort and monitoring are costly, agent i chooses $(a_i, m_i) = (0, 0)$. We are thus left to consider only the case in which $(\hat{a}_i, \hat{m}_i) = (1, 0)$. Since the monitoring choice is unrestricted by the assignment, agent i can be penalized only if the principal verifies that $a_i = 0$. As a result, we claim: Under Assumption 2, agent i is penalized if and only if $(a_i, m_j) = (0, 1)$; whereas under Assumption 1, agent i is penalized if and only if $(a_i, m_j) = (0, 1)$ or $(a_i, m_i, a_j) = (0, 1, 1)$.

As proof of the claim, note that both under Assumption 1 and Assumption 2, if $(a_i, m_j) = (0, 1)$, then $R_j = 0$, and hence the principal verifies that $a_i = 0$ and agent i can be penalized. In addition, under Assumption 1, if $(a_i, m_i, a_j) = (0, 1, 1)$, it follows that output is low and that $R_i = 1$. Since output is low, the principal can verify that $(a_i, a_j) \neq (1,1)$. Since $R_i = 1$ the principal can verify that $a_j = 1$, hence that $a_i \neq 1$, and so agent i can be penalized.

Consider Assumption 2. For any $(a_i, m_j) \neq (0,1)$, it follows that $R_j = n$. Regardless of whether $R_i = 0$ or $R_i = n$, the observation of low output and $R_j = n$ could be the result of high effort by agent i and shirking by agent j. Thus, it is not verifiable that agent i did not complete his assignment and so agent i cannot be penalized.

Consider Assumption 1. Suppose $(a_i, m_j) \neq (0,1)$ and $(a_i, m_i, a_j) \neq (0,1,1)$. If $(a_i, m_j) = (1,1)$, it follows that $R_j = 1$, and hence the principal verifies that $a_i = 1$, and agent i cannot be punished. If $(a_i, a_j) = (1,1)$, output is high and agent i is not punished. For each remaining profile $(\mathbf{a}, \mathbf{m})$, it follows that $R_j = n$ and $R_i \neq 1$. Regardless of whether $R_i = 0$ or $R_i = n$, these observations together with low output could be the result of $(a_i, a_j) = (1,0)$. Since it is not verifiable that $a_i \neq 1$, agent i cannot be penalized. This completes the proof of the claim.

The claim implies that agent i's choice m_i either does not affect whether he is penalized, or, when $(a_i, a_j) = (0,1)$ and Assumption 1 holds, choosing $m_i = 1$ causes the agent to be penalized when he would not have had he chosen $m_i = 0$. Since monitoring is costly, it follows that $m_i = 0$ in equilibrium. This, together with the result that if $(\hat{a}_j, \hat{m}_j) = (0,0)$ then $(a_j, m_j) = (0,0)$, implies that $m_i = m_j = 0$ in equilibrium regardless of $\hat{a}_j$. This, together with our claim, implies that (both under Assumption 1 and Assumption 2) agent i cannot be penalized for $a_i = 0$. Since effort is costly, it follows that agent i chooses $a_i = 0$ when $(\hat{a}_i, \hat{m}_i) = (1,0)$. Q.E.D.

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 Proposition 1, however, it appears that only a second-best outcome may be achieved—successful joint production takes place only if the principal requires the agents engage in costly monitoring of each other.[15]

The intuition that only a second-best solution is achievable is overturned in the main result of this section. Proposition 2 below shows that if the principal requires both agents to work hard and to monitor each other, then in equilibrium the agents work hard *without* cross-checking each other, so that the first-best is achieved. The job assignments that correspond to the second-best solution induce the first-best outcome.

As well as being meaningful for the standpoint of a social planner, Proposition 2 implies that the principal need not compensate agents for the cost of monitoring, even though monitoring is included in the job assignment. When the agents choose whether to participate, they anticipate that in equilibrium they will not monitor each other. As a result they participate as long as their compensation equals the disutility from effort. If they monitored each other in equilibrium, the individual rationality constraint would be more restrictive.

Proposition 2 *In equilibrium, if the principal chooses the assignment* $(\hat{a}_i, \hat{m}_i) = (1, 1)$ *for each* i*, then each agent* i *chooses* $(a_i, m_i) = (1, 0)$*, and the first-best is achieved.*

In order to prove Proposition 2, we begin with a lemma that shows that if the principal requires the two agents to work hard and monitor each other, then the strategy profile in which they both exert effort without monitoring is an equilibrium. To see this, note that if an 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

[15] This intuition is not novel. For instance, it drives the analysis of the peer monitoring second-best solution of moral hazard in teams problem (see, e.g., Arnott and Stiglitz (1991)).

i exerted effort and monitored agent j, then agent j must not have exerted effort, in which case agent i would have produced a report $R_i = 0$. Since this is not the case, the principal verifies that agent i either did not exert effort or failed to monitor the other agent. In either case, he has not fulfilled the job assignment and can be penalized.

