Withholding Performance Information

Korok T. Ray *

September 24, 2004

Abstract

It is a common practice for firms to conduct performance evaluations of their employees and yet to withhold this information from those employees. This paper argues that firms strategically withhold performance information to retain workers. In particular, if the worker enjoys high outside options and is tempted to quit, the firm chooses not to disclose this information to keep him on the job. The firm's equilibrium strategy is to fire if performance is sufficiently low, reveal information if performance is sufficiently high, and withhold information otherwise. This reflects the empirical distribution of performance evaluation ratings observed in practice.

1 Introduction

Firms are often reluctant to precisely reveal an employee's performance evaluation to himself. The literature on performance evaluations extensively documents the tendency of managers to pool employee performance ratings together (Landy and Farr (1980), Mohrman and Lawler (1983), Murphy and Cleveland (1991)). For example, when Morgan Stanley introduced the 360 degree performance evaluation system, the bank explicitly required the "Evaluation Director" to summarize and filter many of the evaluation reports collected from the employee's colleagues (Burton 1998, 2000). This provides firms with the option to withhold performance information, possibly for strategic purposes.

The traditional explanation from industrial psychology is that revealing low output to employees kills morale and generates uncooperative behavior,

*University of Chicago Graduate School of Business. This paper is based on Chapter 2 of my dissertation submitted to Stanford GSB. I'd like to thank Edward Lazear, John Roberts, Andy Skryzpacz, and Madhav Rajan for helpful comments. Stanford GSB and Chicago GSB provided financial support. My email is korok@ChicagoGSB.edu

and thus managers tell all workers that they are above average. This paper offers an economic explanation of this phenomenon: managers withhold performance information to retain workers.

To arrive at this main result, I construct a two stage agency model consisting of a risk-neutral principal (the firm) contracting with a risk-neutral agent (the worker) under limited liability. The key assumption is that the firm observes the worker's performance but the worker himself does not. This is a reasonable assumption in practice since managers who conduct extensive performance evaluations must choose how much of the information from that evaluation to reveal to the worker.

The condition guaranteeing pooling is that the worker's outside options are (sufficiently) high while the firm's outside options are (sufficiently) low. The worker receives an incentive contract, so his productivity determines his pay. Production is complementary across stages, so low performance today signals low performance, and hence low pay, tomorrow. With high outside options, he is tempted to leave. But if the firm's outside options are sufficiently low, the firm prefers to keep the worker rather than capture its low outside options. Thus the worker wants to leave while the firm wants him to stay, so the firm withholds performance information to keep him on the job.

The firm does this by pooling the output together of workers who are tempted to quit. The firm selects a pool large enough such that the average member of the pool stays on the job. Since workers who are pooled do not know where in the pool they sit, they rationally assume their output is average within the pool. Since the average member of the pool stays on the job, so does everyone within the pool. Thus the firm strategically pools to guarantee retention.

Now pooling is costly for the firm. Production is complementary across stages, so high output today increases the marginal productivity of output tomorrow. In particular, the agent works harder after a successful early stage because the marginal return from his labor is higher. The agent bases his second stage effort choice on first-stage output. For workers within the pool, they base this only on what they know: average first stage output within the pool. Therefore, workers at the high end of the pool (the stars) work less than if they knew their early output precisely. Similarly, workers at the low end of the pool (the slugs) work more. Since profits are increasing in second stage effort, the firm loses money on the stars but gains on the slugs. The main result shows that profit from retaining the slugs who otherwise would have quit exceeds the loss in output from the stars. Pooling is costly, but the benefits from retention outweigh those costs.

The firm's equilibrium strategy takes a simple form: fire if output is low, reveal fully if output is high, and pool in between. This is consist with performance evaluations in practice, where workers are either laid off, promoted, or told that they are average (Murphy and Cleveland (1991), Landy and Farr (1980)). Moreover, the model predicts that managers pool more in jobs where workers have higher outside options. This is consistent with a study on performance reviews in the Navy (Bjerke et al (1987)), where performance evaluations among the more prestigious ranks exhibit greater pooling and less performance differentiation.

The paper is organized as follows. Section 2 presents the basic model and solves for the first best outcome. Section 3 establishes the principal and agent's payoff functions and shows that how the parties share the outside option determines their intermediate termination decision. Section 4 presents the main result and gives intuition. Section 5 concludes.

2 The Model

An employee (an agent) works on a project for a firm over two stages, there is no discounting, and both parties are risk neutral. The agent exerts effort e_t in stage $t = 1, 2$ at cost $C(e_t)$, where C', C'', C''' are strictly positive. Costs are separable, increasing, convex, and marginalcosts are convex. He produces output $q_t = e_t + \epsilon_t$. The noise terms ϵ_t are i.i.d., and distributed around a mean of zero, with cdf $G(\cdot)$ and density function $g(\cdot)$. The agent observes his effort but the firm does not.

Production is given by a pair of functions (W, V), where $W : \mathbb{R} \to \mathbb{R}$ and $V : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$. $W(q_1)$ is the value of output after stage one, and $V(q_1, q_2)$ is the value of output after stage two. Thus, q_t is the project's *internal* output within the firm used for planning and evaluation purposes, while $W(q_1), V(q_1, q_2)$ measure the project's *external* value based on market prices. Production is *indivisible* if $W(q_1) = 0$ for all q_1, so work after the first stage alone has no value. The firm cannot sell the output from the first stage q_1, as it fetches a price of zero in the market. Only after the second stage can the firm collect $V(q_1, q_2)$ from the market. Throughout the paper assume that production is indivisible and given by

$$V(q_1, q_2) = \begin{cases} Vq_1q_2 & \text{if } q_1, q_2 > 0 \\ 0 & \text{else} \end{cases}$$

Note that $V_{12}(q_1, q_2) = V > 0$ if $q_1, q_2 > 0$, so production is complementary across stages. In particular, $V_2 = Vq_1$, so a marginal increase in q_1 increases the marginal productivity of second stage output. This production function fits many scenarios (such as R&D, new product development) in which work takes place sequentially in stages, and where early stage output affects later stage productivity.[1] For example, if a product runs through research and development stages, then a highly promising research stage increases the marginal productivity of effort and output in the development stage. It is much easier to create a high quality product if the prototype is very successful.

After the second stage, output (q_1, q_2) is observable to both parties, but only the firm observes q_1 after the first stage. The firm can choose to reveal q_1 to the agent or not. This reflects that firms often have access to performance information before workers do, and that revealing this information is a choice the firm can make. The agent will use q_1 first to decide whether to quit the project or not, and second to select his appropriate second stage effort level. Once revealed, performance information is ex-post verifiable by agent.[2] So if the principal decides to reveal output, he will do it truthfully. This focuses the problem on whether to reveal information at all, and not on whether to distort revealed information.

Let the pooling region $Q \subset \mathbb{R}$ denote the region over which the firm conceals output. Assume the pooling regions are finite unions of disjoint intervals, so $Q = \bigcup_{i \in I} Q_i$ where I is a finite set and Q_i's are disjoint. These intervals Q_i represent performance categories often seen in practice. For example, a worker may only know that is performance is "good," where "good" means a performance rating between 5 to 7 on a scale of 1 to 10. Workers are told in which category their performance lies, so they know not just that $q_i \in Q$ but that $q_i \in Q_k$ for some k. Later I show that in equilibrium the firm selects a Q that is a single interval $(\underline{q}, \bar{q})$.

The principal proposes to the agent a contract (q^p, Q, s, b), where q^p is the firing rule (fire if $q_1 < q^p$), Q is the pooling region, s is a salary paid up-front to the worker, and b is the bonus paid on final output $q_1 q_2$. So the firm receives Vq_1q_2 and the agent $s + bq_1q_2$. I assume the agent is constrained by limited liability (LL), so $s, b \geq 0$. Given this contract, the

[1] A more 'natural' production function might be linear production with persistent ability: $q_t = a + e_t + \epsilon_t$ and $V(q_1, q_2) = Vq_1 + Vq_2$ if $q_1, q_2 > 0$ and zero otherwise. This production function has similar properties as the one I use above, but is significantly less tractable.

[2] This is justified by the usual repeated game arguments. The agent can punish the principal in the future (file a lawsuit, sabotage output, etc), and hence the value of repeated interaction provides incentives to the principal to report truthfully.

agent responds by choosing actions $(q^a, e_1, e_2(\cdot))$. He chooses a quit rule q^a, effort e_1 and an *effort function* $e_2(\cdot)$. Note that the agent's effort function $e_2(\cdot)$ must be constant over each interval Q_i, since the agent cannot condition on information he does not have.

