RAND Journal of Economics
Vol. 32, No. 1, Spring 2001
pp. 27–51

Optimal design of peer review and self-assessment schemes

Sandeep Baliga*

and

Tomas Sjöström**

A principal must decide whether or not to implement a project that originated with one of her employees. Several employees have information about the quality of the project. A successfully implemented project raises the inventor's chance of promotion, at his peer's expense, but a failed project ruins the inventor's career. An employee who has a relatively good reputation (and therefore is happy with the status quo) must be encouraged to promote new ideas. An employee who has a relatively bad reputation (and therefore wants to change the status quo) must be prevented from exaggerating the quality of new ideas. We study incentive-compatible and renegotiation-proof mechanisms, and we find that self-assessment (without any peer reports) is optimal.

1. Introduction

■ Career concerns create incentives for agents to misrepresent the quality of their own work, as well as the work of their colleagues. This makes it costly for the principal to extract information about a project. Indeed, it may be so costly that the principal is better off delegating to an agent the authority to determine whether or not a project should be implemented. In our model, one agent (the *inventor*) has developed a blueprint for a project. Both the inventor and a *peer* (but not the principal) receive some information about the quality of the blueprint. The principal has to decide whether or not to implement the project. The quality of the project is correlated with the talent of the inventor and can be observed by the principal if and only if the project is implemented. The agents' careers are at stake: at the end of the game the principal must promote either the inventor or the peer (but not both), and she would prefer to promote the most talented one. If the project is implemented and succeeds, she prefers the inventor; if it is implemented and fails, she prefers the peer. If the project is cancelled, her preferences are determined by her prior beliefs.

* Northwestern University; baliga@northwestern.edu.
**Pennsylvania State University; jts10@psu.edu.
This is a revised version of our working paper titled "Not Invented Here." We are grateful to Lars Stole and two anonymous referees for comments that significantly improved the article. We also thank Bengt Holmström, Eric Maskin, Suzanne Scotchmer, and many seminar participants. Any errors are our responsibility.

We identify four reasons for the agents to misrepresent their information. First, there is an incentive to support a colleague's *bad* projects to see him fail. Second, there is an incentive to denigrate a colleague's *good* projects to prevent him from getting ahead. Third, there is an incentive for an agent who is not close to a promotion to exaggerate the quality of his own work. Fourth, there is an incentive for an agent who is close to a promotion to suppress his own ideas in order not to have a big failure that ruins his reputation. We refer to these four effects as, respectively, *false praise, denigration, exaggeration,* and *false modesty.*[1] Our first concern is to study the conditions under which these four effects cause problems. An agent who is close to being promoted will be tempted to denigrate the work of others and to be falsely modest about his own work. But an agent who has a long way to promotion is tempted to be enthusiastic about all projects, both his own (exaggeration) and those of his colleagues (false praise), as he has little to lose.[2] Formally, different truth-telling constraints turn out to bind in the two different cases. Our second concern is to design an optimal contract in the presence of these incentive problems.

Authority is defined as the right to make a decision that affects the whole organization (Simon, 1951). Delegation is the deliberate allocation of authority down from a principal to an agent. We study different methods of collecting information and allocating authority within the organization:

(i) *self-assessment:* only the inventor makes an assessment of the project, so in effect he has been delegated the authority to have his project implemented or cancelled;

(ii) *peer review:* only the peer makes an assessment of the project, so in effect he has been delegated the authority to have the inventor's project implemented or cancelled;

(iii) *multiple reports:* both agents make assessments, and the principal retains final authority over the project implementation decision.

The revelation principle implies that if the principal can commit to a mechanism, then centralization can do anything that delegation can do, and usually it can do strictly more. This makes delegation of authority hard to explain. In our model, if the principal can commit, then she will ask for multiple reports. If both agents agree that the project is promising, it is implemented. If they agree it is not promising, it is not implemented. If the agents disagree, then the project is implemented, and the truth-telling constraints are relaxed by promoting the agent whose report best corresponds to the realized project quality. However, that may not be the agent most suitable for a promotion. Anything that happens after a disagreement has zero cost *ex ante* because in equilibrium the agents will not disagree, but it seems unrealistic that the principal can commit not to renegotiate outcomes in which the ‘‘wrong’’ agent is promoted. Thus, we will impose a renegotiation-proofness constraint. This constraint implies that if the project is implemented, then the inventor must be promoted if and only if the project succeeds. Since the principal cannot freely choose her promotion policy following a disagreement,

[1] In the sociology literature, the Not Invented Here (NIH) syndrome corresponds to what we call denigration. Coleman (1990, p. 443–444) explains it as follows: ‘‘the success or failure of others' ideas provides a benchmark for evaluating one's own performance. By demonstrating the defects in another's idea, one justifies not having had the idea oneself; by allowing the idea's potential to be realized, one would be relatively worse off, because that would raise the standard for evaluation of one's own work.''

[2] Gamache and Kuhn (1989, p. 33) report on research suggesting that ‘‘high risk taking described entry level employees. They had little . . . personal stakes in the organization and, consequently, felt they had little or nothing to lose. [T]he lowest risk taking was an arena peopled by middle management. These folks had a large career investment in the organization, essentially no security, lots of peer group pressure and competition and a lot to lose.''

the value of collecting multiple reports is reduced and delegation becomes more appealing.

First, consider the case when the inventor is behind in his career and needs a successfully implemented project to be promoted. In this case, the peer is tempted to denigrate the project, to prevent it from being implemented, while the inventor is tempted to exaggerate its quality. The binding truth-telling constraints are the ‘‘no-denigration’’ and ‘‘no-exaggeration’’ constraints. If the inventor claims the project is promising while the peer disagrees, then the principal should implement the project and promote the inventor if and only if the project succeeds. This policy relaxes the truth-telling constraints and is renegotiation-proof. Moreover, the negative peer report changes neither the principal’s decision to go ahead with the project nor the promotion policy. Therefore, the optimal contract is equivalent to a policy of self-assessment. A policy of peer review would do strictly worse. Intuitively, the principal must prevent denigration and exaggeration. Exaggeration can at most persuade the principal to implement a less promising project that is likely to fail. Because of this, exaggeration is fairly easy to prevent. It is much more difficult to prevent denigration, since if a promising project is stopped because of an unfair peer report, the principal will never learn the project’s true quality. This asymmetry suggests that self-assessment dominates. Still, the fact that exaggeration has to be prevented leads to certain distortions compared to the first best. We characterize the optimal way to use promotion and wage policies to stop exaggeration from occurring.

Second, consider the case where the inventor is ahead. His *ex ante* reputation is so good that he would be promoted even if his project were cancelled (but not if it were implemented and failed). This case is a bit more complicated than the previous one. Now, if the inventor has a good signal about his own project, he is tempted to indulge in ‘‘false modesty’’ (he is concerned about the probability that his signal is wrong and his project will fail). If the peer receives a bad signal about the project, then he thinks it is likely to fail, which would raise his own chances of promotion. Therefore, the peer is tempted to indulge in ‘‘false praise.’’ The two binding truth-telling constraints are the ‘‘no-false-modesty’’ and the ‘‘no-false-praise’’ constraints. If the peer claims the project looks promising but the inventor disagrees, then to relax the truth-telling constraints the principal should implement the project and promote the inventor with a higher probability if the project fails than if it succeeds (since the inventor was the one who sent a negative report). However, renegotiation-proofness forces the principal to promote the inventor if the project succeeds but not when it fails. Since this gives the wrong incentives to the agents, the principal is better off cancelling the project when her employees disagree in this way. This in turn implies that the inventor can have his project terminated simply by reporting a bad signal, irrespective of the peer’s opinion. Therefore, the optimal renegotiation-proof contract is again equivalent to a policy of self-assessment. With a self-assessment policy, false modesty has to be prevented. We characterize the optimal wage and promotion policies that will do so.

We can summarize the two cases by saying that self-assessment is always optimal in the class of incentive-compatible (IC) and renegotiation-proof mechanisms. In effect, the inventor should be given the authority to determine if his project should be implemented or not. Our results are derived allowing for the most general class of mechanisms. For an ‘‘incomplete-contracting’’ approach to delegation, see Aghion and Tirole (1997).

Holmström (1982) presented the first analysis of relative performance evaluation in a team (see also Levitt, 1995). Holmström (1999) analyzes the impact of career concerns on incentives to work. Our focus is instead on adverse selection and on the optimal methods of gathering information. A number of articles consider a supervisor's evaluation of a subordinate (Tirole, 1992; Prendergast and Topel, 1996). They focus on rather different issues than does our article, such as collusion and the effect of favoritism on optimal performance evaluation. Baker, Gibbons, and Murphy (1994) present several models on the use of subjective performance measures in optimal incentive schemes and present results on the substitutability and complementarity of (objective) explicit and (subjective) implicit incentive schemes. The fact that incentive schemes and promotion policies can cause agents to behave destructively to further their own careers at the expense of others has been pointed out by Lazear (1989), Milgrom (1988), and Itoh (1991), although in this literature there is no adverse selection and no predictions about optimal information systems. Rotemberg and Saloner (1995) analyze a different kind of conflict within the firm: the production and sales departments disagree about business strategy and try to present arguments that damage the other side's position. Two interesting articles that are somewhat closer to our setup are Carmichael (1988), who shows the optimality of tenure contracts, and Levitt and Snyder (1997), who analyze a principal-agent model in which the agent is protected by limited liability and receives a signal as well as exerts effort. In the latter model, the agent is tempted to exaggerate to get the high wage he earns if the project is successful, so he has to be compensated for honestly reporting bad signals. An interesting analysis of incentives to be aggressive and conservative in self-assessment is Prendergast and Stole (1996). None of these articles analyze the principal's choice between self-assessment, peer review, and multiple reports.

2. The model

■ **Basic setup.** There are two agents and a principal. All are risk neutral, but there is limited liability: all wages must be nonnegative. Each agent can be a good or bad type. An agent's type is not observable to anyone, including himself. The agents' types are uncorrelated random variables. Let λ_i be the prior probability that agent i is a good type.

Agent 1 has developed a blueprint (or ''project''). He is therefore called the inventor and agent 2 is called the peer. There is no moral hazard: the existence and quality of a project cannot be changed by any action taken by the agents.[3] A project can be of good or bad quality. Good projects are successful if implemented, bad projects are failures if implemented. The principal observes the project's true quality if she implements it, but if she doesn't implement it she never observes the true quality. Implementing a successful project is worth $G > 0$ to the principal, while a failure is worth $B < 0$. The quality of the project is perfectly correlated with the inventor's type: good agents produce good projects, bad agents produce bad projects. (Allowing imperfect but positive correlation between types and projects would not change our basic results. As long as the quality of the project can affect the promotion decision, the incentives to misrepresent information remain.)