Lemma 1 *The profile* $(a_i, m_i) = (1, 0)$ *for each* i *is an equilibrium of the subgame induced by the assignment* $(\hat{a}_i, \hat{m}_i) = (1, 1)$ *for each* i.

Proof. 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 (under either Assumption 1 or 2) 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 $-c_i$. A deviation by agent i to $(a_i, m_i) = (1, 1)$ gives payoff $-c_i - x_i$, 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 penalizes agent i, making the deviation unprofitable. If agent i deviates and chooses $(a_i, m_i) = (0, 1)$, then output is low and either $R_i = 1$ (under Assumption 1) or $R_i = n$ (under Assumption 2), 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. Q.E.D.

The proof of Proposition 2 is concluded in the Appendix by showing that the strategy profile in which both agents exert effort without monitoring each other is the unique equilibrium outcome in the subgame induced by any assignment such that $(\hat{a}_i, \hat{m}_i) = (1, 1)$ for both i.

Proposition 2 shows that if the agents are required to monitor each other's effort, then output will be high and the first-best will be achieved. In fact, the next result shows that the principal can induce agents to exert effort only if she requires them to exert effort and to monitor each other.[16]

Proposition 3 *High output can be achieved with probability one in equilibrium only if* $(\hat{a}_i, \hat{m}_i) = (1, 1)$ *for each* i.

Proof. See the Appendix.

Proposition 3 implies that the principal must require agents monitor each other to guarantee high output.[17]

4 Large Teams

In this section we extend the analysis of the previous section to the case of teams with more than two agents. We let I be the set of agents and assume $\#I \geq 2$ and finite. Again, we assume that each agent i chooses an action $a_i \in \{0, 1\}$, with cost c_i, and we maintain the assumption that output may be either high or low. Specifically, we say that $\pi(\mathbf{a}) = \ell$ if $a_i = 0$, for some i, and that $\pi(\mathbf{a}) = h$ if $a_i = 1$ for all i. Each agent i can monitor each agent j, $j \neq i$, at cost $x_{ij} > 0$. We let $m_{ij} = 1$ denote the

[16] In the above analysis, we have assumed that the only verifiable information contained in the output produced by the agents is whether output was high or low. We can show that the same results hold if upon observing the output, the principal can verify the number of workers who have shirked, as long she cannot determine their identity.

[17] We have not considered the possibility that agents can provide verification of their own effort (self-monitoring). One simple job assignment that induces high effort is the requirement that each agent exert effort and produce verifiable evidence of his own effort. When we view the verifiability as related to the outcome of a trial; however, this job assignment induces a trial that is equivalent to one in which an agent is guilty unless he proves that he has exerted effort, against the presumption that he has not. This is counter to basic legal principles and is also not plausible in the economic environments of interest.

decision by agent i to monitor agent j, and we let $m_{ij} = 0$ denote the decision not to monitor agent j. As before, we assume that for any i and any $j \neq i$, $c_i + x_{ij} < p_i$. However, we allow for the possibility that $c_i + \sum_{j \neq i} x_{ij} > p_i$. Assumptions 1 and 2 are immediately adapted to this environment.

As in the case of two agents, the principal cannot achieve high effort by means of job assignments that do not require the agents to monitor each other's effort. Unlike the two-agent case, the extent of the monitoring requirements necessary to achieve the first-best depends on whether we follow Assumption 1 or 2. Specifically we show that the monitoring requirements inducing high output are more demanding under Assumption 2. In fact, under Assumption 2, high output is achieved only if the principal requires every agent to monitor every other agent.

Proposition 4 *Under Assumption* 2*, high output is achieved with probability one in equilibrium only if the principal's job assignment to each agent* i *is* $(\hat{a}_i, \hat{\mathbf{m}}_i) = (1, \mathbf{1})$.

Proof. See the Appendix.

The above result shows that in equilibrium the principal assigns each agent the task to work hard in the joint production process and monitor each of his teammates. 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 *Under Assumption* 2*, if the principal assigns* $(\hat{a}_i, \hat{\mathbf{m}}_i) = (1, \mathbf{1})$ *for all* i*, then in the unique equilibrium each agent* i *chooses* $(a_i, \mathbf{m}_i) = (1, \mathbf{0})$*, and hence the first-best is achieved.*

Proof. See the Appendix.