Finally, both the principal and the agent have outside options in each stage. This captures the value from quitting the project and dedicating resources (labor, capital) elsewhere. Let $\bar{u_t}^a$ and $\bar{u_t}^p$ denote the agent's and the principal's outside option in stage t. Call $\bar{u_2} = \bar{u_2}^a + \bar{u_2}^p$ the residual surplus: the total surplus from abandoning the project. Finally, the principal and agent must satisfy a participation constraint (PC) that their equilibrium payoffs must exceed $\bar{u_1}^p + \bar{u_2}^p$ and $\bar{u_1}^a + \bar{u_2}^a$ respectively.

2.1 First Best

This section establishes the first best benchmark useful for the rest of the paper: it is efficient for the firm to fully reveal q_1 after the first stage, and it is efficient to quit projects with low early output. The first result states that more information is always socially optimal, and thus suggests that strategic information revelation generates welfare losses. The second result shows that it is efficient to terminate bad projects because firms and workers have outside options elsewhere. All the first best results here hold under general indivisible production functions $V(q_1, q_2)$.

As mentioned earlier (and proved in later sections), the firm will withhold information to retain the worker. Precisely, there are conditions under which a worker prefers to leave the firm but the firm wants him to stay. These conflicts of interest are absent in the social planner's problem. The planner will simply terminate workers at the efficient rate, and hence there is no reason to withhold performance information.

Proposition 1 *It is efficient to fully reveal output after stage one.*

In fact, withholding information makes the planner (weakly) worse off, since it prevents second stage effort from accurately conditioning on first stage output. To see this, suppose Q_i is a nontrivial pooling region. If the planner chooses to pool, then the (second stage) effort function must be constant over Q_i. If the planner reveals information, then the effort function is unconstrained over Q_i. So pooling shrinks the planner's choice set of all possible effort functions. This forces the planner to optimize over a smaller set, and hence this decreases total surplus. This is the logic behind Proposition 1; see the appendix for the proof.

A key figure in the planner's optimization problem is the continuation surplus function

$$S(q_1) = \mathbb{E}V(q_1, e_2(q_1) + \epsilon_2) - C(e_2(q_1)).$$

This is the total surplus from continuing, for each realization of first stage output q_1. In the proof of the proposition below I show that the effort function is increasing if and only if production is complementary across stages ($V_{12}(q_1, q_2) > 0$). So the agent works harder after a successful early stage because the marginal return to his labor is now higher. The social planner's termination rule will take the form of a cut-off rule. In other words, terminating low output workers is efficient.

Proposition 2 *There exists a cut off point q^* such that it is efficient only for workers with $q_1 > q^*$ to advance to the second stage.*

Here, q^* is the efficient termination rule. The proof of this proposition uses the first order conditions on $e_2(\cdot)$ to show that the continuation surplus function is increasing. Now we can write the social planner's problem as

$$\max_{e_1, e_2(\cdot), \tau} \int_{\tau} S(q_1) g(q_1 - e_1) dq_1 + (1 - P_1)\bar{u_2} - C(e_1)$$

where $\tau = q^*$ is the optimal termination rule and $P_1 = Prob(q_1 > \tau) = 1 - G(\tau - e_1)$ is the probability of continuing.

As usual, it is possible to implement the first best with a standard sell-out contract. The principal holds the agent to his participation constraint with negative salary ($s < 0$), and grants full incentives on output ($b = V$). So the agent pays the firm upfront the equilibrium value of output less equilibrium cost of effort, and the firm is able to give the agent a full share in the output from production.

3 Second Best

Under limited liability, the firm cannot implement the first best. The principal no longer can hold the agent to his participation constraint with negative wages, and so the principal cannot seize all the rents. The principal is worse off, the agent is better off, and total surplus shrinks (relative to first best). Now the principal has an incentive to selectively reveal performance in order to maximize her profits. Moreover, the principal will in general set a different target than the agent, i.e. the agent's quit rule will differ from the principal's firing rule.

The analysis proceeds as follows. First I establish properties of the principal and agent's payoff functions. Next, I examine why it is important to assume that the principal can commit to a revelation strategy. Then I show that how the principal and agent share the residual surplus determines their firing and quit rates, respectively. Finally, I use this result to establish when and why the principal will establish a nonempty pooling region.

3.1 Payoffs

Fix a contract (q^p, Q, s, b), where q^p is the firing rule (fire if $q_1 < q^p$), Q is the pooling region, s is the salary, and b is the bonus on final output $q_1 q_2$. Given this contract, the agent responds by choosing $(q^a, e_1, e_2(\cdot))$: a quit rule q^a, first stage effort, and second stage effort function. I will refer to q^p and q^a as the principal and agent's targets, hurdles, and termination rules interchangeably. Let $\tau = \max(q^p, q^a)$. Since output must clear both hurdles for the agent to advance, the probability of advancing is $P_1 = Pr(q_1 > \tau) = 1 - G(\tau - e_1)$.

Now the principal's and agent's continuation utilities are

$$u(q_1) = \mathbb{E}_{q_2}[bq_1q_2 - C(e_2(q_1)) = bq_1e_2(q_1) - C(e_2(q_1))]$$
$$\pi(q_1) = \mathbb{E}_{q_2}[(V-b)q_1q_2 = (V-b)q_1e_2(q_1)]$$

where the expectation is taken over q_2. The shape of these functions depends on the shape of the agent's effort function $e_2(\cdot)$. This function must be constant over each pooling interval Q_i, since the agent cannot condition on information he does not have. Call this the Effort Constraint (EC):

$$\text{(EC)} \qquad e_2(q_1) = \hat{e_i} \qquad \forall q_1 \in Q_i \quad \texttt{for each} \quad i \in I$$

where $\hat{e_i} \in \mathbb{R}$ is the level of the function over Q_i. Recall that the agent is told not just that $q_1 \in Q$ but that $q_1 \in Q_k$ for some k.

Given a contract (q^p, Q, b), the agent's problem is

$$\max_{q^a, e_1, e_2(\cdot)} \int_\tau u(q_1) g(q_1 - e_1) dq_1 + (1 - P_1)\bar{u_2}^a - C(e_1) + s$$
$$\text{subject to (EC), (PC)}$$

where τ is a function of q^a and (PC) is the participation constraint. In words, if $q_1 > \tau$ then the agent advances and receives $u(q_1)$ integrated

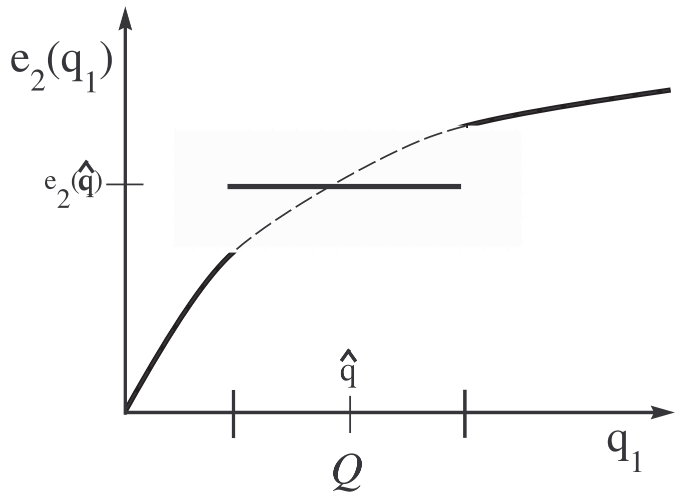

Figure 1: Effort Function.

over each $q_1 > \tau$. If not, he receives $\bar{u_2}^a$. He bears $C(e_1)$ regardless of whether he advances or not, but bears $C(e_2(q_1))$ only if he advances. Note that $C(e_2(q_1))$ is embedded in $u(q_1)$ and so does not appear in the above optimization explicitly.

The first order conditions from the agent's problem shows that q^a satisfies $u(q^a) = \bar{u_2}^a$. At the optimum, the agent is indifferent between staying and leaving. From the FOCs the effort function satisfies $C'(e_2(q_1)) = bq_1$ for each $q_1 \notin Q$ and from (EC) $e_2(q_1) = \hat{e}_i$ for each $q_1 \in Q_i$ for each $i \in I$. Now we can use the agent's problem to solve for $\hat{e}_i$ directly and arrive at the shape of the effort function.

