The Boss's Career Concerns as a Reason for Yes-Men

24 March 2008

Eric Rasmusen

Abstract

Why would an organization be staffed by "yes-men" who always agree with their boss? One reason is that the boss may not welcome disagreement. Once he makes a decision, changing it shows lack of confidence in his ability to observers. This career concern may outweigh his desire to make the best decision for the organization, and lead to his agents concealing information that his initial decision was wrong.

Dan R. and Catherine M. Dalton Professor, Department of Business Economics and Public Policy, Kelley School of Business, Indiana University. visitor (07/08), Nuffield College, Oxford University. Office: 011-44-1865 554-163 or (01865) 554-163. Nuffield College, Room C3, New Road, Oxford, England, OX1 1NF. Erasmuse@indiana.edu. http://www.rasmusen.org. This paper: http://www.rasmusen.org/papers/career-rasmusen.pdf.

Keywords: Yes-men, conservatism, career concerns

I thank Nuffield College, Oxford University for its hospitality.

Introduction

The “yes-man” is a familiar figure of fun in the world of business. One cartoon that depicts a group of executives around a table with their boss goes something like this: “Yes” says one, “Yes”says the next, and the next, and the next and the next. But their thoughts, given in cloud bubbles in the cartoon, are: “No way!”, “We’ll go bankrupt”, “How can he think that?”, “Oh my Lord!”, “Please, somebody say No”.

Why do the subordinates say “yes” when they think “no”? The explanation I will model is that the boss would like them to say yes. Because of his own career concerns, he is not going to show that he made a mistake by reversing course, so the new information is worse than useless, since it will not affect profits but will show that his own signal may have been wrong.

The leading article on yes-men, Prendergast (1993), explains conformism differently, as the result of the incentives the boss gives his agents to exert effort to collect information. Suppose the boss tells an agent to collect information, and the agent wants to avoid the effort. The boss will have to reward the agent based on how closely the agent’s conclusions match the boss’s. The boss does not want the agent to collect information the boss already knows, but he knows that on average any new information will be correlated with the boss’s own information, so he will reward confirmation. If the employee has a noisy signal of the boss’s initial opinion, however, the employee will shade his report in the direction of that opinion.

The Prendergast model is about moral hazard, the problem of agent effort, but it could be adapted to adverse selection, the problem of agent ability. The boss would like to reward a more talented agent, and good information would indicate high ability. The boss’s measure of whether the agent’s information was good would be how closely it matches his own information, so if the agent gets a signal of the boss’s information, he will report that instead of what he truly believes. In this model, as in Prendergast’s moral hazard model, the agent’s information may be much more accurate than the boss’s, but he will still conform, because the boss must use his own signal to judge the agent’s.

There exists a plethora of other models of conformity in general, as opposed to conformity to a boss’s opinion, and these could also explain the yes man, e.g. Scharfstein, & Stein (1990), Avery & Chevalier (1999), Harrison Hong, Jeffrey D. Kubik & Amit Solomon (2000). Most papers in this literature these rely on one of the two explanations just described– the conformer's desire to avoid the effort of forming his own opinion, or his desire to avoid signalling low ability. The adverse selection models are part of the "career concerns" literature originated by Holmstrom (1982, 1999), in which the agent's actions are driven by a desire for a good reputation as much as by his present compensation contract. This can lead to herding behavior, as agents imitate each other– or their boss– to try to jam the signal provided by behavior based on their individual signals. Thus, career concerns provide a reason for "yes men" in firms. Subordinates agree with the boss's opinion because he is more talented or better informed and they do not want to show their own judgement to be inferior. These models can also explain a agent's *non*comformity, as an attempt to signal that he has better information than his boss. In Prendergast & Stole (1996) a young manager is willing to react to new information by reversing his previous decision, but an old manager is not, since one good new decision will not outweigh many old bad decisions in the view of the outside world.