While it may seem infeasible to require each agent to monitor any 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 parameters x_{ij}. An interesting question is whether assignments that impose large costs on agents, should they be fulfilled, are legal. If they are not, we immediately obtain an impossibility result for the case of large teams. For concreteness, suppose that for any agent i, it is illegal to assign $(\hat{a}_i, \hat{\mathbf{m}}_i)$ inducing a cost

$$\hat{a}_i c_i + \sum_{j \neq i} \hat{m}_{ij} x_{ij} > p_i. \tag{1}$$

We refer to assignments satisfying (1) as "overburdening" assignments.

Proposition 6 *Under Assumption* 2*, if $x_{ij} = x_i$ for all $j \neq i$ and overburdening assignments are illegal, then for sufficiently large teams, high output cannot be achieved.*

Proof. By Proposition 4, high output can only be achieved is the job assignment to each agent i is $(\hat{a}_i, \hat{\mathbf{m}}_i) = (1, \mathbf{1})$. For each agent i, following this assignment results in payoff $-c_i - \sum_{j \neq i} x_{ij} = -c_i - (n-1)x_i$. If agent i does not exert effort, he is penalized, and so his payoff is bounded above by $-p_i$. For all i, there exists $\bar{n}_i$ such that for all $n \geq \bar{n}_i$, $-c_i - (n-1)x_i \leq -p_i$. Q.E.D.

If it can be observed whether a report required monitoring effort or not, we need not require 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 this case, to avoid the penalty, each agent prefers to exert effort rather than monitor.

In the following result, for any agent i, we let $M_i(\hat{\mathbf{m}}_i)$ be the set of agents that agent i is assigned to monitor, i.e., given assignment $\hat{\mathbf{m}}_i$, $M_i(\hat{\mathbf{m}}_i) \equiv \{j \mid \hat{m}_{ij} = 1\}$.

Proposition 7 *Under Assumption* 1*, there is an equilibrium in which high output is achieved if and only if for each i, $(\hat{a}_i, \hat{\mathbf{m}}_i)$ is such that $\hat{a}_i = 1$ and either $\sum_{j \in M_i(\hat{\mathbf{m}}_i)} x_{ij} \geq c_i$ or $M_i(\hat{\mathbf{m}}_i) = I \backslash \{i\}$. In equilibrium, if output is high, then there is no monitoring.*

Proof. See the Appendix.

The above result identifies all assignments that allow the principal to achieve high output in equilibrium. However, it may the case that some of those assignments also admit a low-output equilibrium. To see this, suppose that there two agents that are assigned such a heavy task that they prefer the penalty. If they both exerted high effort, they would not be penalized, as output would be high. Still, they can miscoordinate and exert low effort in the belief that the other teammate will shirk. In such cases, agents are penalized in equilibrium. Given a choice, one might expect the principal to choose an assignment such that high output is the unique equilibrium outcome.

In Proposition 8 below, we characterize assignments that admit multiple equilibria. Let $P(\hat{\mathbf{m}}) \equiv \{i \mid \sum_{j \in M_i(\hat{\mathbf{m}})} x_{ij} + c_i \geq p_i\}$ be the set of agents with overburdening assignments, and let $B(\hat{\mathbf{m}}) \equiv \{j \mid \hat{m}_{ij} = 0 \text{ for all } i \notin P(\hat{\mathbf{m}})\}$ be the set of agents that are not assigned to be monitored by any agent outside of $P(\hat{\mathbf{m}})$. Thus, agents in $B(\hat{\mathbf{m}})$ are those who, if they are monitored at all, are monitored by agents with overburdening assignments.

Proposition 8 *Under Assumption* 1, *any assignment* $(\hat{\mathbf{a}}, \hat{\mathbf{m}})$ *that induces a high-output equilibrium, also induces a low-output equilibrium if and only if* $\#(P(\hat{\mathbf{m}}) \cup B(\hat{\mathbf{m}})) \geq 2$.

Proof. See the Appendix.

Proposition 8 implies that if an assignment induces a high-output equilibrium, it also allows the agents to miscoordinate on a low-output equilibrium when there are at least two agents who are either overburdened or who are not assigned to be monitored by a non-overburdened agent.

If overburdening assignment are illegal, the above result yields the following corollary.

Corollary 1 *Under Assumption* 1, *if a non-overburdening assignment induces a high-output equilibrium, this assignment also induces a low-output equilibrium if and only if there are at least two agents who are not assigned to be monitored by any other agent.*

Corollary 1 says that in the absence of overburdening assignments, it is enough that at most one agent is not to be monitored by anyone in order to avoid the low-output equilibrium.