The agents (but not the principal) have specialized knowledge that allows them to decide if the blueprint is promising or not. Formally, both agents (but not the principal) observe the same signal σ, which is imperfectly correlated with the quality of the project. The signal is either good ($\sigma = g$) or bad ($\sigma = b$). A ''bad signal'' might, for

[3] The case where generating good projects is costly in terms of unobserved effort is left for future research. An inventor who is ahead may actually be tempted to sabotage his own project so it emits a bad signal and is not implemented. In Baliga and Sjöström (1997) we show that the possibility of such sabotage does not change our results.

example, simply represent the absence of any information that supports the project, while a ‘‘good signal’’ might represent the existence of such supporting material. The signal σ is accurate with probability q, where $\frac{1}{2} < q < 1$. Let $p(\sigma)$ denote the probability that the project is of good quality (so that it would be successful if implemented), conditional on the signal $\sigma \in \{g, b\}$. By Bayes’ rule,

$$p(g) = \frac{\lambda_1 q}{\lambda_1 q + (1 - \lambda_1)(1 - q)} > \lambda_1 \qquad p(b) = \frac{\lambda_1(1 - q)}{\lambda_1(1 - q) + (1 - \lambda_1)q} < \lambda_1.$$

Under our assumptions, two truthful reports do not contain more information than one truthful report. However, by collecting two reports the principal can relax the truth-telling constraints.

Incentives to misrepresent signals arise because of career concerns. The agents are competing for a desirable promotion. One and only one agent must be promoted. Thus, we rule out the possibility of promoting both agents (it would be too expensive to create a second managerial position), as well as the possibility of promoting *no* agent (the managerial position must be filled).[4] The value to the principal of promoting a bad type is normalized to zero, the value of promoting a good type is $\Delta > 0$. The non-pecuniary value an agent derives from a promotion is denoted $R > 0$. Thus, R is an ‘‘intrinsic reward’’ from the higher position.[5] If an agent is promoted and gets paid w, then his total payoff is $w + R$. Limited liability requires that $w \geq 0$. If limited liability were not imposed, the incentive to misrepresent information could be negated entirely by having a promoted agent pay the principal R. Since we assume limited liability, and since R is nonpecuniary, charging a fee from the promoted agent in this way is impossible. Limited liability also rules out punishing agents with fines when they disagree with each other, or when their predictions about the project turn out to be wrong.

The principal wants to elicit information about σ in order to inform the project implementation and promotion decisions. It is without loss of generality to restrict attention to direct-revelation mechanisms, where the agents are given the incentive to report the signal honestly. Following the revelation principle for general Bayesian games as stated by Myerson (1991), we assume agents’ reports to the mechanism are private and confidential and not observed by the principal. This is a technical device that simplifies our analysis. We will show in Section 6 that we can just as well use public messages in the optimal renegotiation-proof contract. Therefore, our main results do not depend on the possibility of keeping messages secret from the principal. The time line is as follows.

Time $t = 0$. The agents observe the signal $\sigma \in \{g, b\}$.

Time $t = 1$. Each agent i sends a message $m_i \in \{g, b\}$, interpreted as a report on the blueprint’s quality. The principal does not observe these messages.

Time $t = 2$. The principal receives the instruction ‘‘implement the project’’ or the instruction ‘‘don’t implement the project’’ from the mechanism. The instruction ‘‘implement the project’’ is received with probability $h(m)$, where $m = (m_1 m_2)$ is the pair of messages sent at time $t = 1$.

[4] If not promoting any agent is very costly for the principal, then threatening to promote neither agent out of equilibrium, say as a punishment for disagreeing, would not be renegotiation-proof. Hence, we can just as well assume one agent is always promoted.

[5] Following Calvo and Wellisz (1979), an alternative monetary interpretation of R can be given: Suppose that at low levels of a hierarchical firm, employees can be supervised by managers to make sure that they do not shirk. High-level employees instead receive (efficiency) wages, R, which exceed their reservation wages. Notice that even though R is an (extrinsic) monetary reward, the firm cannot make the promoted agent ‘‘pay for his promotion’’ using some part R, as this causes him to shirk.

Time $t = 3$. The outcome of the project is realized and becomes public knowledge. The set of possible outcomes of a project is $Y = \{G, B, 0\}$ with generic element y. Here G indicates that the project was implemented and was a success, B indicates that it was implemented but was a failure, and 0 indicates that the project was not implemented.

Time $t = 4$. The principal receives the instruction ''promote agent 1'' or the instruction ''promote agent 2'' from the mechanism, together with an instruction about what wages to pay. The instruction ''promote agent 1'' is received with probability $\theta^y(m)$ if messages m were sent at $t = 1$ and $y \in Y$ was the outcome at $t = 3$. An agent's wage can depend on the messages, the outcome of the project, and whether or not the agent is promoted.

The outcome of the project, the output (instructions) of the mechanism, and the principal's actions are all verifiable to an outside party (a court), so the principal cannot unilaterally renege on the contract. However, we will consider the possibility of Pareto-improving renegotiation.

□ **Maintained assumptions.** To eliminate some less interesting cases, the following assumptions will be maintained throughout.

Assumption 1. $p(g)G + (1 - p(g))B > 0$.

Assumption 2. $-(p(b)G + (1 - p(b))B) > p(b)(1 - p(b))\Delta$.

Assumption 3. $R < \Delta \min\{\lambda_2, 1 - \lambda_2\}$.

The principal is concerned about the profitability of the project (B or G) as well as its value in generating information about the inventor. Assumption 1 implies that projects with good signals are expected to be profitable (in addition to generating information). This means the principal definitely wants to implement the project when the signal is good.

Assumption 2 implies that projects with bad signals are sufficiently unprofitable that the principal does not want to implement them, even if they generate some information about the inventor. To see this, suppose the signal is bad. If the principal nevertheless implements the project and promotes the inventor if and only if the project succeeds, her expected payoff is

$$p(b)(G + \Delta) + (1 - p(b))(B + \lambda_2\Delta). \tag{1}$$

To understand this, note that the project succeeds with probability $p(b)$, in which case she earns G and promotes the inventor who is good for sure (since there was a success). With probability $1 - p(b)$ the project fails, in which case she earns B and promotes the peer who is good with probability λ_2. Suppose instead that the project is cancelled when the signal is bad, but the inventor is promoted anyway with probability $p(b)$ and both agents are paid the same expected wages they would have earned had the project been implemented. (Of course, the promotion decision and wages can no longer be made contingent on the project's success. However, the principal can use a random device to generate the probability $p(b)$ and pay the expectation of the wages the agents would have earned.) This maintains the agents' incentives to tell the truth. In this case, whenever the inventor is promoted, the principal's expected payoff is $p(b)\Delta$ (since the inventor is good with probability $p(b)$). If the peer is promoted, the principal's expected payoff is $\lambda_2\Delta$. Since the inventor is promoted with probability $p(b)$, the principal's expected payoff is $p(b)p(b)\Delta + (1 - p(b))\lambda_2\Delta$. Therefore, the principal prefers to cancel the project if

$$p(b)p(b)\Delta + (1 - p(b))\lambda_2\Delta > p(b)(G + \Delta) + (1 - p(b))(B + \lambda_2\Delta).$$

But this inequality is equivalent to Assumption 2, which can be interpreted as requiring that the cost, $-(p(b)G + (1 - p(b))B)$, of acquiring more information about the inventor's quality when there is already a bad signal about him is greater than the benefit, $p(b)(1 - p(b))\Delta$, of a better-informed promotion decision. Therefore, it is better to cancel the project after a bad signal.

Assumptions 1 and 2 together imply that if the principal could observe the signal directly, her first-best policy would be to implement the project if and only if the signal is good. Since her first-best decision would depend on the signal, the design problem we study here, which assumes the principal *cannot* observe the signal, is nontrivial. (In contrast, if her first-best decision did not depend on the signal, then even if she could not observe the signal she could trivially achieve the first best.)

Assumption 3 guarantees that the promotion decision will be sensitive to the quality of the project. Notice that $\Delta\lambda_2$ is the principal's gain from promoting agent 2 instead of agent 1 if agent 1 is known to be bad, and $\Delta(1 - \lambda_2)$ is the gain from promoting agent 1 instead of agent 2 if agent 1 is known to be good. Assumption 3 says that the value to an agent of being promoted, R, is small compared to the cost to the principal of promoting the wrong agent in those cases where agent 1's type is known. If this assumption is violated, then the career concerns are so strong that a principal who cannot observe the signal is better off removing the competition among the agents by making the promotion and wage decisions independent of the quality of the project. The problem would then be fairly trivial. (See Baliga and Sjöström, 1997). For example, consider the following mechanism: either of the two agents is asked for a report, the project is implemented if and only if the report is good, the inventor is always promoted, and all wages are always equal to the minimum (zero). This policy is certainly incentive compatible (and renegotiation-proof if $\Delta\lambda_2 < R$; see below), but there is no interesting interaction between career concerns and misrepresentation of information. The inequality $\lambda_2\Delta > R$ means that such a policy is not optimal. Notice that this inequality also implies that the second employee adds value to the firm: the principal prefers having two candidates for promotion to choose from instead of just one, even though wages can be zero when there is only one candidate. By a symmetric argument, the inequality $(1 - \lambda_2)\Delta > R$ implies that always promoting the peer cannot be optimal.

□ **Renegotiation.** The cheapest way to encourage truth telling is by specifying that if the agents disagree, the project is implemented and the agent whose report agrees with the realized project quality is promoted. This policy means lying is less likely to lead to a promotion. It may lead to the less-talented agent getting promoted, but only out of equilibrium. In equilibrium the agents will send the same report, since they have the same information. Realistically, however, there may be some probability that reports differ because agents honestly disagree about the project (or make mistakes). If this happens with positive probability, it will impose more discipline on the principal's policy: promoting the wrong agent in case of disagreement will no longer have zero cost. An alternative way of imposing such discipline, which we follow here, is to insist on renegotiation-proofness. The contract is renegotiation-proof if at no stage of the game (on or off the equilibrium path) would the principal and both agents prefer to replace the old contract with a new one.

The only times when the principal can have any incentive to propose a new contract are $t = 2$ and $t = 4$. Consider first time $t = 2$. In equilibrium, it will be optimal to implement the project if and only if the signal is good. Thus, if the principal gets the instruction ''don't implement the project,'' she infers that the signal is bad and has no incentive to implement the project. Similarly, if the instruction is ''implement the project,'' she infers that the signal was good, so she has no incentive to cancel it. Therefore, there will never be any temptation to renegotiate at $t = 2$. At this point the assumption that the agents' messages are kept secret from the principal simplifies the analysis. If she could see the actual messages at time $t = 2$, we would have to consider the possibility of renegotiation after the principal learns that $m_1 \neq m_2$ (an out-of-equilibrium event). However, our main results do not depend on the possibility of secret messages. In particular, the optimal self-assessment mechanism we will identify in Section 6 can be operated using only public messages.