Proposition 3 *The effort function $e_2(q_1)$ is constant over each Q_i and increasing and concave elsewhere. In particular $e_2(q_1) = e_2(E[q_1|q_1 \in Q_i])$ for all $q_1 \in Q_i$ for each $i \in I$. The principal and agent's payoff functions are increasing and convex.*

The figure above graphs the optimal effort function. For clarity take Q to be a single interval; the proof solves for the general case of $Q = \bigcup_i Q_i$.

Over Q the function is constant at the average value $\hat{e_2}$, where $\hat{e_2}$ satisfies $C'(\hat{e_2}) = b\hat{q}$ and $\hat{q} = E[q_1 | q_1 \in Q]$ is the average output over Q. Even though the effort function is technically only defined off of Q, the proof of the proposition shows how optimally to extend the function over Q. First, create a new function f by projecting $e_2(q_1)$ over Q according to $C'(f(q_1)) = bq_1$. In the picture, f is the dashed line. This function f is the same as the effort function under no pooling. Second, define the effort function over Q by $e_2(q_1) = f(\hat{q})$ for all $q_1 \in Q$. Hence (EC) can be rewritten as $e_2(q_1) = e_2(\hat{q})$ for all $q_1 \in Q$. In words, the effort function is constant over Q at a value based on average output over Q. This is intuitive: since the agent does not know where in Q his output lies, he assumes he produced the average $\hat{q}$ and chooses second stage effort according to this average for *all* $q_1 \in Q$. Thus, pooling simply replaces the effort over the pool with the effort of the average over the pool. The proof of the proposition shows that this not just a natural but in fact the optimal choice of effort over the pool.

The shape of the effort function determines the shape of the principal's (and agent's) payoff functions. The principal and the agent's continuation payoff functions are pictured in the next figure. Even though the effort function is constant over Q and concave off of Q, the function $q_1 e_2(q_1)$ is linear over Q and strictly convex off of Q, and therefore weakly convex everywhere. So the payoff functions are also weakly convex, apparent from their definition. The principal's problem is

$$\max_{q^p, s, b, Q} \int_{\max(q^a, q^p)} \pi(q_1) g(q_1 - e_1) dq_1 + (1 - P_1)\bar{u_2}^p - s$$
$$\text{subject to (EC), (PC), (LL)}$$

The first order conditions show that the principal will set q^p such that $\pi(q^p) = \bar{u_2}^p$. Just like the agent, the principal is indifferent between continuing employment and firing the agent to collect her outside option. The principal will set the lowest possible salary level to guarantee participation, so $s = 0$. In choosing the incentive coefficient b, the principal trades off two separate forces. As she increases b, she induces more effort out of the agent (an incentive effect) which generates more output and hence more profit. But increasing b also reduces the principal's share of final output, since she earns $(V - b)q_1 q_2$. The optimal b lies in $(0, V)$, so the principal gives the agent some but not full incentives.[3] In the analysis of the optimal Q below, take b as fixed at its optimal level.

[3] The first order conditions that solve for b explicitly are messy and uninstructive, hence I omit them.

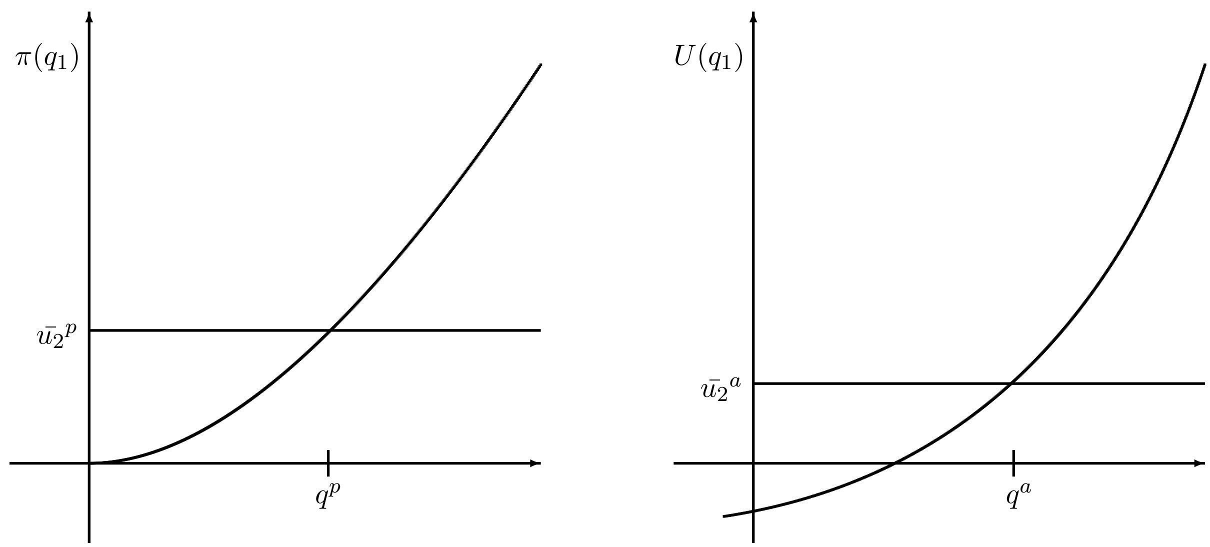

Figure 2: Principal and Agent Payoffs

Suppose the principal cannot commit to a revelation scheme. So even if the principal claims to reveal output ex-ante, she can always reverse this decision ex-post. Then:

Proposition 4 *Without commitment, any pooling region unravels.*

This holds because π is increasing. Suppose the principal claims to withhold information by pooling output over an interval $Q_i = (x, y)$. Let $\hat{q} = E[q_1 | q_1 \in (x, y)]$ be average output over this region. The agent chooses average effort $e_2 = e_2(\hat{q})$ over this region, and the principal earns average profit $\hat{\pi}(q_1) = (V - b) q_1 e_2(\hat{q})$. But if the principal observes $q_1 > \hat{q}$, she can get $\pi(q_1) > \hat{\pi}(q_1)$ if she reveals it. Since she has no commitment, she will do so. So she in fact will only pool over $(x, \hat{q})$ and separate over $(\hat{q}, y)$. Applying this same argument with the candidate pooling interval $(x, \hat{q})$ shows that the principal pools below the average of this interval but separates above it. Repeating this argument ad infinitum, the pooling region unravels. See the appendix for the full formal proof. Therefore the interesting case is when the principal can commit to a revelation strategy; assume this in what follows.

3.2 Intermediate Targets

Recall that the residual surplus $\bar{u_2} = \bar{u_2}^a + \bar{u_2}^p$ is the sum of the principal and agent's outside options and represents the value to both parties of abandoning work after stage one. The logic behind pooling equilibrium stems from two sequential results. First, the distribution of the residual surplus between the principal and agent uniquely determines the ranking of their interim performance targets q^p and q^a, respectively. How the two parties split the residual surplus $\bar{u_2}$ determines their termination decision after the first stage. Second, the ranking of the targets determines whether the principal has an incentive to pool performance. More precisely, if there exists a region in which the principal wants to keep the agent but the agent wants to leave, then the principal will withhold information (pool output) to retain the agent.

Definition 1 *Let $\gamma \equiv \bar{u_2}^a / \bar{u_2}$ denote the agent's share of the residual surplus $\bar{u_2}$.*

Because outside options satisfy $\bar{u_t} = \bar{u_t}^a + \bar{u_t}^p$, it is clear that $\gamma \in [0, 1]$. High γ means the agent captures most of the residual surplus after a failed project, so γ is one measure of the distribution of residual surplus between the two parties. In general one might not expect the intermediate output targets q^a, q^p, q^* to follow any coherent and sensible ordering. Or one might expect difficulty in explicitly comparing targets; for example, it may be possible to show $q^a > q^*$ and $q^p > q^*$ but not know whether $q^a > q^p$. Amazingly, this does not happen. It is remarkable that the distribution of residual surplus completely and uniquely determines the ordering of the targets in a clean and intuitive way:

Proposition 5 *There exists a $\gamma^* \in (0, 1)$ such that $q^a > q^* > q^p$ if and only if $\gamma > \gamma^*$.*

In words, the party that receives most of the residual surplus will set an inefficiently high target, while the other party will set an inefficiently low target. For example, suppose that $\gamma > \gamma^*$, so the agent receives most of the residual surplus. This means his outside option is high (relative to the principal), so the early returns from the project must also be high in order to justify foregoing these attractive outside opportunities. He will tolerate fewer failures since his alternatives are good, so he sets a high output hurdle q^a. In fact, he quits some projects that are efficient to continue, and so he sets $q^a > q^*$. And simultaneously the principal receives a small share of the residual surplus and has low outside options relative to the agent. So the principal does not require a high hurdle rate to justify continuation, since her alternatives outside are weak. So she sets a low output hurdle, and even continues some projects that are efficient to quit, so $q^p < q^*$. Finally, note that it is possible for both parties to quit at the efficient rate, but only if the knife-edge condition $\gamma = \gamma^*$ holds.