The above results show that the effect of ruling out overburdening assignments depends on whether Assumption 1 or 2 holds. We have seen in Proposition 6 that, under Assumption 2, high output cannot be achieved at all if overburdening assignments are illegal. In contrast, under Assumption 1, Proposition 7 shows that high output can be achieved by assigning $\hat{\mathbf{m}}$ such that for all i,

$$c_i \leq \sum_{j \in M_i(\hat{\mathbf{m}}_i)} x_{ij} < p_i - c_i. \tag{2}$$

Moreover, Proposition 8 shows that among the job assignments that induce multiple equilibria are those that require overburdening tasks by at least two agents. Finally, Corollary 1 shows if overburdening assignments are illegal, then the principal can ensure that agents work hard in equilibrium with any assignment $\hat{\mathbf{m}}$ that satisfies condition (2) and that requires that all but at most one agent is monitored by at least one teammate.

5 Conclusion

Our model is deterministic, with each action profile generating a unique output and evidence profile. An agent's failure to fulfill his job assignment is verifiable if it can be determined based on logical inferences about the actions taken given the available information. The model can be reformulated in stochastic terms, with inferences drawn based on the relative likelihood of events, but we expect the basic results would not be affected. It seems reasonable to think that the same results would arise for the relevant parametric cases.[18] In particular, since the payoff structure does not exploit ties, the game is essentially generic, and so we expect our results are robust to the introduction of noise in the parameters.

In addition, the model presented in this paper is discrete, in the sense that effort, monitoring, and output are discrete (binary) variables. These assumptions greatly simplify the analysis and the exposition. However, the structure of the model suggests that similar predictions would continue to hold in a continuous version of the model.

[18] This extension may require the analysis of mixed-strategy equilibria (see Bagwell (1995), and Hjurkens and van Damme (1997)).

A Appendix: Proofs

Proof of Proposition 2. Suppose $(\hat{a}_i, \hat{m}_i) = (1, 1)$ for each i. Proceeding by contradiction, suppose that in an equilibrium an arbitrary agent j chooses $a_j = 0$. In such a case, the output is low, and we show that the best response of agent $i \neq j$ must be such that $m_i = 1$, and hence $R_i = 0$. If agent i chooses $(a_i, m_i) = (1, 1)$, agent i cannot be penalized: in this case, (depending on m_j and whether we use Assumption 1 or 2) either $R_j = 1$ or $R_j = n$, so the principal cannot rule out $(a_i, m_i) = (1, 1)$. If instead agent i chooses $m_i = 0$, proceeding as in the proof of Lemma 1, he is penalized. Since $p_i > x_i + c_i$, it follows that agent i prefers $(a_i, m_i) = (1, 1)$, rather than any profile including $m_i = 0$ if agent j chooses $a_j = 0$.

Suppose now that $m_i = 1$. If $a_j = 0$, then $R_i = 0$, and it is verifiable that $a_j = 0 \neq \hat{a}_j$. If instead agent j chooses $(a_j, m_j) = (1, 1)$, then he cannot be penalized. This implies that j's best response to $m_i = 1$ must be such that $a_j = 1$.

The above arguments show that in equilibrium it must be the case that $a_1 = a_2 = 1$, and thus that $\pi(a_1, a_2) = h$, regardless of m_1 and m_2. Suppose now by contradiction that there is an equilibrium in which $m_i = 1$ for some i. Since in such an equilibrium $\pi(a_1, a_2) = h$, agent i is not penalized, regardless of m_i. Hence i will deviate and choose $m_i = 0$. Q.E.D.

Proof of Proposition 3. When 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 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).

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 there cannot be an equilibrium in which $a_i = a_j = 1$ with probability one.

In conclusion, the only subgame for which $\Pr(\pi(\mathbf{a}) = h) = 1$ in equilibrium is the one following the assignment $(\hat{a}_i, \hat{m}_i) = (1, 1)$ for each agent i. Q.E.D.