At time $t = 4$ the situation is different. The analysis is simplified by the assumption that if the project was implemented, the principal knows agent 1's type for sure. The principal believes agent 2 is good with probability λ_2, independently of what has happened before time $t = 4$, because nothing that happens reveals any information about agent 2's type. If the project is successfully implemented, so that agent 1 is known to be good, then the gain from promoting agent 1 instead of agent 2 is $(1 - \lambda_2)\Delta$. If in such a case the contract specifies that agent 2 should be promoted, the principal can instead propose a new contract as follows. Agent 2 receives an extra monetary transfer of R (on top of what the original contract may have promised him) but is not promoted. Agent 1 is promoted and receives the same wage as in the original contract. By Assumption 3, $(1 - \lambda_2)\Delta > R$, so everybody is made better off (actually agent 2 is indifferent, but he can be given some extra $\epsilon > 0$). So the original contract was not renegotiation-proof. Similarly, if the project fails it would not be renegotiation-proof to promote agent 1. Thus, if (on or off the equilibrium path) the project is implemented, renegotiation-proofness forces the principal to promote the inventor if and only if the project succeeds.

It only remains to discuss the situation when the project is not implemented. It turns out that if the agents' qualifications are fairly similar, then outcomes where the project is cancelled and the ''wrong'' agent is promoted will not be renegotiated. The reason is that, since renegotiation must be voluntary, the principal would have to compensate the agent who gives up his promotion. But this will be more expensive than promoting the wrong agent if the agents have approximately the same qualifications. We need the following definitions.

Definition 1. If $p(b) > \lambda_2$, then the inventor is ahead. If in addition $(p(b) - \lambda_2)\Delta > R$, then he is well ahead.

Definition 2. If $p(b) < \lambda_2$, then the inventor is behind. If in addition $(\lambda_2 - p(b))\Delta > R$, then he is well behind.

Suppose the project is cancelled and the principal thinks the signal is $\sigma = b$. She believes the inventor is good with probability $p(b)$ and the peer is good with probability λ_2. Suppose the inventor is ahead, but according to the contract the peer will be promoted. Will the principal renegotiate the contract and promote the inventor instead? The peer will not accept the new contract unless he is compensated for losing the promotion; this compensation costs the principal R dollars. The expected gain from promoting the inventor instead of the peer is $(p(b) - \lambda_2)\Delta > 0$. If $(p(b) - \lambda_2)\Delta > R$, a Pareto improvement is possible. The principal can pay R dollars to the peer (on top of what he would get under the original contract), pay the inventor the same wage as in the original contract (perhaps zero), and promote the inventor. This makes everybody better off. Thus, promoting the peer would not be renegotiation-proof if the inventor is well ahead. However, suppose the inventor is ahead but not well ahead, and the original contract specifies a zero wage for the inventor. Then, the contract will not be renegotiated despite the fact that the ''wrong'' agent is promoted. The reason is that the peer will insist on a compensation of at least R to give up the promotion, but if the inventor is not well ahead this is not worth it for the principal. Indeed, her gain would only be $(p(b) - \lambda_2)\Delta < R$. Notice that the inventor would be willing to pay up to R dollars to get promoted, but limited liability rules this out.

In exactly the same way we analyze the case where the inventor is behind. If the project is cancelled and the principal thinks the signal is $\sigma = b$, but the contract specifies that the inventor should be promoted, then the gain from promoting the peer instead would be $(\lambda_2 - p(b))\Delta > 0$. The compensation the principal has to pay to the inventor if he is not promoted is R. So promoting the inventor is not renegotiation-proof if the inventor is well behind. But it *is* renegotiation-proof if the inventor is not well behind (as long as the peer's wage according to the original contract is zero, so that the peer cannot ''buy'' the promotion).

Remark. If the project is cancelled and the principal's observations are inconsistent with equilibrium, then she cannot use Bayes' rule to update her beliefs. All we need to assume in this case is that the principal thinks agent 1 is good with at least probability $p(b)$. This is reasonable, since the worst possibility for agent 1 is that the signal was bad. The peer is always thought to be good with probability λ_2 (nothing that happens can reveal any information about the peer). For this article we do not need a more sophisticated analysis of out-of-equilibrium beliefs. In particular, using the terminology of Maskin and Tirole (1992), the concepts of weak and strong renegotiation-proofness coincide.

3. The optimal commitment contract

■ In this section we derive some general properties of the optimal commitment contract, i.e., the optimal contract if the principal can commit not to renegotiate. It will be discussed in more detail in Sections 4 and 5.

□ **Some preliminary results.** By a version of the revelation principle, we may assume the agents play a Nash equilibrium where both agents tell the truth.[6] There are two truth-telling or IC constraints for each agent i, one for each signal $\sigma \in \{g, b\}$. Recall that both agents see the same signal σ, but the two reports can be used to cross-check what one agent reports against the other. Denote by $IC_i(\sigma)$ the constraint that agent i should tell the truth after seeing the signal σ (assuming the other agent tells the truth). Let $w_i^y(m)$ denote agent i's expected wage and $u_i^y(m)$ his expected payoff if the messages are m and the outcome of the project is $y \in \{G, B, 0\}$. The payoff is the sum of the expected wage and the value of being promoted times the probability that a promotion occurs.

[6] Just as in other multi-agent revelation games, in addition to the truth-telling equilibrium there may be other non-truth-telling ones. However, when we consider optimal renegotiation-proof mechanisms, only one agent (the inventor) needs to send a message. This eliminates the problem of multiple equilibria.

For agent 1,

$$u_1^y(m) = w_1^y(m) + \theta^y(m)R,$$

and similarly for agent 2.

Suppose agent 2 always tells the truth. Agent 1's expected payoff when he sees $\sigma = g$ and truthfully announces $m_1 = g$ is

$$h(gg)(p(g)u_1^G(gg) + (1 - p(g))u_1^B(gg)) + (1 - h(gg))u_1^0(gg). \tag{2}$$

If he instead untruthfully announces $m_1 = b$, he expects to get

$$h(bg)(p(g)u_1^G(bg) + (1 - p(g))u_1^B(bg)) + (1 - h(bg))u_1^0(bg). \tag{3}$$

Using (2), (3), and the definition of conditional probabilities, the $IC_1(g)$ constraint can be written as

$$\begin{aligned} &h(gg)(\lambda_1 q u_1^G(gg) + (1 - \lambda_1)(1 - q)u_1^B(gg)) + (1 - h(gg))(\lambda_1 q + (1 - \lambda_1)(1 - q))u_1^0(gg) \\ &\quad \geq h(bg)(\lambda_1 q u_1^G(bg) + (1 - \lambda_1)(1 - q)u_1^B(bg)) \\ &\qquad + (1 - h(bg))(\lambda_1 q + (1 - \lambda_1)(1 - q))u_1^0(bg). \end{aligned}$$

Similarly, we obtain the $IC_1(b)$ constraint for agent 1, and the two IC constraints for agent 2. Finally, the limited liability constraints specify that all wages are nonnegative.

Lemma 1. If the principal can commit not to renegotiate, then the following is optimal: (i) Pay a zero wage to both agents whenever they disagree ($m_1 \neq m_2$). (ii) Pay a zero wage to both agents whenever the project fails. (iii) Implement the project whenever the agents disagree ($h(gb) = h(bg) = 1$). (iv) Set $\theta^B(gg) = 0$ and $\theta^G(gg) = 1$. (v) Set $h(gg) = 1$ and $h(bb) = 0$.

What happens in case of disagreement does not influence the principal's expected payoff directly, because disagreement only happens out of equilibrium. However, it influences the right-hand side of the IC constraints. Since the principal wants this side of the IC constraint to be as small as possible, she pays a zero wage when there is disagreement (part (i) of the lemma). Implementing the project when agents disagree also helps relax the IC constraints by allowing the principal to cross-check agents' messages against the realization of the project quality, so part (iii) is intuitively clear. Part (ii) follows from the fact that everyone is risk neutral and cares only about expected payoffs.

Part (iv) of the lemma says that if the project receives two good reports and is implemented, then the inventor is promoted if and only if the project succeeds. Indeed, if the inventor is promoted when the project fails, then the principal can promote the peer instead and compensate the inventor by a payment of R dollars. The peer is good with probability λ_2, and the inventor is known for sure to be bad when the project fails. Assumption 3 implies that the gain of $\lambda_2\Delta$ from promoting the peer instead of the inventor is greater than the cost R. Moreover, no IC constraint will be violated. The argument for why the inventor should be promoted when the project succeeds is similar, using the fact that when the project succeeds the inventor is known for sure to be good, and the assumption that $(1 - \lambda_2)\Delta > R$. (The promotion policy when $m_1 \neq m_2$ will be discussed later.) Finally, Assumptions 1 and 2 imply that projects with good signals make expected profits, and projects with bad signals are too costly to be worth implementing. Thus, in equilibrium the principal wants to implement the project if and only if the signal is good. This explains part (v) of the lemma.

□ **The principal's simplified problem.** We can now simplify the principal's problem by assuming the equalities derived in Lemma 1 hold. Consider the principal's payoff. With probability λ_{1q} both the project and the signal is good. In this case, assuming the agents tell the truth, the project is successfully implemented (by Lemma 1 part (v)), agent 1 is promoted (by Lemma 1 part (iv)), and the principal's payoff is

$$G + \Delta - w_1^G(gg) - w_2^G(gg).$$

The principal's payoff for other cases is similarly computed using Lemma 1. Overall, the principal's expected payoff is

$$\begin{aligned}&\lambda_1 q(G + \Delta - w_1^G(gg) - w_2^G(gg)) + (1 - \lambda_1)(1 - q)(B + \lambda_2\Delta)\\ &\quad + \lambda_1(1 - q)(\theta^0(bb) + (1 - \theta^0(bb))\lambda_2)\Delta + (1 - \lambda_1)q(1 - \theta^0(bb))\lambda_2\Delta \\ &\quad - (\lambda_1(1 - q) + (1 - \lambda_1)q)(w_1^0(bb) + w_2^0(bb)).\end{aligned} \tag{4}$$

She maximizes this expression subject to the IC constraints, which using Lemma 1 we can simplify as follows.
$IC_1(g)$:

$$\lambda_1 q(w_1^G(gg) + R) \geq \lambda_1 q\theta^G(bg)R + (1 - \lambda_1)(1 - q)\theta^B(bg)R. \tag{5}$$

$IC_1(b)$:

$$\begin{aligned}&(\lambda_1(1 - q) + (1 - \lambda_1)q)(w_1^0(bb) + \theta^0(bb)R)\\ &\quad \geq \lambda_1(1 - q)\theta^G(gb)R + (1 - \lambda_1)q\theta^B(gb)R.\end{aligned} \tag{6}$$

$IC_2(g)$:

$$\begin{aligned}&\lambda_1 q w_2^G(gg) + (1 - \lambda_1)(1 - q)R\\ &\quad \geq \lambda_1 q(1 - \theta^G(gb))R + (1 - \lambda_1)(1 - q)(1 - \theta^B(gb))R.\end{aligned} \tag{7}$$

$IC_2(b)$:

$$\begin{aligned}&(\lambda_1(1 - q) + (1 - \lambda_1)q)(w_2^0(bb) + (1 - \theta^0(bb))R)\\ &\quad \geq \lambda_1(1 - q)(1 - \theta^G(bg))R + (1 - \lambda_1)q(1 - \theta^B(bg))R.\end{aligned} \tag{8}$$

The limited liability constraints are

$$w_1^G(gg),\ w_2^G(gg),\ w_1^0(bb),\ w_2^0(bb) \geq 0. \tag{9}$$

To explain the left-hand side of $IC_1(g)$, for example, use Lemma 1 to set $h(gg) = 1$, $u_1^G(gg) = w_1^G(gg) + \theta^G(gg)R = w_1^G(gg) + R$, and $u_1^B(gg) = w_1^B(gg) + \theta^B(gg)R = 0$ in the left-hand side of the $IC_1(g)$ constraint stated before Lemma 1. For $m_1 \neq m_2$, we get $u_1^G(bg) = \theta^G(bg)R$, $u_2^G(bg) = (1 - \theta^G(bg))R$, $u_1^0(bg) = \theta^0(bg)R$, etc.