This result suggests that there will be two classes of equilibria, depending on whether $\gamma < \gamma^*$. The distribution of residual surplus will determine whether the equilibrium is separating or pooling.

4 Main Result

To build intuition behind when the principal will pool and when she will separate, first note that in general pooling is costly. To see this, suppose the principal pools over an interval $Q = (\underline{q}, \bar{q})$. Let $\hat{q} = E[q|q_1 \in Q]$ be the average value over the pool. Outside Q the principal earns $\pi(q_1) = (V-b)q_1 e_2(q_1)$ where $e_2(q_1)$ solves $C'(e_2(q_1)) = bq_1$. Within the pool she earns $\hat{\pi}(q_1) = (V-b)q_1 e_2(\hat{q})$. Since π is convex over Q, Jensen's inequality shows that

$$E[\pi(q_1)|q_1 \in Q] > \pi(E[q_1|q_1 \in Q]) = \pi(\hat{q}) = \hat{\pi}(\hat{q}) = E[\hat{\pi}(q_1)|q_1 \in Q].$$

The last equality holds since $\hat{\pi}$ is linear. Multiply both sides by $1/\Pr(Q)$:

$$\int_Q \pi(q_1) g(q_1 - e_1) dq_1 > \int_Q \hat{\pi}(q_1) g(q_1 - e_1) dq_1.$$

Hence separating earns more profit for the principal than pooling.

Intuitively, the principal reports only the average output $\hat{q} = E[q_1|q_1 \in Q]$ to everyone in the pool, and the agent chooses second stage effort based on this report. In particular, if an agent produces $q_1 > \hat{q}$, instead of choosing $e_2(q_1)$ as he would if he knew q_1, he chooses $e_2(\hat{q}) < e_2(q_1)$ since the effort function $e_2(\cdot)$ is increasing. In general the stars (those with $q_1 > \hat{q}$) slack off and the slugs ($q_1 < \hat{q}$) work harder, since both think that they are average. Profits are increasing in second stage effort, so the principal loses money on the stars and gains on the slugs. But because the profit function is convex, the loss exceeds the gain. The principal makes so much money off the stars that the cost of telling them that they are average exceeds the benefit of telling the slugs that they're average.

The figure above shows the costs and benefits of pooling. The convex and linear dashed lines are the principal's payoff under separating and pooling, respectively. The area under these functions is the total profit from separating and pooling. Since $\pi(q_1) > \hat{\pi}(q_1)$ iff $q_1 > \hat{q}$, the darkly shaded

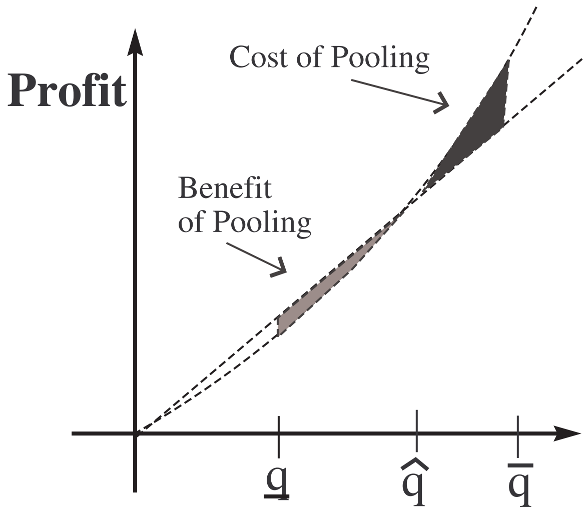

Figure 3: Costs and Benefits of Pooling.

region is the cost of pooling while the lightly shaded region is the benefit of pooling. The convexity of π and Jensen's inequality shows that the cost of pooling (the loss in output from the stars) exceeds the benefits of pooling (gain in output from the slugs). It is better to tell everyone where they stand.

Therefore, because pooling is costly, the principal will never pool if she gains nothing from it. This happens precisely when the principal receives most of the residual surplus. To see this, suppose that $\gamma < \gamma^*$, and so $q^p > q^* > q^a$ by Proposition 5. If $q_1 < q^a < q^p$, both parties want the agent to quit, so there is no conflict of interest. The principal can fire the agent, or, equivalently, simply reveal q_1 to the agent and he will quit on his own (since $q_1 < q^a$). If $q_1 > q^p$, the principal wants to retain the agent, and since $q^p > q^a$, the agent also wants to stay, so again there is no conflict of interest. By the Jensen inequality argument above, pooling is costly and yields no additional benefits to the principal. So the principal will reveal q_1 to the agent and he will choose to stay. If output lies at either extreme ($q_1 < q^a < q^p$ or $q_1 > q^p > q^a$), the interests of both parties are aligned.

Now if $q_1 \in (q^a, q^p)$, the principal wishes to fire the agent while the agent wishes to stay. This represents a conflict of interest. In at-will employment contracts, both parties are free to leave at anytime, and so output must clear

$\max(q^a, q^p)$ to justify continuation. If $\gamma < \gamma^*$, then $q^p = \max(q^a, q^p)$ is the relevant hurdle. Output fails this hurdle if $q^a < q_1 < q^p$, so the principal can implement her optimal termination decision by firing the agent. Note that simply revealing q_1 is not sufficient (as it was earlier) because of the conflict of interest; the principal must fire the agent. Thus the principal fires if $q_1 < q^p$ and fully reveals output otherwise. Her payoffs are given by $\max(\pi(q_1), \bar{u_2}^p)$ as shown in Figure 1.2(a). In sum, if the principal receives most of the residual surplus, she sets a higher output target than the agent, and hence can implement her optimal termination rule without resorting to pooling.

Life is different if the tables are turned. Now suppose that $\gamma > \gamma^*$, so the agent receives most of the residual surplus. By Proposition 5, $q^p < q^* < q^a$. As before, there is alignment of interest if output is extreme but conflict of interest otherwise. If output is very low ($q_1 < q^p < q^a$), both parties prefer separation. The principal fires the agent, or equivalently reveals q_1 and the agent leaves on his own. If output is very high ($q_1 > q^a > q^p$), both parties prefer continuation. The principal reveals q_1 and the agent chooses to stay. By Jensen's inequality, pooling is costly and yields no benefits to the principal.

If $q_1 \in (q^p, q^a)$ the agent prefers to quit but the principal prefers him to stay. If the principal reveals q_1, the agent will quit, leading to a suboptimal outcome for the principal. The key insight is that the principal can retain the worker by withholding information. More precisely, the principal will pool over a large enough region such that the average output level within the pool $\hat{q}$ exceeds q^a. This ensures that the agent will stay on the job if he is told that he is average. Above this pooling interval, the principal will separate because pooling is costly (by Jensen's inequality). The principal pools as little as possible, only enough to keep workers on the job.

To see this visually, refer to Figure 4. The graph on the left plots the principal's profit if she reveals output: she earns $\bar{u_2}^p$ if $q_1 < q^a$ and $\pi(q_1)$ if $q_1 > q^a$. Note that the principal cannot collect $\pi(q_1)$ if $q_1 \in (q^p, q^a)$ because the agent leaves if he learns q_1; instead the principal only collects her outside option $\bar{u_2}^p$. Therefore full revelation implements the agent's termination rule, as $q^a = \max(q^p, q^a)$ is the relevant hurdle to clear. The graph on the right plots profit under pooling. $Q = (\underline{q}, \bar{q})$ is the pooling region and $\hat{q} = E[q_1 | q_1 \in Q]$. The straight dashed line is the principal's profit over the pool, given by the linear equation $\hat{\pi}(q_1) = (V - b) q_1 e_2(\hat{q})$. So the principal's overall profit is $\hat{\pi}(q_1)$ over Q and $\max(\pi(q_1), \bar{u_2}^p)$ outside of Q.