Proof of Proposition 4. First note that, regardless of the assignment $(\hat{\mathbf{a}}, \hat{\mathbf{m}})$, there cannot be any equilibrium in which $\pi(\mathbf{a}) = h$ and some agent i chooses $m_{ij} = 1$ for some $j \neq i$. Since $\pi(\mathbf{a}) = h$, no agent is penalized, and since $x_{ij} > 0$, the choice $m_{ij} = 1$ cannot be part of a best response. So we are left to show that the profile $(a_j, \mathbf{m}_j) = (1, \mathbf{0})$ for all j cannot be an equilibrium in the subgame induced by any assignment where for some agent i, $\hat{a}_i = 0$, or $\hat{m}_{ij} = 0$ for some $j \neq i$. Suppose first that for some agent i, $\hat{a}_i = 0$. Then, regardless of $\hat{\mathbf{m}}$, the profile $(a_k, \mathbf{m}_k) = (1, \mathbf{0})$ for all k cannot be an equilibrium, because agent i deviates and chooses $(a_i, \mathbf{m}_i) = (0, \mathbf{0})$. Since for any $k \neq i$, the principal cannot rule out $a_k = 1$ 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. Second, consider any assignment such that $\hat{a}_i = 1$, but $\hat{m}_{ij} = 0$ for some $j \neq i$. The profile $(a_k, \mathbf{m}_k) = (1, \mathbf{0})$ for all k cannot be an equilibrium, because agent i can profitably deviate by choosing $(a_i, \mathbf{m}_i) = (0, \mathbf{0})$. In such a case, since $R_{ik} = n$ for all k, the principal cannot rule out $a_j = 0$, $a_i = 1$, and $\mathbf{m}_i = \hat{\mathbf{m}}_i$. Therefore, i cannot be penalized. Q.E.D.

Proof of Proposition 5. Consider the assignment $\hat{a}_i = 1$ and $\hat{m}_{ij} = 1$ for all i and $j \neq i$. First, we show that the profile $(a_i, \mathbf{m}_i) = (1, \mathbf{0})$ for all i is an equilibrium in the induced subgame. Suppose that each agent $j \neq i$ chooses $(a_j, \mathbf{m}_j) = (1, \mathbf{0})$. Any deviation such that $a_i = 1$ and $\mathbf{m}_i \neq \mathbf{0}$ is not profitable because $x_{ij} > 0$ for all j. Suppose that agent i deviates and chooses $a_i = 0$. As a result output is low, and so it is verifiable that $\mathbf{a} \neq \mathbf{1}$. Letting $j \neq i$, since $a_j = 1$, by Assumption 2 it must be that $R_{ij} = n$. Hence it is verifiable that $(a_j, m_{ij}) \neq (0, 1)$. Since this is true for all $j \neq i$, it is verifiable that either $m_{ij} = 0$ for some $j \neq i$, or $a_j = 1$ for all $j \neq i$. Thus, it is verifiable that either $a_i = 0$ or $m_{ij} = 0$ for some $j \neq i$, and hence agent i is penalized, making the deviation unprofitable.

Second, we show that the profile $(a_i, \mathbf{m}_i) = (1, \mathbf{0})$ for all i is the unique equilibrium in the subgame. We first see that any profile with $a_i = 1$ for all i, and $\mathbf{m}_i \neq \mathbf{0}$ for some i cannot be an equilibrium. In this case, output is high, so the principal does not penalize any agent, regardless of the monitoring choices. Since $x_{ij} > 0$ for any j, agent i will deviate and choose $\mathbf{m}_i = \mathbf{0}$. Thus, we are only left to show that there is no equilibrium in which $a_i = 0$ for some i. Proceeding by contradiction, suppose there is an equilibrium profile such that the set $J \equiv \{j \in I \mid a_j = 0\}$ is non-empty (implying that output is low).

Case 1. $m_{ij} = 1$ for some i and some $j \in J$. Then $R_{ij} = 0$, so it is verifiable that $a_j = 0 \neq \hat{a}_j$, and agent j is penalized. Suppose first that $J = \{j\}$. Then if agent j deviates and chooses $(a_j, \mathbf{m}_j) = (1, \mathbf{0})$, output is high and no agent is penalized. Since $p_j > c_j$, the deviation is profitable for agent j, a contradiction. Second, suppose that there exists $k \neq j$ such that $k \in J$. If j deviates by choosing $a_j = 1$, $m_{jk} = 1$, and $m_{jl} = 0$ for all $l \neq k$, he is not penalized because $R_{jk} = 0$, and so the principal cannot rule out $(a_j, \mathbf{m}_j) = (1, \mathbf{1})$. Hence agent j's payoff is $-c_j - x_{jk} > -p_j$, and the deviation is profitable, a contradiction.