Intuitively, the solution to this program depends on who is the better candidate for a promotion when the project is cancelled, since this determines who has an incentive to push through the project and who has an incentive to suppress it. In Section 4 we consider the case when the inventor is behind. Section 5 deals with the case when the inventor is ahead.

4. The inventor is behind: exaggeration and denigration

■ The inventor is behind if $p(b) < \lambda_2$. In the terminology of Prendergast and Stole (1996), the inventor is an ''impetuous youngster'' who must prove himself against a colleague with a better reputation, while the peer is a conservative ''old-timer'' who likes the status quo. Since the peer is the best candidate for promotion conditional on a bad signal, the inventor is tempted to exaggerate the quality of the project. The peer is tempted to denigrate the project. Formally, the binding IC constraints are $IC_1(b)$ and $IC_2(g)$ (see the proof of Proposition 1). On the right-hand side of these constraints, we find the variables that relate to the message profile $m = gb$. Thus, the principal has to worry about what to do when the inventor says the signal is good and the peer says it is bad.

□ **Optimal commitment contract.** If the agents disagree, then the optimal commitment contract implements the project and gives the promotion to the agent whose report agrees with the actual project quality. Thus, following the critical message profile $m = gb$, the inventor is promoted if and only if the project succeeds. This minimizes the incentives to lie. With such a policy, denigration does not increase the peer's probability of promotion. Indeed, even if only the inventor supports the project, the principal implements it and the promotion policy is the same as if both agents had supported the project. The inventor, on the other hand, can expect $p(b)R > 0$ by exaggerating, since exaggeration guarantees that the project will be implemented, and conditional on a bad signal it will be successful (and he promoted) with probability $p(b)$. To prevent exaggeration, in equilibrium the contract must give the inventor at least $p(b)R$ when the signal is bad. If the inventor is not well behind, then the cheapest way to do this is by promising the inventor a probability $p(b)$ of getting a promotion even if the project is cancelled. But, if the inventor is well behind, then the cost of promoting the inventor when the project is cancelled is significant. The cheapest way to prevent exaggeration is instead to give a monetary compensation to the inventor (but no chance of a promotion) when the project is cancelled.

Proposition 1. If the principal can commit not to renegotiate and the inventor is behind, then the following is optimal. When the messages contradict each other, implement the project and promote the inventor if and only if his project succeeds ($\theta^G(m) = 1$ and $\theta^B(m) = 0$ if $m_1 \neq m_2$). If the inventor is not well behind, promote him with positive probability when both agents report the signal is bad ($\theta^0(bb) = p(b) > 0$), and all wages are zero. If the inventor is well behind, never promote him when both reports are bad ($\theta^0(bb) = 0$), and all wages should be zero except $w_1^0(bb) = p(b)R$.

Together with Lemma 1, Proposition 1 describes the optimal commitment contract when $p(b) < \lambda_2$. The salient features are, first, that the project is implemented following all messages except $m = bb$, and when the project is implemented the inventor is promoted if and only if it succeeds (whatever the message profile). Second, to prevent exaggeration the inventor is compensated when the project is cancelled. The compensation takes the form of a possibility of promotion ($\theta^0(bb) > 0$) when the inventor is not well behind, but it takes a purely monetary form ($w_1^0(bb) > 0$) when he is well behind.

□ **Renegotiation.** The only potential problem with the contract derived in Proposition 1 is that the inventor is sometimes promoted when $m = bb$ and the project is cancelled, even though he is behind and the principal infers from the messages that the signal was bad. However, this only happens when he is not well behind, and since the peer's wage is zero, such a policy is renegotiation-proof, as was shown in Section 2. Thus, we have

Proposition 2. If the inventor is behind, then the optimal commitment contract is renegotiation-proof.

5. The inventor is ahead: false modesty and false praise

■ The inventor is ahead if $p(b) > \lambda_2$. He is preferred for promotion even after a bad signal but would not remain so after a failed project (since a failure reveals that he is a bad type). Thus, the inventor likes the status quo and has an incentive to underestimate the project's quality (playing it safe rather than suffering a costly failure, which we call false modesty). The peer, on the other hand, thinks he can get promoted only if the inventor has a failed project, and a necessary condition for this to happen is that a project is implemented. Thus, the peer is tempted to overestimate the project's quality, which we call false praise. The binding IC constraints are $IC_1(g)$ and $IC_2(b)$. Thus, the principal has to worry about what to do when the inventor's report is bad but the peer report is good ($m = bg$).

□ **Optimal commitment contract.** If the agents disagree, then the optimal commitment contract minimizes the incentive to lie by implementing the project and promoting the agent whose message best corresponds to the project quality. Therefore, following the critical message $m = bg$, the inventor is promoted with a higher probability if the project fails than if it succeeds.

Proposition 3. Suppose the principal can commit not to renegotiate and the inventor is ahead. Then the following policy is optimal. If $m = bg$ and the project fails, promote the inventor with positive probability ($\theta^B(bg) > 0$). If $m = bg$ and the project succeeds, promote the peer with positive probability ($\theta^G(bg) < 1$). If the inventor is well ahead, then promote him for sure when $m = bb$ (i.e., $\theta^0(bb) = 1$), and all wages are zero except $w_2^0(bb)$ and possibly $w_1^G(gg)$. If the inventor is not well ahead, then promote the peer with positive probability when $m = bb$ (i.e., $\theta^0(bb) < 1$), and all wages are zero except possibly $w_1^G(gg)$.

□ **Renegotiation.** There are two potential problems with the optimal commitment contract described in Proposition 3. First, if both agents report bad signals, the project is cancelled but the peer is sometimes promoted ($\theta^0(bb) < 1$) even though the inventor is ahead. Since this happens only when the inventor is not well ahead, this does not cause renegotiation (see Section 2). However, renegotiation-proofness fails out of equilibrium because, in order to provide incentives to tell the truth, the peer can get promoted if he is the only one who supported a successful project ($\theta^G(bg) < 1$), and the inventor can get promoted if he is the one who did not support an unsuccessful project ($\theta^B(bg) > 0$). As shown in Section 2, this is not renegotiation-proof.

We now consider the optimal renegotiation-proof contract for the case where the inventor is ahead. When the project is implemented, renegotiation-proofness requires that the inventor is promoted if and only if his project succeeds:

$$\theta^G(m) = 1 \quad \text{and} \quad \theta^B(m) = 0 \quad \text{for all } m. \tag{10}$$

If the project is cancelled, then renegotiation-proofness implies that the inventor must be promoted whenever he is well ahead:

$$\theta^0(m) = 1 \quad \text{for all } m. \tag{11}$$

(Even if the principal's observations are inconsistent with equilibrium, (11) must hold when the inventor is well ahead, since she cannot think the inventor is good with probability less than $p(b)$ when the project is cancelled). However, if the inventor is not well ahead, then (11) can be violated.

With these constraints, it is still easy to prove parts (i), (ii), and (iv) of Lemma 1. But it is no longer clear that $h(gb) = h(bg) = h(gg) = 1$ is optimal: if the principal cannot choose her promotion policy freely after the project is implemented, implementing the project becomes less desirable. Therefore, to find the optimal renegotiation-proof contract when the inventor is ahead, we modify the simplified program of Section 3 by not assuming

$$h(gg) = h(gb) = h(bg) = 1. \tag{12}$$

Instead we assume (10) holds and, if the inventor is well ahead, also (11).

Lemma 2. Suppose the inventor is ahead but not well ahead. The following is the solution if the simplified program of Section 3 is modified by not imposing (12) but instead imposing (10). Set $h(gg) = h(gb) = 1$ and $h(bb) = h(bg) = 0$. All wages except possibly $w_1^G(gg)$ are zero. If

$$\frac{qp(b)}{(1-q)p(g)}R + (\lambda_2 - p(b))\Delta < 0, \tag{13}$$

then $\theta^0(bb) = \theta^0(bg) = 1$ and

$$w_1^G(gg) = \frac{1-p(g)}{p(g)}R.$$

If

$$\frac{qp(b)}{(1-q)p(g)}R + (\lambda_2 - p(b))\Delta \geq 0, \tag{14}$$

then

$$\theta^0(bb) = \theta^0(bg) = p(g) \tag{15}$$

and $w_1^G(gg) = 0$.

Lemma 3. Suppose the inventor is well ahead. When the simplified program of Section 3 is modified by not imposing (12) but instead imposing (10) and (11), the solution involves $h(gg) = h(gb) = 1$ and $h(bb) = h(bg) = 0$. All wages are zero, except

$$w_1^G(gg) = \frac{1 - p(g)}{p(g)}R.$$

The contracts characterized in Lemmas 2 and 3 are renegotiation-proof. This is certainly true in the case of Lemma 3 because the right agent is always promoted by construction. That is, agent 1 is promoted except when his project has failed. In the case of Lemma 2 the only problematic aspect is equation (15). Agent 2 is promoted with probability $1 - p(g) > 0$ when a bad signal is received and the project cancelled, even though agent 1 is ahead. However, since agent 1 is not well ahead, this policy is renegotiation-proof (see Section 2). Since any renegotiation-proof contract must satisfy the constraints of the programs analyzed in Lemmas 2 and 3, we conclude the following:

Proposition 4. The contracts described in Lemmas 2 and 3 are optimal within the set of renegotiation-proof contracts (for the cases where the inventor is ahead but not well ahead, or well ahead, respectively).

When the inventor is ahead, the optimal renegotiation-proof contract involves not implementing the project when only the peer supports it, i.e., $h(bg) = 0$. If the project is implemented, the principal discovers the inventor's true type, but this is not advantageous when renegotiation-proofness ''ties her hands.'' (Dewatripont and Maskin (1995) discuss the fact that receiving more information can hurt the principal if she cannot commit.) By cancelling the project when the inventor says it is bad and the peer says it is good, the principal gains flexibility in using promotions as a reward. However, this in effect gives the inventor veto power over the implementation of the project. To prevent false modesty, the inventor is either given a monetary reward when he reports a good signal and the project succeeds ($w_1^G(gg) > 0$), or he is promoted with a probability strictly less than 1 when the project is cancelled ($\theta^0(bb) < 1$). Notice that if the inventor is well ahead, then the monetary reward is used to prevent false modesty, not the promotion policy.