Of course, pooling is still costly, but the additional benefits from reten-

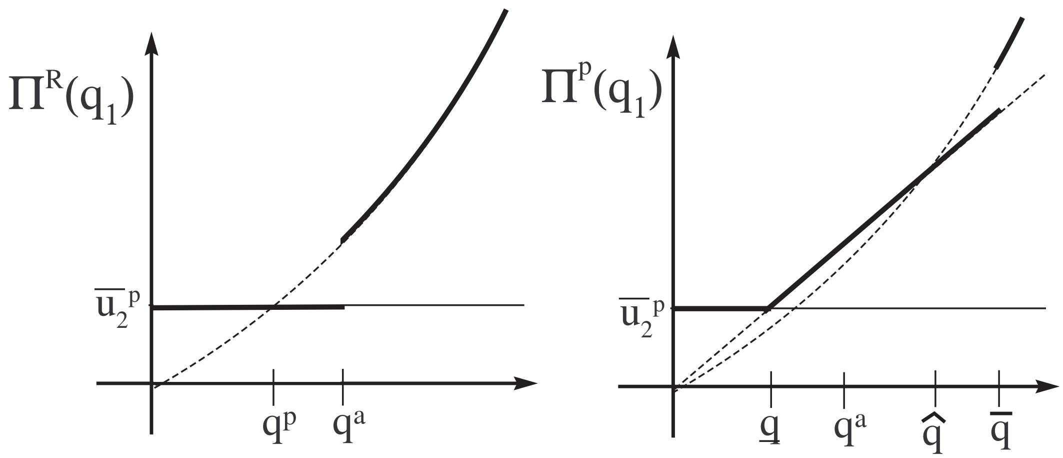

Figure 4: Principal's profit under Revelation (a) and Pooling (b)

tion outweigh these costs. The principal will pool if the area under $\Pi^P(q_1)$ exceeds the area under $\Pi^R(q_1)$. The next figure shows this difference, illustrating the cost and benefit of pooling. As before, the cost is the value of lost output from the stars ($q_1 > \hat{q}$) who slack off, and the benefit is the value of additional output from the slugs ($q_1 > \hat{q}$) who work harder. But the benefit of pooling now expands to include the value of output from slugs who would have quit if they knew their q_1. This Profit from Retention is the lightly shaded triangle lying beneath the dashed line for $q_1 \in (\underline{q}, q^a)$.

Under certain conditions, this profit is large enough such that the benefits of pooling outweigh the costs. So the value of retention overturns the costs of pooling generated by the convex profit function and Jensen's inequality. In sum, the principal fires the worker if output is sufficiently low, reveals output if it is sufficiently high, and pools if output is in between. Collecting these results, we arrive at the main theorem. The full proof and construction of equilibrium lies in the appendix.

Theorem 1 *If $\gamma < \gamma^*$, the principal fully reveals output. If $\gamma > \gamma^*$ and costs are sufficiently convex, there exists a nontrivial pooling interval $Q \equiv (\underline{q}, \bar{q})$ such that the principal pools over Q and reveals output outside of Q.*

The convexity of the cost function is necessary in order to make it worthwhile for the principal to pool. Precisely, as the costs of effort become more convex, this flattens the principal's profit function: the principal makes less

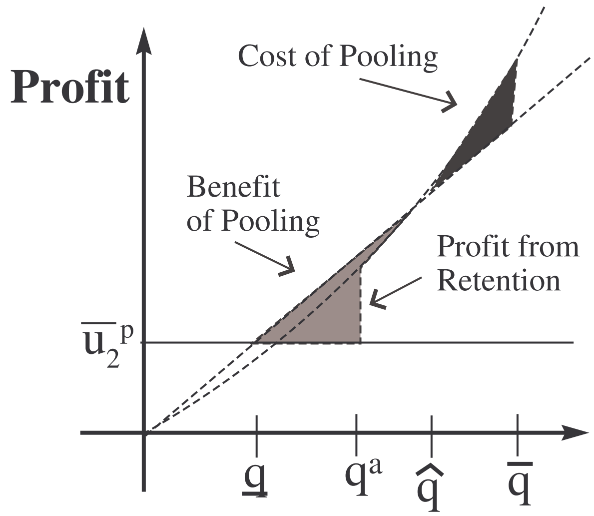

Figure 5: Costs and Benefits of Pooling.

money off of revealing high output because labor is so expensive. This crushes down the dark shaded region in the figure: the costs of pooling shrinks since the value of lost output from the stars falls. All the while, the principal still earns profit from retention. Eventually, costs of pooling sinks below the benefit, and hence it is worthwhile for the principal to pool. In the limit, the two dashed lines in the figure converge, at which point the principal's profit function is linear and the principal pools everybody. But for any finite amount of convexity in the cost function, the principal will pool up to some finite $\bar{q}$. Hence for a sufficiently convex cost function there exists a nontrivial pooling equilibrium. For the full argument, see the proof in the appendix.

To summarize, if the principal receives most of the residual surplus, she will fully reveal output to the agent. If instead the agent receives most of the residual surplus, the principal will partially reveal output to the agent. Precisely, the principal will fire the agent (or simply reveal his output) if output is sufficiently low, fully reveal output if it is sufficiently high, and withhold information (pool) if it is in between. The principal withholds information as a retention mechanism.

4.1 Comparative Statics

Comparative statics on the model are straightforward and follow from the proof of the theorem.

Corollary 1 *The principal pools more as the agent's labor supply becomes less elastic.*

As the agent's costs become more convex, it becomes more expensive for the principal to induce any given amount of effort. The principal's profit function flattens out, so the principal pools over a larger region. The key is that the agent's effort function is increasing in q_1, which the principal reveals. Workers with high early output will work hard in the second stage. If their labor is inelastic, this effort is expensive for the principal to induce. So it is costly to reveal performance information to workers with inelastic labor supply. Thus the principal withholds performance information more by expanding the pool.

Corollary 2 *The principal pools more as the agent's outside options increase.*

The agent's target q^a increases with $\bar{u_2}^a$. The pooling region Q must satisfy $E[q_1|Q] \geq q^a$, so the principal expands the pooling region. In words, as the agent's alternatives improve, he sets a higher standard to justify staying in the game and foregoing good options elsewhere. But to stay on the job he must believe that his average output exceeds his quit rate. So the principal expands the pool (at the top) so that the average output over the pool matches the increase in q^a. This implies that jobs or industries where agents have high outside options should exhibit more pooling in employee performance ratings. Employers withhold information to prevent workers with good alternatives from leaving.

A study on performance evaluations in the Navy (Bjerke et al 1987) is consistent with this theory. The authors find that higher ranked combat communities in the Navy (nuclear submarine community) had more compression and inflation of ratings than in lower ranked support communities (supply and tanker fleets). If a soldier's outside opportunities increase with rank, then this fits the story here.

5 Conclusion

In surveying the extensive literature on subjective performance evaluation, Prendergast (1999) finds that "supervisors distort subjective performance ratings by not sufficiently differentiating good from bad performance." There is sufficient documentation of this phenomenon indicating that this is a common feature of performance evaluation systems in firms. While previous explanations for this have been offered (based on psychology), the literature lacks a convincing economic reason to pool employee performance.

This paper advances one: retention. When workers are tempted to quit firms because of weak performance inside and strong options outside, firms pool performance together to keep them on the job. Workers do not know where in the pool they sit, and so both the stars (at the top of the pool) and the slugs (at the bottom) believe that they are average. By complementarity in production, second stage effort increases in (the agent's belief of) first stage output. So the stars think they're average and work less, while the slugs think they're average and work more. But the profit gain from retaining slugs who would have otherwise quit exceeds the profit loss from the stars, so it pays for the firm to pool workers together. The pool is large enough that the average member of the pool stays on the job. Since the pooled workers think that they're average, this guarantees retention.

6 Appendix

Proof of Proposition 1

Proof: Let $X \equiv \{q_1 : S(q_1) > \bar{u_2}\}$ be the social planner's continuation set: it is efficient to allow a worker with output q_1 to advance to the second stage if and only if $q_1 \in X$. Let $P_1 = Pr(X)$.