Case 2. $m_{ij} = 0$ for all i and all $j \in J$. In this case, for all i and all $j \in J$, $R_{ij} = n$. In addition, for any agent $t \notin J$, $a_t = 1$, so it follows that for all i and all $t \notin J$, $R_{it} = n$. Pick an arbitrary i. For all $\ell \neq i$, $R_{i\ell} = n$, so it is verifiable that $(a_i, \mathbf{m}_i) \neq (1, \mathbf{1})$, and hence i is penalized. Since i is arbitrary, every agent in J is also penalized. Pick any

$j \in J$. If $J = \{j\}$, agent j can profitably deviate by choosing $(a_j, \mathbf{m}_j) = (1, \mathbf{0})$, so that output is high and no agent is penalized, a contradiction. Suppose that there exists $k \neq j$ such that $k \in J$. If agent j deviates by choosing $a_j = 0$, $m_{jk} = 1$, and $m_{jl} = 0$ for all $l \neq k$, he is not penalized because $R_{jk} = 0$, and so the principal cannot rule out $(a_j, \mathbf{m}_j) = (1, \mathbf{1})$. Hence agent j's payoff is $-c_j - x_{jk} > -p_j$, and the deviation is profitable, a contradiction. Q.E.D.

Proof of Proposition 7. If high output is achieved, then the principal does not penalize agent i regardless of the monitoring choices, and hence there is no equilibrium in which $a_i = 1$ for all i, and $m_{ij} = 1$ for some i and j. Take the strategy profile where $a_i = 1$ for all i, and $m_{ij} = 0$ for all i and $j \neq i$. Each agent i has payoff $-c_i$. Suppose that agent i deviates and chooses $a_i = 0$. By the same argument as in Proposition 4, if $\hat{m}_{ij} = 1$ for all $j \neq i$, then agent i will not take any deviation such that $a_i = 0$. Suppose that $M_i(\hat{\mathbf{m}}_i) \neq I \backslash \{i\}$. Using Assumption 1, 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 $p_i > c_i$, this deters any deviation such that $a_i = 0$, $m_{ij} = 0$, 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 $-\sum_{j \in M_i(\hat{\mathbf{m}}_i)} x_{ij}$. If $\sum_{j \in M_i(\hat{\mathbf{m}}_i)} x_{ij} \geq c_i$, then any such deviation is deterred. If $\sum_{j \in M_i(\hat{\mathbf{m}}_i)} x_{ij} < c_i$, the deviation such that $a_i = 0$ and $m_{ij} = 1$ if and only if $\hat{m}_{ij} = 1$ is not deterred.

Proof of Proposition 8. We consider three cases.

Case 1. $\#P(\hat{\mathbf{m}}) \geq 2$. Suppose that the profile $(\mathbf{a}, \mathbf{m})$ is such that $(a_i, \mathbf{m}_i) = (0, \mathbf{0})$ for each $i \in P(\hat{\mathbf{m}})$. As a result, output is low. If agent i is not penalized, then he has no reason to deviate from $(a_i, \mathbf{m}_i) = (0, \mathbf{0})$. Suppose that agent i is penalized in equilibrium. Even if he unilaterally deviates and chooses $a_i = 1$, output will still be low. Agent i can avoid being penalized only if, in addition to $a_i = 1$, he also chooses $\mathbf{m}_i = \hat{\mathbf{m}}_i$. Since $i \in P(\hat{\mathbf{m}})$, it follows that $\sum_{j \in M_i(\hat{\mathbf{m}}_i)} x_{ij} + c_i \geq p_i$, and hence i will not deviate. This concludes the proof that if $\#P(\hat{\mathbf{m}}) \geq 2$, then there is a low-output equilibrium.

Case 2. $\#P(\hat{\mathbf{m}}) = 0$. A profile $(\mathbf{a}, \mathbf{m})$ such that some agent $i \notin P(\hat{\mathbf{m}})$ is penalized cannot be an equilibrium. Since $\sum_{j \in M_i(\hat{\mathbf{m}}_i)} x_{ij} + c_i < p_i$, agent i would fare better by deviating and choosing $a_i = 1$ and $m_{ij} = 1$ for all $j \in M_i(\hat{\mathbf{m}}_i)$. When output is low, any agent i who does not choose $m_{ij} = 1$ for all $j \in M_i(\hat{\mathbf{m}}_i)$ is penalized. Thus, if the profile $(\mathbf{a}, \mathbf{m})$ is an equilibrium such that output is low, it must be that $m_{ij} = 1$ for all $j \in M_i(\hat{\mathbf{m}}_i)$ and for all i. At the same time, any agent j such that $a_j = 0$ and $m_{ij} = 1$ for some i is penalized. Hence, if the profile $(\mathbf{a}, \mathbf{m})$ is an equilibrium such that output is low, it must be that $a_j = 1$ for all $j \notin B(\hat{\mathbf{m}})$. This implies that, for the case $\#P(\hat{\mathbf{m}}) = 0$, if $\#B(\hat{\mathbf{m}}) = 0$, then there is no low-output equilibrium.