These explanations of yes men are based on the idea that the yes man is trying to impress the boss, to convey favorable or suppress unfavorable information about himself. The present paper looks at a different explanation: the agent is acting on behalf of the manager's career concerns, not his own. Rather than saying "yes" to protect his own reputation, the yes-man is saying "yes" to protect the boss's. He may do this because his own career is linked to his boss's, or because the boss wishes to suppress disagreeing opinions and would punish dissent. If the boss has already made a decision, then even if new information would lead him to realize it was mistaken, he may wish to stay with that decision because to change would reveal lack of confidence in his ability. In such a circumstance he would not welcome dissenting arguments.

This is important because sometimes the boss could try to develop a reputation for tolerating dissent or try to give the employee a contract that rewards truthful reporting. Ewerhart & Schmitz (2000) point out that in the Prendergast (1996) model, the boss could construct a truth-eliciting contract for the agent. The boss could ask for the agent's own opinion and for his best guess as to the boss's opinion, and reward the agent only for that best guess.

The agent would then be willing to report his own true opinion, and would show his high effort by the quality of his guess about his boss's opinion. But would the boss like to elicit the truth? That is what this paper will examine.

I will start by constructing a model in which the boss's ability to tell whether a project will be successful is unknown to some audience he cares about. This audience will observe his decision on whether to adopt the project, a later report by his agent on whether the project will be successful, the boss's decision on whether to continue with the project, and its ultimate success or failure. The audience will not observe the agent's actual information, or whether the boss chooses to hedge his decision by taking some action that equally affects its costs and benefits. The boss cares about his reputation with the audience as well as profits. The first part of the paper will show when the boss will be overly reluctant to change his decision. The second part will look into when he will hire yes-men.

The result will not be as simple as that an untalented boss prefers yes-men and a talented boss prefers honesty. We will see that the boss will hire a yes-man if the agent's report would not change the boss's public decision of whether to continue a project or drop it, even if his report might affect the boss's private decision about how much to support the project. This happens in two contexts, if the relevant audience cannot observe whether a boss's policy is to hire a yes-man or an honest agent: in a pooling equilibrium in which neither type of boss would react to bad news by dropping the project, and in a separating equilibrium in which a talented boss would continue the project and an untalented boss would drop it. In that separating equilibrium, the untalented boss will hire an honest agent, but the talented boss will hire a yes-man. The reason is that even a talented boss's reputation is harmed when the agent's signal differs from his own, and the talented boss does not find the information as useful for increasing profits as the untalented boss does. Finally, if the boss can convince the audience that his policy is to hire an honest agent, he will do so regardless of his type in any equilibrium. The boss is still hurt whenever the agent reports bad news, and the news might not affect his decision about whether to continue the project, but if the agent is honest then a report of good news will help the boss's reputation. The talented boss is helped more than the untalented boss, and so will commit to a truthful agent; the untalented boss has no choice but to imitate or be exposed as untalented.

The Model

The model has one true player: the boss. A passive "audience" observes but takes no actions, and an agent reports a signal but using a strategy determined by the boss. The boss and agent each receive an independent signal of whether a project will be successful: *good* or *bad*. If the project is successful it yields $W < 1$; if a failure, it yields -1. The prior is that the project will be successful with probability .5. To be profitable ex ante, the probability of *success* must be $p^* > .5$, derived from $Wp^* + (-1)(1-p^*) = 0$ so $p^* = \frac{1}{1+W}$.

With probability θ the boss is talented and his signal is correct with probability $\alpha = \alpha_1$. Otherwise he is untalented and it is correct with probability α_2 such that $.5 < \alpha_2 < \alpha_1$. We will assume that α_2 is big enough for its success posterior to exceed p^* and justify investment. The agent's signal is correct with probability $\gamma > .5$. We do not need to order γ with respect to α. The boss and audience do not observe the agent's signal, but they do observe a report of *good* or *bad* made by the agent.