Let $Q = \bigcup_{i \in I} Q_i$ be an arbitrary pooling region. Let $B \equiv \{f : \mathbb{R} \to \mathbb{R}\}$ be the space of all real-valued functions. Let A be the space of all functions such that are constant over Q_i for each i . So

$$A = \{e_2 : \mathbb{R} \to \mathbb{R} | e_2(q_1) = e_2(E[q_1 | q_1 \in Q_i]) \texttt{ if } q_1 \in Q_i \texttt{ for some } i \in I\}$$

If the planner pools, he solves

$$\max_{e_1, e_2(\cdot) \in A} \int_X S(q_1) g(q_1 - e_1) dq_1 + (1 - P_1)\bar{u_2} - C(e_1)$$

If he does not pool he solves

$$\max_{e_1, e_2(\cdot) \in B} \int_X S(q_1) g(q_1 - e_1) dq_1 + (1 - P_1)\bar{u_2} - C(e_1)$$

since $A \subset B$, the planner achieves a (weakly) higher maximum by not pooling. Since the Q was chosen arbitrarily, the planner is never better off pooling. ■

Proof of Proposition 2

Proof: Let $X \equiv \{q_1 : S(q_1) > \bar{u_2}\}$ be the social planner's continuation set: it is efficient to allow a worker with output q_1 to advance to the second stage if and only if $q_1 \in X$. Let $P_1 = Pr(X)$. I will show that X is in fact an interval.

Continuation surplus is

$$S(q_1) = \mathbb{E}V(q_1, e_2(q_1) + \epsilon_2) - C(e_2(q_1))$$

The social planner solves

$$\max_{e_1, e_2(\cdot)} \int_X S(q_1) g(q_1 - e_1) dq_1 + (1 - P_1)\bar{u_2} - C(e_1)$$

Note that by Proposition 1, it is never efficient to pool, so the effort function $e_2(\cdot)$ is unconstrained.

Maximizing the integral pointwise yields

$$e_2(q_1) \in \arg\max S(q_1) g(q_1 - e_1)$$

for almost every $q_1 \in X$. For every such q_1, $e_2(q_1)$ solves

$$\frac{\partial S(q_1)}{\partial e_2(q_1)} = 0.$$

Expanding this gives

$$C'(e_2(q_1)) = \mathbb{E}V_2(q_1, e_2(q_1) + \epsilon_2)$$

Hence

$$S'(q_1) = \mathbb{E}V_1(q_1, e_2(q_1) + \epsilon_2) + e_2'(q_1)[\mathbb{E}V_2(q_1, e_2(q_1) + \epsilon_2) - C'(e_2(q_1))].$$

The term in brackets is zero by the FOC for $e_2(q_1)$. Hence $S'(q_1) > 0$ since $V_1 > 0$. Hence the ex post social value of a worker increases in his first stage output. Since $S(\cdot)$ is continuous, by the intermediate value theorem, there exists a q^* such that $S(q^*) = \bar{u_2}$. Since $S(\cdot)$ is strictly increasing, this means that $X = \{q_1 : q_1 > q^*\}$.

Let $\hat{S} = S(q_1)g(q_1 - e_1)$. Then

$$\frac{\partial \hat{S}}{\partial e_2 \partial q_1} = g(q_1 - e_1)\mathbb{E}V_{21}(q_1, e_2(q_1) + \epsilon_2).$$

So $e_2'(q_1) > 0$ iff $V_{21} > 0$. ■

Lemma 1 *Suppose* $C(x), C'(x), C''(x), C'''(x) > 0$ *for* $x > 0$ *and* $C(0) = C'(0) = 0$. *Then*

$$\frac{C'''(x)}{C''(x)} = \lambda \frac{C''(x)}{C'(x)} \quad \texttt{for some } x > 0 \Longrightarrow \lambda < 1$$

Proof: Let $x > 0$ satisfy the hypothesis. Let $f = C'$. Then

$$\frac{C'''(x)}{C''(x)} = \lambda \frac{C''(x)}{C'(x)} \Longleftrightarrow \frac{d}{dx}(\frac{f(x)}{f'(x)}) = 1 - \lambda$$

Integrating both sides of the equality on the right gives

$$\frac{f(x)}{f'(x)} = ((1-\lambda)x + k$$

where k is a constant. Therefore

$$\frac{d}{dx}(\log\, f(x)) = \frac{f'(x)}{f(x)} = \frac{1}{((1-\lambda)x + k}$$

Integrating again gives

$$\log\, f(x) = \int \frac{1}{((1-\lambda)x + k} dx = \log\, ((1-\lambda)x + k)^{\frac{1}{1-\lambda}} + c$$

where c is a constant. Apply the exponential function to both sides:

$$f(x) = [\exp(c)][((1-\lambda)x + k)^{\frac{1}{1-\lambda}}]$$

Now $f(0) = C'(0) = 0$, so this means $k = 0$. Since costs are strictly increasing, $f(x) > 0$ for $x > 0$. If $\lambda = 1$, then $f = 0$, a contradiction. If $\lambda > 1$, then simple algebra shows that $f(x) < 0$, another contradiction. Hence it must be that $\lambda < 1$. ■

Proof of Proposition 3
Proof:

Step One: Effort on $\mathbb{R}\backslash Q$.

Outside of Q the agent solves

$$\max_{q^a, e_1, e_2(\cdot)} \int u(q_1) g(q_1 - e_1) dq_1 + (1 - P_1)\bar{u_2}^a - C(e_1)$$

whose FOC gives $C'(e_2(q_1)) = bq_1$ for all $q_1 \notin Q$. By the implicit function theorem,

$$e_2'(q_1) = \frac{b}{C''(e_2(q_1))} > 0$$

$$e_2''(q_1) = -C'''(e_2(q_1)))(\frac{be_2'(q_1)}{C''(e_2(q_1)))})^2 < 0$$

So the agent's effort function is increasing and concave off of Q.

Step Two: Effort on Q.

From (EC) the agent's effort function must be constant over each Q_i, so he earns $\hat{u}(q_1) = bq_1\hat{e}_i - C(\hat{e}_i)$ over Q_i, where $\hat{e}_i$ is the single effort choice over Q_i. The agent solves

$$\max_{\hat{e}_i} E[\hat{u}(q_1)|q_1 \in Q_i]$$

FOCs show that $C'(\hat{e}_i) = b\hat{q}_i$ where $\hat{q}_i = E[q_1|q_1 \in Q]$.

Now the effort function is defined only off of Q by $C'(e_2(q_1)) = bq_1$ for all $q_1 \notin Q$. Extend the effort function over $Q = \bigcup_{i \in I} Q_i$ in the following way. Construct the function $f : \mathbb{R} \to \mathbb{R}$ by $C'(f(q_1)) = bq_1$. So f projects the effort function over $\mathbb{R}$. Let $e_2(\hat{q}_i) \equiv f(\hat{q}_i)$, defining the effort function at $\hat{q}_i$. So

$$C'(e_2(\hat{q}_i)) = C'(f(\hat{q}_i)) = b\hat{q}_i = C'(\hat{e}_i)$$

And therefore let

$$e_2(q_1) = \hat{e}_i = e_2(\hat{q}_i) \texttt{ for all } q_1 \in Q_i.$$

This defines the effort function over Q. Combined with Step One, the function is defined over $\mathbb{R}$.

Step Three: Payoffs.

Recall that $u(q_1) = bq_1e_2(q_1) - C(e_2(q_1))$ and $\pi(q_1) = (V - b)q_1e_2(q_1)$. Now

$$u'(q_1) = e_2'(q_1)[bq_1 - C'(e_2(q_1))] + be_2(q_1) = be_2(q_1) > 0$$

$$u''(q_1) = be_2'(q_1) > 0$$

using the FOC $C'(e_2(q_1)) = bq_1$. So the agent's continuation utility is increasing and convex. Moreover, $\pi'(q_1) = q_1 e_2'(q_1) + e_2(q_1) > 0$ so π is increasing. Some computations show that

$$\pi''(q_1) > 0 \Leftrightarrow \frac{C'''(e_2(q_1)))C'(e_2(q_1)))}{(C''(e_2(q_1))))^2} < 2.$$

Let $q_1 \in \mathbb{R}$, and let $x = e_2(q_1) > 0$. By Lemma 1,

$$\frac{C'''(x)C'(x)}{(C''(x))^2} = \lambda < 1 < 2$$

And so $\pi''(q_1) > 0$. Since q_1 is arbitrary, this means π is convex. ■

Proof of Proposition 4

Proof: Suppose the principal pools over some $Q_i = (x, y)$. Construct a sequence of sets F_n such that

$$F_0 \equiv Q_i = (x, y)$$

$$\hat{q_n} = E[q_1 | q_1 \in F_n]$$

$$F_{n+1} = (x, \hat{q_n})$$

Lemma 2 *If the principal pools over F_n, then without commitment he pools over F_{n+1} and separates over $F_n \backslash F_{n+1}$.*

Proof: (By induction)

Step One. Her profit over F_0 is the linear function $\hat{\pi}_0(q_1) = (V - b)q_1 e_2(\hat{q_0})$ for all $q_1 \in F_0$. Since $e_2(\cdot)$ is increasing,

$$\pi(q_1) = (V - b)q_1 e_2(q_1) > (V - b)q_1 e_2(\hat{q_0}) = \hat{\pi}(q_1)$$

for all $q_1 > \hat{q_0}$. So the principal prefers to reveal q_1 for all $q_1 \in (\hat{q_0}, y) = F_0 \backslash F_1$ and pool for all $q_1 \in (x, \hat{q_0}) = F_1$.