Suppose $B(\hat{\mathbf{m}}) = \{i\}$ for some i. Then it must be the case that $a_j = 1$ for all $i \neq j$. By hypothesis, it must be the case that either $M_i(\hat{\mathbf{m}}_i) = I \backslash \{i\}$ or that $\sum_{j \in M_i(\hat{\mathbf{m}}_i)} x_{ij} \geq c_i$. In the second case, agent i will deviate and choose $(a_i, \mathbf{m}_i) = (1, \mathbf{0})$. As a result, output is high, agent i is not penalized, and his payoff is $-c_i$ which is greater than $-\sum_{j \in M_i(\hat{\mathbf{m}}_i)} x_{ij}$, the payoff for choosing $a_i = 0$ and $m_{ij} = 1$ for all $j \in M_i(\hat{\mathbf{m}}_i)$. In the first case, $M_i(\hat{\mathbf{m}}_i) = I \backslash \{i\}$. If agent i chooses $a_i = 0$ and $m_{ij} = 1$ for all $j \in M_i(\hat{\mathbf{m}}_i)$, then for all $j \neq i$, $(a_j, m_{ij}) = (1, 1)$. Since output is low and, under Assumption 1, $R_{ij} = 1$ for all $j \neq i$, it is verifiable that it is impossible that j has fulfilled his assignment. Thus, if $\#P(\hat{\mathbf{m}}) = 0$ and $\#B(\hat{\mathbf{m}}) = 1$, there is no low-output equilibrium.

Suppose $\#B(\hat{\mathbf{m}}) \geq 2$. Suppose that for all i, $m_{ij} = 1$ for all $j \in M_i(\hat{\mathbf{m}}_i)$ and $a_i = 0$ if and only if $i \in B(\hat{\mathbf{m}})$. Given any $i \in B(\hat{\mathbf{m}})$, by construction $m_{ji} = 0$ for any j, and since $\#B(\hat{\mathbf{m}}) \geq 2$, the principal cannot rule out $a_i = 1$, and so agent i is not penalized. If agent i deviates and plays $a_i = 1$, output will still be low, and hence he will be penalized unless he chooses $m_{ij} = 1$ for all $j \in M_i(\hat{\mathbf{m}}_i)$. Hence he will not deviate. Given any $i \notin B(\hat{\mathbf{m}})$, by construction $m_{ij} = 1$ for some j, and thus i will be penalized if he deviates and chooses $a_i = 0$. Thus, if $\#P(\hat{\mathbf{m}}) = 0$ and $\#B(\hat{\mathbf{m}}) \geq 2$, there is a low-output equilibrium in which all agents $i \in B(\hat{\mathbf{m}})$ choose $a_i = 0$ and all agents $i \notin B(\hat{\mathbf{m}})$ choose $a_i = 1$.

Case 3. $\#P(\hat{\mathbf{m}}) = 1$. Let $P(\hat{\mathbf{m}}) = \{j\}$. If the profile $(\mathbf{a}, \mathbf{m})$ is an equilibrium such that output is low, it must be the case that $m_{ik} = 1$ for all $k \in M_i(\hat{\mathbf{m}}_i)$ and for all $i \neq j$. Suppose that for all $i \neq j$, it the case that $i \notin B(\hat{\mathbf{m}})$, and hence that $a_i = 1$. If $j \notin B(\hat{\mathbf{m}})$, then agent j chooses $(a_j, \mathbf{m}_j) = (1, \mathbf{0})$. In such a case, output is high and j is not penalized. Agent j is instead penalized when $a_j = 0$ because, by definition of $B(\hat{\mathbf{m}})$, there is an agent i such that $m_{ij} = 1$. In the case that $B(\hat{\mathbf{m}}) = P(\hat{\mathbf{m}}) = \{j\}$, we apply the same argument presented for the case when $P(\hat{\mathbf{m}}) = \emptyset$ and $B(\hat{\mathbf{m}}) = \{i\}$ for some i, which implies that $a_j = 1$. Thus, if $B(\hat{\mathbf{m}}) = P(\hat{\mathbf{m}}) = \{j\}$, so that $\#\left(P(\hat{\mathbf{m}}) \cup B(\hat{\mathbf{m}})\right) = 1$, then there is no low-output equilibrium.