The boss must choose first whether to hire a yes-man who always reports *good* if the boss has chosen *adopt* or an honest agent who reports his own signal. The boss then decides whether to adopt the project. After he hears the agent's report, he can revise the decision, choosing *continue/drop* or *adopt−now/reject−again*. At the same time, he chooses *hedge* or *not hedge*, where *hedge* replaces the project decision profits with profits of θW from success and $-\theta$ from failure, for $\theta < 1$. The audience observes the boss's first and second adoption decisions, the agent's report, and the success of the project, but not the signals or the *hedge* decision. It may or may not observe the decision to hire yes-men; we will look at both cases.

Let us call the audience's posterior on the boss's probability of being talented after observing his decision and the agent's report the boss's *interim reputation*, and its posterior after observing his decision, the agent's report, and the success or failure of the project his *final reputation*. The audience does not observe profit, only success.[1] We will assume the boss is risk-neutral and his payoff function is separable with weights on his profits, interim, and

[1] This assumption implies that the unobservable action *hedge* cannot be deduced ex post by observing profits.

final reputation such that with $x > 0, y > 0, 1 - x - y > 0$, so

$$U = x(profits) + y * (interim\ reputation) + (1 - x - y)(final\ reputation) \tag{1}$$

This is a model of a boss who must make a public decision with partial information and who must decide whether to acquire more information from an honest agent or to hire an uninformative yes-man. He cares about both his short-run and long-run reputation with a particular audience as well as profits. This audience might be outside investors, future employers, or present employers; it could be voters if the boss is an elected official. The audience might also be internal to the organization the boss heads; he may value his employees' esteem or believe they will work better if they have more confidence in his ability. Whoever the audience may be, if the agent reports bad news to the boss, the audience will find out, and they will also observe some but not all of his decisions, including the decision to drop a project that the boss earlier thought profitable. Thus, the boss may have to choose between maximizing profit by dropping bad projects and maintaining a good reputation. If he chooses to sacrifice profit, but does have useful information, the model also allows him to *hedge*, an action unobservable by the audience which moderates the effect of an unprofitable decision. The *hedge* decision might, for example, be how many resources to invest in the project, or how much to tell the rest of the organization to depend on its success. This will create a tradeoff between hiring yes-men and acquiring information that could increase profits even without sacrificing reputation.

The decision problem is symmetric between initial adoption or rejection of the project because we have assumed that the audience observes the agent's report and the true state of *success* or $failure$ in either case, so without loss of generality we will only go into detail about the case in which the boss has heard the signal *good*.[2]

How Beliefs Evolve when the Agent's Report Is Uninformative

[2] For analysis of the effects of career concerns when rejecting a project leads to less public information about its profitability than acceptance, see Milbourn Shockley & Thakor (2001). In the present model, if the success of the project is never observed that is equivalent to setting $(1 - x - y) = 0$. Nothing would change in the qualitative features of the equilibrium.

First, let us consider what happens if the agent is a yes-man and always reports *good*. The boss and audience will then ignore his report, so all we need calculate are the posterior probabilities resulting from the boss's signal and actions.

The boss's *success posterior*, his posterior probability on the success of the project given all available information. would be after hearing his own signal *good*

$$Pr(success|good) = \frac{Pr(good|success)Pr(success)}{Pr(good)} = \frac{\alpha .5}{\alpha(.5) + (1-\alpha)(1-.5)} = \alpha \tag{2}$$

This is monotonic in α, so $Pr(success|good, talented) > Pr(success|good, untalented)$. We have assumed that $Pr(success|good, untalented) > p^*$, so both types of boss maximize profits by accepting a project after the signal *good*.[3] Assume for the moment that in equilibrium, when he cares about reputation as well as profits, the boss adopts the project if his signal is *good* and rejects it if his signal is *bad*, something we will prove in Proposition 1.