6. Self-assessment: delegation of authority to the inventor

■ So far we have assumed all contracts involve multiple reports. It turns out, however, that the optimal contract can always be replicated by a simple self-assessment mechanism.[7] Let the inventor report the signal, but ask for no message from the peer. If the inventor supports his own project, then implement it, and promote the inventor if and only if the project is successful. If the project is successfully implemented and the inventor is ahead, then pay him $R(1 - p(g))/p(g)$, except when he is not well ahead and (14) holds, in which case pay him zero. If the inventor does not support his own project, then don't implement it, and use the following promotion policy:

[7] If it is impossible to prevent the peer from voicing his opinion, the principal should behave as suggested by the optimal renegotiation-proof mechanism: if the peer disagrees with the inventor, the principal should simply disregard the peer. In this case, the principal's beliefs cannot be determined by Bayes' rule, and we can assume that she chooses to believe the inventor's report. The policy is then renegotiation-proof and the peer may as well be quiet, since his opinion will be disregarded anyway.

Case 1. Suppose agent 1 is behind. If agent 1 is well behind, then pay him $p(b)R$ and promote agent 2; but if agent 1 is not well behind, then promote agent 1 with probability $p(b)$ and pay him zero.

Case 2. Suppose agent 1 is ahead. Then if agent 1 is not well ahead and (14) holds, promote him with probability $p(g)$. Otherwise, promote him for sure.

Except as mentioned, all wages are zero. It is easy to check that the inventor will tell the truth, and the project implementation decision, wages, and promotion policies mimic the optimal renegotiation-proof contracts derived above. Notice that the outcome of the self-assessment mechanism actually replicates the optimal commitment contract when the inventor is behind, but not when he is ahead. (This suggests that if the principal can commit, then multiple reports are useful if the inventor is a ‘‘senior worker’’ who is ahead in his career, but self-assessment is fine for ‘‘junior workers’’ who are behind.)

The binding IC constraint when the inventor is behind is that he should not say the signal is good when it is bad, i.e., exaggeration must be prevented. The binding IC constraint when the inventor is ahead is that he should not say the signal is bad when it is good, i.e., false modesty must be prevented (compare the low risk taking of middle management and high risk taking of entry-level employees identified in Gamache and Kuhn, 1989). Truth telling is guaranteed by the design of wage and promotion policies. If after the project is cancelled the agents have approximately the same qualifications (so that the principal does not care too much about who is promoted), then the promotion policy is made less high-powered than it would be in the first best. Thus, an inventor who is behind but not well behind is promoted with probability $p(b) > 0$ even when the project is cancelled, in order to prevent exaggeration. If the inventor is ahead but not well ahead and (14) holds, then to prevent false modesty he is promoted only with probability $p(g) < 1$ when the project is cancelled. If one agent is a much better candidate for promotion than the other one after the project is cancelled, then distorting the promotion policy is very costly. Instead, the principal prefers to distort the wage policy. Namely, if the inventor is well ahead (or not well ahead but (14) holds), then to prevent false modesty he is given an extra-high wage if his project succeeds. If the inventor is well behind, then to prevent exaggeration he is given an extra-high wage when his project is cancelled.

Williamson (1983) argues that large organizations are bureaucratic and stifle drastic innovations, and proposes that allowing inventors more autonomy and discretion may lead to more successful research and development.[8] Our optimal self-assessment mechanism can be interpreted as the delegation of authority over project-implementation decisions to the inventor. This simple mechanism has two additional advantages. First, while multi-agent revelation games often are plagued by multiple Nash equilibria, a self-assessment mechanism avoids this problem. Second, there is no need for ‘‘secret messages’’: the principal may just as well observe the inventor’s message directly.

In deriving these results we have assumed that both agents observe the same signal. Scotchmer (1990) points out that eliciting multiple opinions may be useful when different agents get different signals. (However, she stresses that asking for reports from several agents may lead to the existence of multiple equilibria.) Our results have implications also for the case when the agents observe different (imperfectly correlated)

[8] Williamson (1983), gives the example of General Electric setting up a partially owned ‘‘Technical Ventures Operation’’ (TVO) to develop projects and to share in their rewards. Williamson (1987) reports that similar arrangements are becoming more popular, and he offers other examples including General Motors’ partial ownership position in a firm producing innovations.

signals. In this case the advantages of self-assessment would be traded off against an added advantage of multiple reports: multiple reports carry information about multiple signals. For example, suppose the inventor is ahead. If the inventor says his project is bad but the peer says it is good, either one agent is lying or they have observed different signals. In the latter case, if the peer's signal is very informative, then the principal may well want to implement the project. However, implementing the project when the agents disagree in this way encourages ''false praise.'' Indeed, the peer would be promoted when the project fails, and this is likely to happen when the peer's signal is bad. To prevent false praise, the principal would have to offer a large surplus to the peer when the project is cancelled, but this would be costly. In designing the optimal mechanism, the principal would trade off the value of getting a report from the peer against the cost of satisfying the peer's truth-telling constraint. If the signals are sufficiently highly correlated, the principal prefers a self-assessment mechanism.

Our model also provides some intuition for why a policy of self-assessment (without any peer report) may dominate a policy of peer review (without any report from the inventor). For example, suppose the inventor is well behind. With a policy of peer review, the peer can guarantee himself a promotion by denigrating the project. On the other hand, if the peer truthfully reports that the signal is good, then he is only promoted when the project fails, which happens with probability $1 - p(g)$. Therefore, to prevent denigration, the peer must get paid $p(g)R$ dollars if he reports that the signal is good. The *ex ante* cost of this is $\Pr(\sigma = g) \times p(g)R = \lambda_1 qR$. Intuitively, since the principal will never learn the project's true quality if the project is cancelled, the incentive to denigrate is large. With self-assessment, exaggeration needs to be prevented, but to do this it suffices to pay the inventor $p(b)R$ dollars if he reports that the signal is bad, at an *ex ante* cost of $\Pr(\sigma = b) \times p(b)R = \lambda_1(1 - q)R$. This is so because if the principal implements the project when the signal is bad, the project will only succeed with probability $p(b)$. In other words, exaggeration is unlikely to lead to a promotion. As long as $q > ½$, the peer's incentive to denigrate is stronger than the inventor's incentive to exaggerate, and it is cheaper to guarantee truth telling with self-assessment than with peer review.

Suppose instead that the inventor is well ahead. With a policy of peer review, by false praise the peer can persuade the principal to implement a project with a bad signal. This raises the peer's chance of a promotion from 0 to $1 - p(b)$, so the incentive to engage in false praise is great. To prevent false praise, the principal must pay the peer $(1 - p(b))R$ dollars when he says the signal is bad, at an *ex ante* cost of $\Pr(\sigma = b) \times (1 - p(b))R = (1 - \lambda_1)qR$. With a policy of self-assessment the inventor can guarantee himself a promotion by false modesty. If the inventor truthfully reports the signal is good, the project is implemented and he is promoted with probability $p(g)$. Therefore, to prevent false modesty the principal only needs to pay the inventor $(1 - p(g))R$ dollars when he says that the signal is good, at an *ex ante* cost of $\Pr(\sigma = g) \times (1 - p(g))R = (1 - \lambda_1)(1 - q)R$. Indeed, false modesty only raises the possibility of a promotion from $p(g)$ to 1. Intuitively, the inventor's temptation to be falsely modest is small, since a project with a good signal is likely to succeed anyway. As long as $q > ½$, the peer's incentive to engage in false praise is greater than the inventor's incentive to engage in false modesty, so self-assessment is again cheaper than peer review.

Appendix

■ Proofs to Lemmas 1–3 and Propositions 1 and 3 follow.

Proof of Lemma 1. (i) This follows from the fact that the disagreement payoffs $u_1^B(bg)$, $u_1^G(bg)$, etc., only enter on the right-hand side of the IC constraints.

(ii) If $m_1 \neq m_2$, then this follows from (i). So suppose $m_1 = m_2 = g$. The wages $w_1^G(gg)$ and $w_1^B(gg)$ enter in the principal's payoff and the $IC_1(g)$ constraints through the term

$$p(g)w_1^G(gg) + (1 - p(g))w_1^B(gg),$$

which is the expected wage to agent 1 when $\sigma = g$ and the project is implemented. Since both the principal and the agent only care about this expectation, it is without loss of generality to set $w_1^B(gg)$ as low as possible. A similar argument holds for $w_2^B(gg)$ and for $m_1 = m_2 = b$ (although the latter case is irrelevant, since the principal does not implement the project when $m = bb$; see below).

(iii) From (i) we can suppose all the disagreement wages are zero. Notice that $h(bg)$ only appears on the right-hand side of $IC_1(g)$ and $IC_2(b)$. Suppose a proposed contract has $h(bg) < 1$. Suppose the principal changes the contract in the following way: following the message bg, the project is implemented with probability one, and if the project succeeds, agent 1 is promoted with probability

$$h(bg)\theta^G(bg) + (1 - h(bg))\theta^0(bg)$$

(where $h(bg)$, $\theta^G(bg)$, $\theta^0(bg)$ are as specified in the original contract). If the project fails, he promotes agent 1 with probability

$$h(bg)\theta^B(bg) + (1 - h(bg))\theta^0(bg).$$

This leaves the right-hand sides of $IC_1(g)$ and $IC_2(b)$ unchanged. The principal's welfare is unchanged, since disagreements never happen in equilibrium. Hence, we can assume without loss of generality that $h(bg) = 1$ and, by a similar argument, $h(gb) = 1$.

(iv) Suppose $\theta^B(gg) > 0$. Consider changing the contract by reducing $\theta^B(gg)$ by ϵ. Compensate agent 1 by raising $w_1^B(gg)$ by ϵR. In the new contract the IC constraints are obviously still satisfied. If the signal is good and the project is implemented but fails, the increase in expected wages is only ϵR, while the principal gains $\epsilon\lambda_2\Delta$ by promoting agent 2 more often. By Assumption 3, this improves the principal's payoff. The argument for $\theta^G(gg) = 1$ is similar.

(v) Consider $h(gg)$. By implementing the project when the signal is good, the principal gets more information because she can observe the outcome of the project. Since she can always disregard this information if she wants (as in the proof of part (iii)), she can design a policy with $h(gg) = 1$ that implies no greater wage payments than a policy with $h(gg) < 1$. As message profile $m = gg$ is received in equilibrium whenever $\sigma = g$, there is also a direct effect on the principal's revenue from increasing $h(gg)$. But this is positive by Assumption 1. Therefore, $h(gg) = 1$ is optimal.