Step n. Her profit over F_n is the linear function $\hat{\pi}_n(q_1) = (V - b)q_1 e_2(\hat{q_n})$ for all $q_1 \in F_n$. Since $e_2(\cdot)$ is increasing,

$$\pi(q_1) = (V - b)q_1 e_2(q_1) > (V - b)q_1 e_2(\hat{q_n}) = \hat{\pi}(q_1)$$

for all $q_1 > \hat{q_n}$. So the principal prefers to reveal q_1 for all $q_1 \in (\hat{q_n}, y) = F_n \backslash F_{n+1}$ and pool for all $q_1 \in (x, \hat{q_n}) = F_{n+1}$. ■

From the lemma, without committment the principal pools over F_{n+1} for each n. But $F_{n+1} \subset F_n$ and $\hat{q_n} \longrightarrow x$ so $F_n \longrightarrow \emptyset$. So the principal does not pool over any interval, and hence not on any finite disjoint unions of intervals. So any pooling region Q unravels. ■

Proof of Proposition 5

Proof:

Continuation payoffs are $u(q_1) = bq_1e_2(q_1) - C(e_2(q_1))$ and $\pi(q_1) = (V-b)q_1e_2(q_1)$. Recall from Proposition 3 that $C'(e_2(q_1)) = bq_1$ for all $q_1 \notin Q$ so $e_2(q_1)$ depends implicitly on b. Let

$$U(q_1, b) = bq_1e_2(q_1, b) - C(e_2(q_1, b))$$

be the continuation utility for each b. Observe that

$$U(q^a) = u(q^a) = \bar{u_2}^a$$
$$U(q_1, V) = Vq_1e_2(q_1, V) - C(e_2(q_1, V)) = S(q_1)$$
$$U(q^*, V) = S(q^*) = \bar{u_2}$$

where $S(q_1)$ is the continuation surplus function. The equalities $u(q^a) = \bar{u_2}^a$ and $S(q^*) = \bar{u_2}$ follow from the FOCs of the agent's and planner's problem, respectively. Now U increases in both its arguments:

$$\frac{\partial U(q_1, b)}{\partial q_1} = u'(q_1) = be_2(q_1) > 0$$
$$\frac{\partial U(q_1, b)}{\partial b} = q_1e_2(q_1, b) > 0 \texttt{ for each } q_1.$$

So $U(q^a, V) > U(q^a, b)$ since $b < V$. Let

$$\gamma^* \equiv \frac{U(q^a, b)}{U(q^a, V)} < 1$$

And thus

$$\gamma U(q^a, V) > U(q^a, b) \iff \gamma > \gamma^*. \tag{1}$$

By definitions of γ, q^*, q^a respectively, $\gamma\bar{u_2} = \bar{u_2}^a$, $U(q^*, V) = \bar{u_2}$ and $U(q^a, b) = \bar{u_2}^a$. Combining these gives

$$q^a > q^* \iff U(q^a, V) > U(q^*, V) = \bar{u_2} = \bar{u_2}^a/\gamma = U(q^a, b)/\gamma.$$

Combining with (1) gives

$$q^a > q^* \Longleftrightarrow \gamma > \gamma^*. \quad (2)$$

Some algebra combined with manipulating definitions gives:

$$\begin{aligned}
q^a > q^* &\overset{(1)}{\Longleftrightarrow} U(q^a, V) > U(q^*, V) \\
&\overset{(2)}{\Longleftrightarrow} Vq^a e_2(q^a) - C(e_2(q^a)) = U(q^a, V) > U(q^*, V) = \bar{u_2} = \bar{u_2}^a + \bar{u_2}^p \\
&\overset{(3)}{\Longleftrightarrow} Vq^a e_2(q^a) - C(e_2(q^a)) > U(q^a, b) + \bar{u_2}^p = bq^a e_2(q^a) - C(e_2(q^a)) + \bar{u_2}^p \\
&\overset{(4)}{\Longleftrightarrow} \pi(q^a) = (V - b)q^a e_2(q^a) > \bar{u_2}^p = \pi(q^p) \\
&\overset{(5)}{\Longleftrightarrow} q^a > q^p
\end{aligned}$$

Since (1) U is increasing, (2) definition of $\bar{u_2}$, (3) $U(q^a, b) = \bar{u_2^a}$, (4) subtract $bq^a e_2(q^a)$ from both sides and add $C(e_2(q^a))$ to both sides, and (5) π is increasing. Furthermore,

$$\begin{aligned}
q^a > q^p &\overset{(6)}{\Longleftrightarrow} u(q^a) > u(q^p) \\
&\overset{(7)}{\Longleftrightarrow} \bar{u_2}^a = u(q^a) > u(q^p) = bq^p e_2(q^p) - C(e_2(q^p)) \\
&\overset{(8)}{\Longleftrightarrow} \bar{u_2}^a + Vq^p e_2(q^p) - bq^p e_2(q^p) > Vq^p e_2(q^p) - C(e_2(q^p)) \\
&\overset{(9)}{\Longleftrightarrow} \bar{u_2} = \bar{u_2}^a + \bar{u_2}^p = \bar{u_2}^a + \pi(q^p) > Vq^p e_2(q^p) - C(e_2(q^p)) \\
&\overset{(10)}{\Longleftrightarrow} U(q^*, V) = \bar{u_2} > Vq^p e_2(q^p) - C(e_2(q^p)) = U(q^p, V) \\
&\overset{(11)}{\Longleftrightarrow} q^* > q^p
\end{aligned}$$

Because: (6) u is increasing, (7) definition of u, (8) add $Vq^p e_2(q^p)$ to both sides, (9) $\pi(q^p) = Vq^p e_2(q^p) - bq^p e_2(q^p)$, (10) definition of U, and (11) U is increasing.

Combining all this with (2) gives

$$q^a > q^* > q^p \Longleftrightarrow \gamma > \gamma^*. \quad (3)$$

■

Proof of Theorem 1

Proof: As before, the agent's effort function $e_2(\cdot)$ is defined by $C'(e_2(q_1)) = bq_1$ for all $q_1 \notin Q$. Recall that

$$u(q_1) = bq_1 e_2(q_1) - C(e_2(q_1))$$

$$\pi(q_1) = (V-b)q_1e_2(q_1).$$

Let $Q \subset \mathbb{R}$ be a candidate pooling itnerval. Let $\hat{q} = E[q_1|q_1 \in Q]$ be the average output over Q. As in the proof of Proposition 4, the agent's and principal's payoffs over the pool are

$$\hat{u}(q_1) = bq_1e_2(\hat{q}) - C(e_2(\hat{q}))$$

$$\hat{\pi}(q_1) = (V-b)q_1e_2(\hat{q})$$

for all $q_1 \in Q$. These functions are linear over Q. To guarantee participation, it must be that both parties prefer the pool over quitting. That is,

$$E[\hat{u}(q_1)|Q] = \hat{u}(\hat{q}) = u(\hat{q}) > \bar{u_2}^a = u(q^a) \Leftrightarrow \hat{q} > q^a$$
$$E[\hat{\pi}(q_1)|Q] = \hat{\pi}(\hat{q}) = \pi(\hat{q}) > \bar{u_2}^p = \pi(q^p) \Leftrightarrow \hat{q} > q^p$$

since the payoff functions are increasing. Thus participation holds iff $\hat{q} > \max(q^a, q^p)$.