The boss's interim reputation, the probability of being talented after he adopts the project but before its success is observed, is then

$$\begin{aligned} Pr(talented|adopt) &= \frac{Pr(adopt|talented)Pr(talented)}{Pr(adopt)} \\ &= \frac{[\alpha_1(.5)+(1-\alpha_1)(1-.5)]\theta}{[\alpha_1(.5)+(1-\alpha_1)(1-.5)]\theta+[\alpha_2(.5)+(1-\alpha_2)(1-.5)](1-\theta)} = \theta. \end{aligned} \tag{3}$$

The interim reputation is the same as the prior talent estimate.[4]

[3]This case is simple because if the boss adopts and the agent later sends the report *good*, the boss will clearly *continue*, and we can focus on the case of what happens when the agent's signal conflicts with the boss's. If we assumed that $Pr(success|good, talented) > Pr(success|good, untalented)$, then one type of boss would adopt after *good* and clearly *continue* after the agent reported *good* also, while the other type would *reject* after the agent reported *good* and clearly *continue to reject* after *bad*. The model would be more intricate, but not more illuminating.

[4]This is a consequence of the assumption that the prior probability of *success* is .5. If it were greater, then if the boss follows his signal, *adopt* would raise the audience's talent estimate, because a successful project is more likely than a failure, and a boss with a more accurate signal is more likely to adopt. This would lead to mixed-strategy semi-pooling

The boss's final talent estimate, the audience's belief as to his probability of being talented after success or failure is observed, is

$$
\begin{aligned}
Pr(talented|adopt, success) &= \tfrac{Pr(adopt,success|talented)Pr(talented)}{Pr(adopt,success)} \\
&= \tfrac{\alpha_1(.5)\theta}{\alpha_1(.5)\theta+\alpha_2(.5)(1-\theta)}. \\
&= \tfrac{\alpha_1\theta}{\alpha_1\theta+\alpha_2(1-\theta)} > \theta.
\end{aligned}
\tag{4}
$$

or

$$
\begin{aligned}
Pr(talented|adopt, failure) &= \tfrac{Pr(adopt,failure|talented)Pr(talented)}{Pr(adopt,failure)} \\
&= \tfrac{(1-\alpha_1)(1-.5)\theta}{(1-\alpha_1)(1-.5)\theta+(1-\alpha_2)(1-.5)(1-\theta)}. \\
&= \tfrac{(1-\alpha_1)\theta}{(1-\alpha_1)\theta+(1-\alpha_2)(1-\theta)} < \theta.
\end{aligned}
\tag{5}
$$

These expressions were premised on the boss following his signal, and we must now see whether he will do this in equilibrium. This is a signalling game in the other sense of "signal": the boss may take costly activities to try to persuade the audience that he is talented. Depending on the weight of the talent estimate in the utility function, this can result in implausible equilibria. For example, one equilibrium is for both types of agents to always *adopt*, regardless of the signal, with the out-of-equilibrium belief being that a agent who chooses $reject$ is untalented with probability one. If the interim and final reputations have high enough weight in the utility function, even the talented boss will choose *adopt|bad* because a deviation to *adopt|bad* would make both talent estimates fall to zero. Even if $failure$, was observed, the audience would believe that the boss was untalented because he deviated, and was only lucky in choosing $reject$.

To rule out such absurd equilibira, we will use the following refinement of perfect Bayesian equilibrium:

No Perverse Beliefs: If deviation X would reduce boss type 1's payoff more than type 2's in an equilibrium, the audience's posterior after observing X must not raise the probability of type 1.

---

equilibria in which the untalented boss adopts with some probability even if his signal is *bad*. Since this paper's focus is on the agent's additional information, I have assumed away this motive for pooling.

This is a less controversial refinement than conventional ones such as the intuitive criterion. It will, however, rule out the perverse equilibrium just described, something the intuitive criterion would not do.[5] Note that No Perverse Beliefs always allows the out-of-equilibrium belief of Passive Conjectures, in which the prior is unchanged by a deviation. We will also limit our analysis to pure-strategy equilibria, though allowing mixed-strategy deviations.

Proposition 1: The boss's initial action will always be efficient. He will adopt the project if his signal is good and reject it if his signal is bad.