Consider $h(bb)$. Suppose a contract has $h(bb) > 0$. Suppose the message profile bb is received and the project is about to be implemented (which it is with probability $h(bb)$). By the same argument as in the proof of part (iv), we may suppose $\theta^B(bb) = 0$ and $\theta^G(bb) = 1$. That is, the inventor will be promoted if and only if the project succeeds, which it does with probability $p(b)$. It was shown in Section 2 that the principal's expected payoff is given by (1). Suppose instead the project is cancelled, but the inventor nevertheless is promoted with probability $p(b)$. All wages (conditional on the promotion decision) are the same as before. Then, the agents' expected payoffs do not change, so the IC constraints are not violated. The principal's expected payoff now becomes $(p(b)p(b) + (1 - p(b))\lambda_2)\Delta$ (this is so since the inventor is good with probability $p(b)$, and the peer with probability λ_2). Cancelling the project makes the principal better off, since by Assumption 2 we have

$$(p(b)p(b) + (1 - p(b))\lambda_2)\,\Delta > p(b)(G + \Delta) + (1 - p(b))(B + \lambda_2\Delta).$$

This proves $h(bb) = 0$ is optimal. *Q.E.D.*

Proof of Proposition 1. Suppose the inventor is behind. We split the proof into two cases (1) the inventor is not well behind, and (2) the inventor is well behind.

Case 1 (the inventor is not well behind):

$$p(b) < \lambda_2 < p(b) + \frac{R}{\Delta}. \tag{A1}$$

We will show that it is optimal to set all wages equal to zero, to set $\theta^G(gb) = \theta^G(bg) = 1$, to set $\theta^B(gb) = \theta^B(bg) = 0$, and to set $\theta^0(bb) = p(b)$. Assume $IC_1(g)$ and $IC_2(b)$ are not binding, so these constraints can be disregarded (we will check this later). Now we need a series of claims.

Claim 1. $w_1^G(gg) = w_2^0(bb) = 0$ is optimal.

Proof. Obvious.

Claim 2. $IC_1(b)$ binds at the optimum.

Proof. Suppose not. Then $w_1^0(bb) = 0$, or else the principal can lower $w_1^0(bb)$ without violating any constraints. Thus, $\theta^0(bb) > 0$ if $IC_1(b)$ does not bind. But lowering $\theta^0(bb)$ raises the principal's profit by (A1) without violating any incentive constraints.

Claim 3. $IC_2(g)$ binds at the optimum.

Proof. Suppose not. Then we must have $w_2^G(gg) = 0$ (or else it should be reduced). Then $IC_2(g)$ slack means we cannot have $\theta^B(gb) = \theta^G(gb) = 0$. So one of these variables can be reduced without altering the principal's payoff and without violating $IC_2(g)$. This relaxes $IC_1(b)$, but then profit can be increased as in the proof of claim 2.

Claim 4. $w_1^0(bb) = 0$.

Proof. Suppose $w_1^0(bb) > 0$. $IC_1(b)$ binding means $\theta^0(bb) < 1$. Lower $w_1^0(bb)$ by ϵR and increase $\theta^0(bb)$ by ϵ. This raises the principal's profit by

$$\epsilon(\lambda_1(1 - q) + (1 - \lambda_1)q)(R - \Delta(\lambda_2 - p(b))) > 0$$

by (A1) without violating any constraints.

Claim 5. $\theta^B(gb) = 0$, $\theta^G(gb) = 1$, $\theta^0(bb) = p(b)$ and $w_2^G(gg) = 0$.

Proof. Since $w_1^0(bb) = 0$, we can solve for $\theta^0(bb)$ and $w_2^G(gg)$ from the two binding constraints $IC_1(b)$ and $IC_2(g)$. Substituting this into the principal's objective function, and maximizing subject to the constraint $w_2^G(gg) \geq 0$, results in $\theta^B(gb) = 0$ and $\theta^G(gb) = 1$. This implies $\theta^0(bb) = p(b)$ and $w_2^G(gg) = 0$. This proves the claim.

Finally, we can make sure $IC_1(g)$ and $IC_2(b)$ are satisfied by setting $\theta^B(bg) = 0$ and $\theta^G(bg) = 1$.

Case 2 (the inventor is well behind):

$$\lambda_2 > p(b) + \frac{R}{\Delta}. \tag{A2}$$

We claim it is optimal to set all wages equal to zero except $w_1^0(bb) = p(b)R$, and to set

$$\theta^0(bb) = \theta^B(bg) = \theta^B(gb) = 0 \quad \text{and} \quad \theta^G(gb) = \theta^G(bg) = 1.$$

Suppose $\theta^0(bb) > 0$. Then lower $\theta^0(bb)$ by ϵ and raise $w_1^0(bb)$ by $R\epsilon$. This changes the principal's payoff by

$$\epsilon(\lambda_1(1 - q) + (1 - \lambda_1)q)((\lambda_2 - p(b))\Delta - R) > 0$$

using (A2), without violating any incentive constraints. Therefore, $\theta^0(bb) = 0$.

The gb-variables only appear in the $IC_2(g)$ and $IC_1(b)$ constraints. Notice that $IC_1(b)$ must hold with equality: otherwise just lower $w_1^0(bb)$. Therefore, as $q > 1 - q$, it is optimal to set $\theta^B(gb) = 0$, as it reduces total expected wage payments. Then $IC_2(g)$ binds at the optimum: otherwise it must be the case that $w_2^G(gg) > 0$, but then $w_2^G(gg)$ can be reduced. Therefore, as $q > 1 - q$, it is optimal to set $\theta^G(gb) = 1$, as it minimizes expected wage payments (the principal cares about the sum of the right-hand sides of $IC_1(b)$ and $IC_2(g)$). But then $IC_2(g)$ is satisfied with $w_2^G(gg) = 0$, so this is optimal. From $IC_1(b)$ we obtain $w_1^0(bb) = p(b)R$.

The bg-variables only appear on the right-hand side of $IC_2(b)$ and $IC_1(g)$. These constraints are satisfied at minimum cost if $w_2^0(bb) = w_1^G(gg) = 0$, $\theta^B(bg) = 0$, and $\theta^G(bg) = 1$. *Q.E.D.*

Proof of Proposition 3. Suppose the inventor is ahead. We split the proof into two cases: (1) the inventor is well ahead, and (2) the inventor is not well ahead.

Case 1 (the inventor is well ahead):

$$\lambda_2 < p(b) - \frac{R}{\Delta}. \tag{A3}$$

Then, we claim it is optimal to set $\theta^0(bb) = \theta^G(gb) = \theta^B(bg) = 1$, and $\theta^B(gb) = 0$, and

$$\theta^G(bg) = \max\left\{0,\ 1 - \frac{(1-\lambda_1)(1-q)}{\lambda_1 q}\right\}.$$

Each agent is paid a zero wage, except that $w_2^0(bb) = p(b)(1 - \theta^G(bg))R$, and if $(1-\lambda_1)(1-q) \geq \lambda_{1q}$, then

$$w_1^G(gg) = \left(\frac{(1-\lambda_1)(1-q)}{\lambda_1 q} - 1\right)R.$$

To show this, first consider increasing $\theta^0(bb)$ by $\epsilon > 0$ and $w_2^0(bb)$ by $R\epsilon$. By (A3), this changes the principal's payoff by

$$\begin{aligned}&(\lambda_1(1-q)(1-\lambda_2) - (1-\lambda_1)q\lambda_2)\epsilon\Delta - (\lambda_1(1-q) + (1-\lambda_1)q)R\epsilon\\ &\quad = (\lambda_1(1-q) + (1-\lambda_1)q)\epsilon[(p(b) - \lambda_2)\Delta - R] > 0\end{aligned}$$

without violating any constraints. Thus, $\theta^0(bb) = 1$ is optimal.

The bg-variables only appear on the right-hand side of $IC_1(g)$ and $IC_2(b)$ constraints. $IC_2(b)$ will hold with equality, otherwise $w_2^0(bb)$ can be reduced ($w_2^0(bb) > 0$ as $\theta^0(bb) = 1$). We claim also that $IC_1(g)$ holds with equality. If not, then $w_1^G(gg) = 0$, and the principal will set the bg-variables to minimize the right-hand side of $IC_2(b)$, as it is the only binding constraint involving the bg-variables. This implies $\theta^G(bg) = \theta^B(bg) = 1$. But then $IC_1(g)$ is violated, a contradiction. Thus, $IC_1(g)$ holds with equality. By inspection of the principal's expected wage payments, we see that she cares about the sum of the left-hand side of the $IC_1(g)$ and $IC_2(b)$ constraints. As both these constraints are satisfied with equality, she should set the bg-variables to minimize the sum of the right-hand side of the $IC_1(g)$ and $IC_2(b)$ constraints, with the restriction that the right-hand side of $IC_1(g)$ must exceed $\lambda_1 qR$, for otherwise equality in $IC_1(g)$ is incompatible with limited liability, (9).

We claim $\theta^B(bg) = 1$ is optimal. For if $\theta^B(bg) < 1$, then raising $\theta^B(bg)$ lowers the sum of expected wage payments as the right-hand side of $IC_1(g)$ increases more slowly than the right-hand side of $IC_2(b)$ falls, by $q > 1 - q$. Also, by a similar argument, the principal should set $\theta^G(bg)$ as low as possible. However, the right-hand side of $IC_1(g)$ must exceed $\lambda_1 qR$. This yields two cases: (a) if $(1-\lambda_1)(1-q) \geq \lambda_1 q$, then set $\theta^G(bg) = 0$, and from $IC_1(g)$, $w_1^G(gg) + R = [(1-\lambda_1)(1-q)/\lambda_1 q]R \geq R$; (b) if $(1-\lambda_1)(1-q) < \lambda_1 q$, then set $\theta^G(bg) = 1 - [(1-\lambda_1)(1-q)/\lambda_1 q]$, and from $IC_1(g)$, $w_1^G(gg) = 0$. Finally,

$$w_2^0(bb) = \{[\lambda_1(1-q)]/[\lambda_1(1-q) + (1-\lambda_1)q]\}(1 - \theta^G(bg))R$$

from $IC_2(b)$.

Finally, by setting $\theta^G(gb) = 1$, $\theta^B(gb) = 0$, $IC_1(b)$ and $IC_2(g)$ are satisfied with $w_1^0(bb) = w_2^G(gg) = 0$, and this is clearly optimal. This completes the proof for Case 1.