If the principal reveals over Q, payoffs are

$$\Pi^R(q_1) = \begin{cases} \pi(q_1) & \text{if } q_1 > \max(q^a, q^p) \\ \bar{u_2}^p & \text{otherwise} \end{cases} \qquad U(q_1) = \begin{cases} u(q_1) & \text{if } q_1 > \max(q^a, q^p) \\ \bar{u_2}^a & \text{otherwise} \end{cases}$$

Since π is convex over Q, Jensen's inequality shows that

$$E[\pi(q_1)|Q] > \pi(E[q_1|Q]) = \pi(\hat{q}) = \hat{\pi}(\hat{q}) = E[\hat{\pi}(q_1)|Q].$$

The last equality holds since $\hat{\pi}$ is linear. Multiply left and right hand sides by $1/\Pr(Q)$:

$$\int_Q \pi(q_1)g(q_1-e_1)dq_1 > \int_Q \hat{\pi}(q_1)g(q_1-e_1)dq_1.$$

Part One. Let $\gamma < \gamma^*$. By Proposition 5, $q^a < q^* < q^p$, so $q^p = \max(q^a, q^p)$ and $\Pi^R(q_1) = \max(\pi(q_1), \bar{u_2}^p)$. Thus

$$\int_Q \Pi^R(q_1)g(q_1-e_1)dq_1 > \int_Q \pi(q_1)g(q_1-e_1)dq_1 > \int_Q \hat{\pi}(q_1)g(q_1-e_1)dq_1.$$

and so the principal prefers to separate, so Q cannot be a pooling interval. Hence the principal will never pool over any interval, and hence never over any finite union of disjoint intervals. This proves part one.

Part Two. Now let $\gamma > \gamma^*$. By Proposition 5, $q^p < q^* < q^a$, so $q^a = \max(q^a, q^p)$. So

$$\Pi^R(q_1) = \begin{cases} \pi(q_1) & \text{if } q_1 > q^a \\ \bar{u_2}^p & \text{otherwise} \end{cases} \qquad U(q_1) = \begin{cases} u(q_1) & \text{if } q_1 > q^a \\ \bar{u_2}^a & \text{otherwise} \end{cases}$$

The equilibrium is by construction. Let Q be an arbitrary pooling region. Without loss, let Q be a finite unions of disjoint intervals, since the set of such intervals is dense in $\mathbb{R}$. Write $Q = \bigcup_{i \in I} Q_i$ for a finite set I. Suppose $q_1 > q^a$ for all $q_1 \in Q_i$ for some k, i.e. suppose there exists an interval Q_i that lies above q^a. By convexity of π and Jensen's inequality, the same computations from Part One shows that the principal will separate over Q_i. So Q will not consist of an interval whose elements all exceed q^a. Now suppose $E[q_1|q_1 \in Q_i] < q_a$ for some i. That is, there exists an interval whose average output over the interval falls below q^a. Then the agent will leave, and pooling is pointless. So every Q_i must satisfy $E[q_1|q_1 \in Q_i] > q^a$ and $q^a \in Q_i$. Every interval must contain q^a and the average of the interval must exceed q^a. Since each Q_i is disjoint, this means that Q is a single interval. Write $Q = (\underline{q}, \bar{q})$ such that $\hat{q} = E[q_1|q_1 \in Q] > q^a > q^p$ and $q^a \in Q$.

The principal's payoffs under pooling are

$$\Pi^P(q_1) = \begin{cases} \bar{u_2}^p & \text{if } q_1 < \underline{q} \\ \hat{\pi}(q_1) & \text{if } q_1 \in (\underline{q}, \bar{q}) \\ \pi(q_1) & \text{if } q_1 > \bar{q} \end{cases}$$

Figure 4 shows Π^R and Π^P. In choosing the optimal Q the principal solves

$$\max_{\underline{q}, \bar{q}} \int^{\underline{q}} \bar{u_2}^p + \int_{\underline{q}}^{\bar{q}} \hat{\pi}(q_1) + \int_{\bar{q}} \pi(q_1)$$

where I suppress the integrand $g(q_1 - e_1)dq_1$ and the integration limits ($-\infty, \infty$) both here and throughout the rest of the proof. Let $(\underline{q}, \bar{q})$ solve this problem. Solving for $(\underline{q}, \bar{q})$ explicitly is messy and uninstructive; it is enough to show that the principal is better off pooling instead of separating. The principal pools if $\int \Pi^P(q_1) > \int \Pi^R(q_1)$. Now

$$\int \Pi^P(q_1) - \int \Pi^R(q_1) = \int_{\underline{q}}^{\bar{q}} \hat{\pi}(q_1) - [\int_{\underline{q}}^{q^a} \bar{u_2}^p + \int_{q^a}^{\bar{q}} \pi(q_1)] = B - C$$

where

$$B = \int_{\underline{q}}^{q^a} (\hat{\pi}(q_1) - \bar{u_2}^p) + \int_{q^a}^{\hat{q}} \hat{\pi}(q_1) - \pi(q_1)$$

$$C = \int_{\hat{q}}^{\bar{q}} \pi(q_1) - \hat{\pi}(q_1)$$

are the (B)enefits and (C)osts of pooling, respectively. These are the light and dark regions of Figure 4. By Lemma 3, as costs become convex (in a precise sense), $\hat{\pi}(q_1) - \pi(q_1)$ converges pointwise to zero. So the principal's payoffs under pooling and separation converge, and by the Lebesgue Convergence Theorem, so do their integrals. Thus $C \longrightarrow 0$ and

$$B \longrightarrow \int_{\underline{q}}^{q^a} [(V - b)(q_1) - \bar{u_2}^p] > 0$$

since $\hat{\pi}(q_1) \longrightarrow (V - b)q_1$ by Lemma 3. So if costs are sufficiently convex, the principal prefers to pool. Q is a nontrivial pooling equilibrium. ∎

Lemma 3 *If costs become increasingly convex,* $\hat{\pi}(q_1) - \pi(q_1) \longrightarrow 0$ *pointwise. Moreover, they both converge to the linear function* $(V - b)q_1$.

Proof: Let Q be a pooling interval and $\hat{q} = E[q_1 | q_1 \in Q]$. Parameterize the cost function by $C(e_t) = \frac{c}{n+1} e_t^{n+1}$ for each n. So $C'(e_t) = ce_t^n$, and $e_2(q_1) = \sqrt[n]{(b/c)q_1}$. For all $q_1 \in Q$,

$$\pi(q_1) - \hat{\pi}(q_1) = (V - b)\sqrt[n]{\frac{b}{c}} q_1(\sqrt[n]{q_1} - \sqrt[n]{\hat{q}}) \longrightarrow 0 \quad \texttt{as} \quad n \longrightarrow \infty$$

$$\hat{\pi}(q_1) = (V - b)\sqrt[n]{\frac{b}{c}} q_1(\sqrt[n]{\hat{q}}) \longrightarrow (V - b)q_1 \quad \texttt{as} \quad n \longrightarrow \infty$$

∎

References

[1] Bjerke, D.G., J.N. Cleveland, R. Morrison, W. Wilson (1987): "Officer Fitness Report Evaluation Study," Navy Personnel Research and Development Center Report, San Diego, CA.

[2] Burton, Diane M (1998) "The Firmwide 360-degree Performance Evaluation Process at Morgan Stanley." HBS Case 9-498-053.

[3] Burton, Diane M (2000) "Morgan Stanley: Becoming a 'One-Firm Firm'." HBS Case 9-400-043.

[4] Fama, Eugene (1980): "Agency Problems and the Theory of the Firm," Journal of Political Economy, Vol. 88, No. 2, pp. 288-307.

[5] Landy, Frank and J. Farr (1980): "Performance Rating," Psych. Bul., 87, pp. 72107.

[6] Lazear, Edward P. (2001): "The Peter Principle: A Theory of Decline," Mimeo.

[7] Lazear, Edward P. and Sherwin Rosen (1981): "Rank order Tournaments as Optimum Labor Contracts," *Journal of Political Economy*, 89, 841-864.

[8] Lizzeri, Alessandro, Margaret Meyer, and Nicola Persico (2002): "The Incentive Effects of Interim Performance Evaluations," Mimeo.

[9] Mohrman, A. and Edward Lawler (1991): "Pay for Performance: Evaluating Performance Appraisal and Merit Pay." Washington: National Academy Press.

[10] Murphy, Kevin J. (1990): "Merck & Co., Inc.". HBS Case No. 9-491-006, Sep 12, 1990.

[11] Murphy, Kevin R. and Jeannette Cleveland 1991: "Performance Appraisal: An Organizational Perspective." Boston: Allyn and Bacon.

[12] Prendergast, Canice "The Provision of Incentives in Firms." Journal of Economic Literature, Vol. 37, No. 1. (Mar., 1999), pp. 7-63.