Proof. Consider any equilibrium in which either type of boss chooses $adopt|bad$ or $reject|good$. First, suppose it is an equilibrium which is pooling in the sense that one action is never observed, e.g., without loss of generality, the one in which both types of boss always choose *adopt*. Either type would then deviate by choosing $reject|bad$. Doing so would not reduce his interim or final talent estimate under the No Perverse Beliefs refinement, and it would increase his expected profit. Second, suppose it is an equilibrium in which both actions, *adopt* and *reject*, are observed. Either the two types of players have equal probabilities of choosing counter to their signals or unequal. If they are equal, a player will deviate to choose in accordance with his signal, because the audience, not observing his signal, will not know he is deviating and his deviation will not change .... I NEED TO WORK ON THIS.

This can be proved for the case where there is an informative agent too. QED

Proposition 1's result that the boss's initial decision is efficient seems obvious, but it contrasts with the result in Prendergast & Stole (1996) that the boss will overinvest in the first period. In their model, the decision is about a continuous amount of investment in a function quadratic in investment, and the signal is continuous also, its variance being greater for less talented bosses. A more talented boss would maximize profit by a bigger choice of investment, and so all bosses increase their investments beyond the

[5]The intuitive criterion would require that the audience believe that the boss's type was *talented* if so doing would make a deviation profitable for the talented type but not for the untalented type. For a limited range of α_1 and α_2 this would eliminate the *always adopt* equilibrium, but it would not eliminate it when, for example, both types would benefit from rejecting on hearing the signal *bad*.

profit-maximizing level to try to increase their talent estimate. Here, payoffs are symmetric between overinvestment and underinvestment, a mistake in each direction being equally unprofitable. A talented boss is no more likely to invest from motives of pure profit, and so the untalented boss does not imitate him. Any inefficiency will arise later in the game, in connection with the agent's signal and report.

Beliefs with an Honest Agent

Now suppose that the agent truthfully reports his signal. We will look at what happens when the boss hears a *good* signal of his own, adopts, and then hears*bad* from the agent, since if he hears *good* from the agent it is easy to see that he will not drop the project.

The boss's success posterior after hearing the signal *good* himself, adopting, and hearing the report *bad* from the agent is

$$Pr(success|good, bad) = \frac{Pr(good, bad|success)Pr(success)}{Pr(good, bad)} = \frac{\alpha(1-\beta)(.5)}{\alpha(1-\beta)(.5) + (1-\alpha)\beta(1-.5)}. \tag{6}$$

We can see that $Pr(success|good, bad) < Pr(success|good)$.

The boss's interim reputation after project adoption and the agent report *bad* is

$$\begin{aligned} Pr(talented|adopt, bad) &= \frac{Pr(adopt,bad,|talented)Pr(talented)}{Pr(adopt,bad)} \\ &= \frac{[(\alpha_1(1-\gamma)(.5)+(1-\alpha_1)\gamma(1-.5)]\theta}{(\alpha_1(1-\gamma)(.5)+(1-\alpha_1)\gamma(1-.5)]\theta+(\alpha_2(1-\gamma)(.5)+(1-\alpha_2)\gamma(1-.5)][1-\theta]}. \end{aligned} \tag{7}$$

The interim reputation in equation (7) is less than θ because it is of the form $Pr(talented|adopt, bad) = \frac{K\theta}{K\theta+L(1-\theta)}$, and $K < L$. That $K < L$ can be seen from $K = (\alpha_1(1-\gamma) + (1-\alpha_1)\gamma < L = \alpha_2(1-\gamma) + (1-\alpha_2)\gamma$, which is true if $\alpha_1(1-2\gamma) < \alpha_2(1-2\gamma)$, which is true because $\gamma > .5$ and $\alpha_1 > \alpha_2$.