Case 2 (the inventor is not well ahead):

$$p(b) - \frac{R}{\Delta} < \lambda_2 < p(b). \tag{A4}$$

Then, we claim the following is optimal: $\theta^G(gb) = 1$, $\theta^B(gb) = 0$, $\theta^0(bb) = 1 - p(b)(1 - \theta^G(bg))$, $w_2^0(bb) = w_1^0(bb) = w_2^G(gg) = 0$,

$$\theta^G(bg) = \begin{cases} 0 & \text{if } \dfrac{(1-\lambda_1)(1-q)}{\lambda_1 q} \geq 1 \\ 1 - \dfrac{(1-\lambda_1)(1-q)}{\lambda_1 q} & \text{otherwise} \end{cases} \tag{A5}$$

$$\begin{aligned}\theta^B(bg) &= \begin{cases} \dfrac{\lambda_1 q}{(1-\lambda_1)(1-q)} & \text{if } \dfrac{(1-\lambda_1)(1-q)}{\lambda_1 q} \geq 1 \quad \text{and} \quad p(b) - \lambda_2 - \dfrac{1-q}{q}\dfrac{R}{\Delta} < 0 \\ 1 & \text{otherwise} \end{cases}\\ w_1^G(gg) &= \begin{cases} \left(\dfrac{(1-\lambda_1)(1-q)}{\lambda_1 q} - 1\right)R & \text{if } \dfrac{(1-\lambda_1)(1-q)}{\lambda_1 q} \geq 1 \quad \text{and} \quad p(b) - \lambda_2 - \dfrac{1-q}{q}\dfrac{R}{\Delta} \geq 0 \\ 0 & \text{otherwise.} \end{cases}\end{aligned} \tag{A6}$$

Assume $IC_1(b)$ and $IC_2(g)$ do not bind so these constraints can be neglected. (We will check this later.)

Claim 1. It is optimal to set $w_2^G(gg) = w_1^0(bb) = 0$.

Proof. Obvious.

Claim 2. $IC_2(b)$ binds at the optimum.

Proof. Suppose not. Then $w_2^0(bb) = 0$ (or else the principal should lower $w_2^0(bb)$), and $\theta^0(bb) < 1$. But without violating any constraints we can raise $\theta^0(bb)$ by $\epsilon > 0$, which raises the principal's welfare by

$$\Delta(\lambda_1(1 - q) + (1 - \lambda_1)q)(p(b) - \lambda_2)\epsilon. \tag{A7}$$

This is positive by (A4).

Claim 3. $IC_1(g)$ binds at the optimum.

Proof. Suppose not. Then $w_1^G(gg) = 0$ (or else profit can be increased by reducing $w_1^G(gg)$) and $\theta^G(bg) < 1$. Hence, $\theta^G(bg)$ can be increased without violating $IC_1(g)$ while relaxing $IC_2(b)$. But when $IC_2(b)$ is relaxed the principal can be made better off as in the proof of claim 2.

Claim 4. It is optimal to set $w_2^0(bb) = 0$.

Proof. Suppose $w_2^0(bb) > 0$. As $IC_2(b)$ binds, $\theta^0(bb) > 0$. Now $\theta^0(bb)$ can be decreased by ϵ and $w_2^0(bb)$ decreased by ϵR. (Recall we are neglecting $IC_1(b)$ and $IC_2(g)$). This increases the principal's payoff by

$$\epsilon\Delta(\lambda_1(1 - q) + (1 - \lambda_1)q)\left(\lambda_2 - p(b) + \frac{R}{\Delta}\right),$$

which is positive by (A4). This proves the claim.

Since $w_2^0(bb) = 0$, we can use the two binding constraints $IC_1(g)$ and $IC_2(b)$ to solve for $w_1^G(gg)$ and $\theta^0(bb)$ as functions of $\theta^G(bg)$ and $\theta^B(bg)$. More precisely,

$$\theta^0(bb) = p(b)\theta^G(bg) + (1 - p(b))\theta^B(bg) \tag{A8}$$

from $IC_2(b)$ and

$$w_1^G(gg) = \theta^G(bg)R + \frac{(1 - \lambda_1)(1 - q)}{\lambda_1 q}\theta^B(bg)R - R \tag{A9}$$

from $IC_1(g)$. Substituting into the principal's objective function and maximizing subject to $w_1^G(gg) \geq 0$ yields (A5) and (A6). We obtain $w_1^G(gg)$ by substituting these expressions into (A9). Finally, it can be checked that by setting $\theta^G(gb) = 1$ and $\theta^B(gb) = 0$, the omitted constraints are automatically satisfied. In fact, $IC_2(g)$ is trivial and $IC_1(b)$ becomes $\theta^0(bb) \geq p(b)$. Using (A8), this requires

$$p(b)\theta^G(bg) + (1 - p(b))\theta^B(bg) \geq p(b). \tag{A10}$$

Using (A5) and (A6) and $q > 1 - q$, it is straightforward to check that (A10) is satisfied. *Q.E.D.*

We now turn to the optimal renegotiation-proof contracts. An argument along the lines of Lemma 1 shows that $h(bb) = 0$. Suppose the simplified program in Section 3 is modified by not imposing $h(gg) = h(gb) = h(bg) = 1$, but instead imposing (10). We can assume without loss of generality that $w_1^0(gg) = 0$. Otherwise, if $h(gg) > 0$, increase $w_1^G(gg)$ to keep expected wage payments to agent 1 when $m = gg$ constant. This leaves the principal's payoff and the left-hand side of $IC_1(g)$ unchanged. (If $h(gg) = 0$, the principal is simply not implementing any projects and can set all wages equal to zero.) A similar argument also establishes that we can set $w_2^0(gg) = 0$ without loss of generality.

Therefore, the program becomes: maximize the principal's expected payoff

$$\begin{aligned}
&h(gg)(\lambda_1 q(G + \Delta - w_1^G(gg) - w_2^G(gg)) + (1 - \lambda_1)(1 - q)(B + \lambda_2\Delta))\\
&\quad + (1 - h(gg))(\lambda_1 q(\theta^0(gg)\Delta + (1 - \theta^0(gg))\lambda_2\Delta) + (1 - \lambda_1)(1 - q)(1 - \theta^0(gg))\lambda_2\Delta)\\
&\quad\quad + \lambda_1(1 - q)(\theta^0(bb)\Delta + (1 - \theta^0(bb))\lambda_2\Delta - w_1^0(bb) - w_2^0(bb))\\
&\quad\quad\quad + (1 - \lambda_1)q((1 - \theta^0(bb))\lambda_2\Delta - w_1^0(bb) - w_2^0(bb))
\end{aligned}$$

subject to

$IC_1(g)$:

$$h(gg)\lambda_1 q(w_1^G(gg) + R) + (1 - h(gg))(\lambda_1 q + (1 - \lambda_1)(1 - q))\theta^0(gg)R \\ \geq h(bg)\lambda_1 qR + (1 - h(bg))(\lambda_1 q + (1 - \lambda_1)(1 - q))\theta^0(bg)R \tag{A11}$$

$IC_1(b)$:

$$(\lambda_1(1 - q) + (1 - \lambda_1)q)(w_1^0(bb) + \theta^0(bb)R) \\ \geq h(gb)\lambda_1(1 - q)R + (1 - h(gb))(\lambda_1(1 - q) + (1 - \lambda_1)q)\theta^0(gb)R \tag{A12}$$

$IC_2(g)$:

$$h(gg)(\lambda_1 q w_2^G(gg) + (1 - \lambda_1)(1 - q)R) + (1 - h(gg))(\lambda_1 q + (1 - \lambda_1)(1 - q))R(1 - \theta^0(gg)) \\ \geq h(gb)(1 - \lambda_1)(1 - q)R + (1 - h(gb))(\lambda_1 q + (1 - \lambda_1)(1 - q))R(1 - \theta^0(gb)) \tag{A13}$$

$IC_2(b)$:

$$(\lambda_1(1 - q) + (1 - \lambda_1)q)(w_2^0(bb) + (1 - \theta^0(bb))R) \\ \geq h(bg)(1 - \lambda_1)qR + (1 - h(bg))(\lambda_1(1 - q) + q(1 - \lambda_1))(1 - \theta^0(bg))R \tag{A14}$$

and the limited liability constraints:

$$w_1^G(gg),\ w_2^G(gg),\ w_1^0(bb),\ w_2^0(bb) \geq 0. \tag{A15}$$

Proof of Lemma 2. We claim that if the inventor is ahead but not well ahead, then maximizing the principal's expected payoff subject to (A11)–(A15) results in $h(bg) = 0$, $h(gg) = h(gb) = 1$. If (13) holds, then $\theta^0(bb) = \theta^0(bg) = 1$ and

$$w_1^G(gg) = \frac{1 - p(g)}{p(g)}R.$$

Otherwise, $\theta^0(bb) = \theta^0(bg) = p(g)$ and $w_1^G(gg) = 0$. All other wages are zero.

We shall disregard $IC_1(b)$ and $IC_2(g)$ for now, showing later that they are satisfied. Without these constraints, $w_1^0(bb) = w_2^G(gg) = 0$ is certainly optimal.

Claim 1. It is optimal to set $h(gg) = 1$.

Proof. If $h(gg) < 1$, then increase $h(gg)$ by ϵ and, if possible, alter $w_1^G(gg)$ to keep the left-hand side of $IC_1(g)$ constant. This will change the expected wage payments, which contains the component $\lambda_1 q h(gg) w_1^G(gg)$. Let ϕ denote the change in this component due to the changes in $h(gg)$ and $w_1^G(gg)$. There are two cases:

Case 1. If (i) $\theta^0(gg) \geq p(g)$ or (ii) if $\theta^0(gg) < p(g)$ and $w_1^G(gg) > 0$, the change in expected wages $\lambda_1 q h(gg) w_1^G(gg)$ needed to keep the left-hand side of $IC_1(g)$ constant is

$$\phi = \epsilon(\lambda_1 q + (1 - \lambda_1)(1 - q))\theta^0(gg)R - \epsilon\lambda_1 qR = \epsilon(1 - \lambda_1)(1 - q)\theta^0(gg)R - \epsilon\lambda_1 q(1 - \theta^0(gg))R. \tag{A16}$$

If (i) holds, this expression is nonnegative. If (ii) holds, it is negative, so $w_1^G(gg)$ may have to be reduced, but this is possible since $w_1^G(gg) > 0$. In any case, the principal's expected payoff changes by

$$\epsilon(\lambda_1 q(G + \Delta) + (1 - \lambda_1)(1 - q)(B + \lambda_2\Delta)) - \epsilon\lambda_1 q\theta^0(gg)\Delta \\ - \epsilon(\lambda_1 q + (1 - \lambda_1)(1 - q))(1 - \theta^0(gg))\lambda_2\Delta - \phi. \tag{A17}$$

Substituting from (A16), (A17) becomes

$$\epsilon(\lambda_1 qG + (1 - \lambda_1)(1 - q)B) + (1 - \lambda_1)(1 - q)\theta^0(gg)(\lambda_2\Delta - R) + \epsilon\lambda_1 q(1 - \theta^0(gg))R \\ + \epsilon\lambda_1 q(1 - \theta^0(gg))\Delta(1 - \lambda_2),$$

which is strictly positive by Assumptions 1 and 3. Hence, it is optimal to raise $h(gg)$.