Finally, letting $P(\hat{\mathbf{m}}) = \{j\}$, suppose there exists $i \neq j$ such that $i \in B(\hat{\mathbf{m}})$. Say that $a_i = 0$ and $m_{ij} = 1$ for all $j \in M_i(\hat{\mathbf{m}}_i)$. As a result, output is low regardless of agent j's action a_j. If $j \in B(\hat{\mathbf{m}})$, then agent j chooses $a_j = 0$ and $m_{jk} = 1$ for all $k \in M_j(\hat{\mathbf{m}}_j)$. If $j \notin B(\hat{\mathbf{m}})$, since $j \in P(\hat{\mathbf{m}})$, agent j will prefer to choose $(a_j, \mathbf{m}_j) = (0, \mathbf{0})$ and be penalized, rather than $a_j = 1$ and $m_{jk} = 1$ for all $k \in M_j(\hat{\mathbf{m}}_j)$. As a result, since $i \in B(\hat{\mathbf{m}})$, it follows that for any $k \neq i$, it is the case that $m_{ki} = 0$. Since the principal cannot rule out $a_i = 1$, agent i is not penalized. Since $a_j = 0$, it follows that output is low and agent j is penalized, regardless of the choice a_j, and hence agent j is better off choosing $a_j = 0$. Thus, if $\#P(\hat{\mathbf{m}}) = 1$ and $\#\left(P(\hat{\mathbf{m}}) \cup B(\hat{\mathbf{m}})\right) \geq 2$, there is a low-output equilibrium. Q.E.D.

References

Arlen, Jennifer H. (1992), "Liability for Physical Injury when Injurers as well as Victims Suffer Losses," *Journal of Law, Economics and Organization* 8, 410–426.

Arnott Richard and Joseph Stiglitz (1991), "Moral Hazard and Nonmarket Institutions: Dysfunctional Crowding Out or Peer Monitoring?" *American Economic Review* 179-190

Bagwell, Kyle (1995), "Commitment and Observability in Games," *Games and Economic Behavior* 8, 271–280.

Emons, Winand (1990), "Efficient Liability Rules for an Economy with Non-identical Individuals," *Journal of Public Economics* 42, 89–104.

Emons, Winand and Joel Sobel (1991), "Liability Rules When Agents are not Identical," *Review of Economic Studies* 57, 375–390.

Feess, Eberhard and Ulrich Hege (1998), "Efficient Liability Rules for Multi-Party Accidents with Moral Hazard," *Journal of Institutional and Theoretical Economics* 154.

Feess, Eberhard and Ulrich Hege (1997), "The Role of Insurance in the Adjudication of Multi-Party Accidents," Working Paper, University of Frankfurt.

Green, Jerry. (1976), "On the Optimal Structure of Liability Laws," *Bell Journal of Economics* 7, 553–574.

Holmström, Bengt (1982), "Moral Hazard in Teams," *Bell Journal of Economics* 13, 324–340.

Hurkens, Sjaak, and Eric van Damme (1997), "Games with Imperfectly Observable Commitment," *Games and Economic Behavior* 21, 282–308.

Kornhauser, Lewis A. and Richard L. Revesz (1989), "Sharing Damages Among Multiple Tortfeasors," *Yale Law Journal* 98, 831–884.

Landes, William S. and Richard A. Posner (1987), The Economic Structure of Tort Law (Cambridge, MA: Harvard University Press).

Legros, Patrick and Hitoshi Matsushima (1991), "Efficiency in Partnerships," *Journal of Economic Theory* 55, 296–322.

Legros, Patrick and Steven A. Matthews (1993), "Efficient and Nearly-Efficient Partnerships," *Review of Economic Studies* 60, 599–611.

Lizzeri Alessandro, Margaret Meyer, and Nicola Persico (2001) "The Incentive and Sorting Effects of Interim Performance Evaluations" mimeo, Nuffield College, Oxford University

Nalbantian, Haig R. and Andrew Schotter (1997), "Productivity Under Group Incentives: An Experiment Study," *The American Economic Review* 87, 314–341.

Polinski, A. Mitchell and Yeon-Koo Che (1991), "Decoupling Liability: Optimal Incentives for Care and Litigation," *RAND Journal of Economics* 22, 562–570.

Posner, Richard A. (1986), *Economic Analysis of Law, 3rd ed.* (Boston: Little, Brown).

Shavell, Steven (1987), *Economic Analysis of Accident Law* (Cambridge, MA: Harvard University Press).

Squintani, Francesco (2001),"Verifiability, Generalized Information Correspondences, and Optimal Contracts," Working Paper, University of Rochester.