Case 2. If $\theta^0(gg) < p(g)$ and $w_1^G(gg) = 0$, the left-hand side of $IC_1(g)$ will increase when $h(gg)$ is increased by ϵ. As $w_1^G(gg) = 0$, this perturbation changes the principal's expected payoff by

$$\epsilon(\lambda_1 q(G + \Delta) + (1 - \lambda_1)(1 - q)(B + \lambda_2\Delta)) - \epsilon\lambda_1 q\theta^0(gg)\Delta - \epsilon(\lambda_1 q + (1 - \lambda_1)(1 - q))(1 - \theta^0(gg))\lambda_2\Delta$$
$$= \epsilon(\lambda_1 qG + (1 - \lambda_1)(1 - q)B) + \epsilon\lambda_1 q(1 - \theta^0(gg))\Delta(1 - \lambda_2) + (1 - \lambda_1)(1 - q)\theta^0(gg)\lambda_2\Delta,$$

which is positive by Assumption 1. Hence, it is optimal to increase $h(gg)$ in Case 2 too, and the claim is proved.

Claim 2. $IC_1(g)$ and $IC_2(b)$ bind, and either $\theta^0(bb) = 1$ or $w_1^0(bb) = 0$.

Proof. Suppose $IC_2(b)$ does not bind. Then $w_2^0(bb) = 0$ or else $w_2^0(bb)$ could be lowered, and hence $(1 - \theta^0(bb))R > 0$. Now raise $\theta^0(bb)$ by $\epsilon > 0$. This respects all constraints and increases the principal's payoff by

$$\epsilon\Delta(\lambda_1(1 - q) - (\lambda_1(1 - q) + (1 - \lambda_1)q)\lambda_2) = (\lambda_1(1 - q) + (1 - \lambda_1)q)\epsilon\Delta(p(b) - \lambda_2) > 0, \quad \text{(A18)}$$

since the inventor is ahead. Therefore, $IC_2(b)$ must bind.

Suppose $IC_1(g)$ does not bind and recall from Claim 1 that $h(gg) = 1$. Then, $w_1^G(gg) = 0$. Except for the slack constraint $IC_1(g)$, $h(bg)$ and $\theta^0(bg)$ only enter $IC_2(b)$, which binds. Reducing the right-hand side of $IC_2(b)$ is advantageous because the principal can either lower $w_2^0(bb)$ or raise $\theta^0(bb)$ (the latter is strictly advantageous from (A18)). Therefore, if $IC_1(g)$ is slack, the right-hand side of $IC_2(b)$ must already be zero, which implies $\theta^0(bg) = 1$ and $h(bg) = 0$. But then $IC_1(g)$ is violated, a contradiction. Therefore $IC_1(g)$ must bind.

Finally, suppose $\theta^0(bb) < 1$ and $w_1^0(bb) > 0$. Then, by raising $\theta(bb)$ by ϵ, reducing $w_1^0(bb)$ by ϵR, and increasing $w_2^0(bb)$ by ϵR, the sum of the wages is constant, all constraints are respected, and the principal's payoff goes up by (A18)). This proves the claim.

Claim 3. At the optimum, $\theta^0(bb) > 0$ and $w_2^0(bb) = 0$.

Proof. If $\theta^0(bb) = 0$, then as $IC_2(b)$ binds we must have $h(bg) = \theta^0(bg) = 0$. Then the right-hand side of $IC_1(g)$ is zero, but this contradicts the fact that $IC_1(g)$ is binding and $h(gg) = 1$. Thus, $\theta^0(bb) > 0$.

Now suppose $w_2^0(bb) > 0$. As $\theta^0(bb) > 0$, we can lower $\theta^0(bb)$ by ϵ and $w_2^0(bb)$ by ϵR without violating any constraints (recall we are omitting $IC_1(b)$ from the program). The principal's expected payoff goes up by

$$\epsilon\Delta(\lambda_1(1 - q) + (1 - \lambda_1)q)\left(-p(b) + \lambda_2 + \frac{R}{\Delta}\right) > 0,$$

since the inventor is not well ahead. This proves the claim.

From $IC_2(b)$ we now have

$$\theta^0(bb) = h(bg)p(b) + (1 - h(bg))\theta^0(bg) \quad \text{(A19)}$$

and from $IC_1(g)$ (as $h(gg) = 1$),

$$w_1^G(gg) = (1 - h(bg))\left[\frac{\theta^0(bg)}{p(g)} - 1\right]R. \quad \text{(A20)}$$

Substituting (A19) and (A20) into the principal's objective function, we find that the principal should choose $h(bg)$ and $\theta^0(bg)$ to maximize the expression

$$-(1 - h(bg))\lambda_1\left[q\left[\frac{\theta^0(bg)}{p(g)} - 1\right]R + (1 - q)(\lambda_2 - p(b))\left[\frac{\theta^0(bg)}{p(b)} - 1\right]\Delta\right] \quad \text{(A21)}$$

subject to $w_1^G(gg) \geq 0$, where $w_1^G(gg)$ is given by (A20). The derivative of (A21) with respect to $\theta^0(bg)$ is

$$-(1 - h(bg))\lambda_1\frac{1 - q}{p(b)}\left(\frac{q}{1 - q}\frac{p(b)}{p(g)}R + (\lambda_2 - p(b))\Delta\right).$$

The expression in the big parenthesis is the expression that occurs in (13). Now it can be verified that the solution involves $\theta^0(bg) = 1$ if (13) holds, and $\theta^0(bg) = p(g)$ otherwise. In both cases, $h(bg) = 0$ is optimal.

Equation (A19) implies $\theta^0(bg) = \theta^0(bb)$. Finally, $IC_1(b)$ and $IC_2(g)$ are satisfied at zero cost by setting $h(gb) = 1$. *Q.E.D.*

Proof of Lemma 3. Suppose the inventor is well ahead. We need to show that the problem of maximizing the principal's expected payoff subject to (A11)–(A15), and with the extra constraint $\theta^0(m) = 1$ for all m, is solved by $h(gb) = 1$ and $h(bg) = 0$. All wages are zero, except that

$$w_1^G(gg) = \frac{1 - p(g)}{p(g)}R.$$

We shall disregard $IC_1(b)$ and $IC_2(g)$ for the moment. Then $w_1^0(bb) = w_2^G(gg) = 0$ is optimal. As in the proof of Lemma 2, one can show that $IC_1(g)$ and $IC_2(b)$ must bind. This implies

$$w_1^G(gg) = (1 - h(bg))\frac{1 - p(g)}{p(g)}R \qquad w_2^0(bb) = h(bg)(1 - p(b))R.$$

Substituting into the principal's expected payoff, and using $\theta^0(m) = 1$ for all m, we find that the principal should choose $h(bg)$ to maximize the expression

$$-(1 - \lambda_1)((1 - h(bg))(1 - q) + h(bg)q)R.$$

The solution is $h(bg) = 0$. Moreover, by setting $h(gb) = 1$, we guarantee that $IC_1(b)$ and $IC_2(g)$ hold. *Q.E.D.*

References

AGHION, P. AND TIROLE, J. ''Formal and Real Authority in Organizations.'' *Journal of Political Economy,* Vol. 105 (1997), pp. 1–29.

BAKER, G., GIBBONS, R., AND MURPHY, K.J. ''Subjective Performance Measures in Optimal Incentive Schemes.'' *Quarterly Journal of Economics,* Vol. 109 (1994), pp. 1125–1156.

BALIGA, S. AND SJÖSTRÖM, T. ''Optimal Design of Peer Review and Self-Assessment Schemes.'' Northwestern University Math Center Paper no. 1290, 1997 (www.kellogg.nwu.edu/research/math/Down-papers.htm).

CALVO, G.A. AND WELLISZ, S. ''Hierarchy, Ability and Income Distribution.'' *Journal of Political Economy,* Vol. 87 (1979), pp. 991–1010.

CARMICHAEL, H.L. ''Incentives in Academics: Why Is There Tenure?'' *Journal of Political Economy,* Vol. 96 (1988), pp. 453–472.

COLEMAN, J.S. *Foundations of Social Theory.* Cambridge, Mass.: Harvard University Press, 1990.

DEWATRIPONT, M. AND MASKIN, E. ''Contractual Contingencies and Renegotiation.'' *RAND Journal of Economics,* Vol. 26 (1995), pp. 704–719.

GAMACHE, R.D. AND KUHN, R.L. *The Creativity Infusion.* New York: Harper and Row, 1989.

HOLMSTRÖM, B. ''Moral Hazard in Teams.'' *Bell Journal of Economics,* Vol. 13 (1982), pp. 324–340.

———. ''Managerial Incentive Problems: A Dynamic Perspective.'' *Review of Economic Studies,* Vol. 66 (1999), pp. 169–182.

ITOH, H. ''Incentives to Help in Multi-Agent Situations.'' *Econometrica,* Vol. 59 (1991), pp. 611–636.

LAZEAR, E.P. ''Pay, Inequality and Industrial Politics.'' *Journal of Political Economy,* Vol. 97 (1989), pp. 561–580.

LEVITT, S.D. ''Optimal Incentive Schemes When Only the Agent's 'Best' Output Matters to the Principal.'' *RAND Journal of Economics,* Vol. 24 (1995), pp. 744–760.

——— AND SNYDER, C.M. ''Is No News Bad News? Information Transmission and the Role of 'Early Warning' in the Principal-Agent Model.'' *RAND Journal of Economics,* Vol. 28 (1997), pp. 641–661.

MASKIN, E. AND TIROLE J. ''The Principal-Agent Relationship with an Informed Principal, II: Common Values.'' *Econometrica,* Vol. 60 (1992), pp. 1–42.

MILGROM, P. ''Employment Contracts, Influence Activities and Efficient Organization Design.'' *Journal of Political Economy,* Vol. 96 (1988), pp. 42–60.

MYERSON, R. *Game Theory.* Cambridge, Mass.: Harvard University Press, 1991.

PRENDERGAST, C. AND STOLE, L. ''Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning.'' *Journal of Political Economy,* Vol. 104 (1996), pp. 1105–1134.

——— AND TOPEL, R. ''Favoritism in Organizations.'' *Journal of Political Economy,* Vol. 104 (1996), pp. 958–978.

ROTEMBERG, J. AND SALONER, G. ''Overt Interfunctional Conflict (and Its Reduction Through Business Strategy).'' *RAND Journal of Economics*, Vol. 26 (1995), pp. 630–653.

SCOTCHMER, S. ''Puzzles: Professional Advice and Other Hazards.'' *Journal of Economic Perspectives,* Vol. 4 (1990), pp. 189–195.

SIMON, H. ''A Formal Theory of the Employment Relationship.'' *Econometrica* Vol. 19 (1951), pp. 293–305.

TIROLE, J. ''Collusion and the Theory of Organizations.'' In J.-J. Laffont, ed., *Advances in Economic Theory: Sixth World Congress.* New York: Cambridge University Press, 1992.

WILLIAMSON, O.E. *Markets and Hierarchies,* 2d ed. New York: Free Press, 1983.

———. *The Economic Institutions of Capitalism.* New York: Free Press, 1987.