The boss's final reputation is, if his decision to continue with the project conveys no information (something to which we will have to return),

$$\begin{aligned} Pr(talented|adopt, bad, success) &= \frac{Pr(adopt,bad,success|talented)Pr(talented)}{Pr(adopt,bad,success)} \\ &= \frac{\alpha_1(1-\gamma)(.5)\theta}{\alpha_1(1-\gamma)(.5)\theta+\alpha_2(1-\gamma)(.5)(1-\theta)} > \theta. \end{aligned} \quad (8)$$

The value of $Pr(talented|adopt, bad, success)$ is the same as for $Pr(talented|adopt, success)$ from equation (4). The agent's bad report makes no difference in the end. This is because the signals of boss and agent are independent of each other conditional on *success*. If the audience observes *adopt* and *success*, its posterior probability of the boss having the more dependable talented signal rises. Knowing whether the agent's signal was *good* or *bad*, however, is irrelevant information, since the boss's decision did not depend on it and since the audience observes *success* directly and does not have to infer its probability.

In contrast, the interim reputation $Pr(talented|adopt, bad)$ did not take the same value as $Pr(talented|adopt)$. The reason is that before *success* is observed directly, the agent's signal *bad* is useful information about success, and hence, indirectly, about the accuracy of the boss's signal. If the agent's signal is *bad*, that reduces the audience's probability of *success*, which reduces the probability that the boss's signal of *good* was accurate.

Thus, a *bad* signal from the agent is informative about the project's success probability, and about the boss's talent if success or failure has not been observed, but it does not affect the final reputation unless via some effect it may have on the boss's continue/drop decision.

The two questions we wish to answer are (1) When will the boss change his decision after hearing a *bad* report from the agent? and (2) Will the agent report truthfully?

Maximizing Profit

What if the boss simply maximizes profit? Note first that the boss would never choose *hedge*. That decision cuts both profits and losses in half, which cuts expected profits in half. Since we have assumed that the boss is risk-neutral, he will either choose *drop* or *continue*, depending on whether his success posterior is less than p^* or not.

If the agent is a yes-man, the boss will stick with his initial decision, choosing *continue* if he chose *adopt*, since his success posterior will be unchanged.

If the agent is honest, what the boss does depends on the accuracy of the agent's signal, γ.

(1) If the agent's signal is accurate enough ($\gamma \geq \gamma^{**}$ for some value $\gamma^{**} < 1$), then

$$Pr(success|adopt, bad, talented) < p^*, \tag{9}$$

and both types of boss drop the project on hearing a bad report from the agent.

We are assuming for the moment that the boss does not care about the interim or final reputation, but these would be unchanged if he dropped the project in case (1). If he deviated by continuing the project, however, the No Perverse Beliefs refinement would require that the audience's probability that he is talented not fall, and it could even rise to certainty. Thus, later when we add career concerns, the boss may be tempted to deviate from maximizing profitability.

(2) If the agent's signal is inaccurate enough ($\gamma \leq \gamma^*$ for some value $\gamma^* > 0$), then

$$Pr(success|adopt, bad, untalented) > p^*, \tag{10}$$

and both types of boss will continue with the project on hearing a bad report from the agent, because $Pr(success|good, bad, untalented) < Pr(success|good, bad, talented)$.

The audience's talent estimate would be unaffected by the decision to continue. If the boss deviated by dropping the project, the No Perverse Beliefs refinement requires that the audience's probability that he is talented not rise. Adding career concerns thus will not affect the outcome in case (2).

(3) If the agent's signal take an intermediate accuracy level ($\gamma \in [\gamma^*, \gamma^{**})$, so that

$$Pr(success|adopt, bad, untalented) < p^* < Pr(success|good, bad, talented) \tag{11}$$

then the talented boss will continue with the project but the untalented boss will drop it.

Career concerns will matter in case (3), since if both types of boss maximize profits, the interim and final reputations will change to 0 after *drop* and 1 after *continue*.

In case (1) the agent's report can make a difference in the actions of both types of bosses, increasing expected profits from a negative amount to zero, so the bosses would ex ante welcome hearing the agent's report. In case (2), the report could not affect behavior, so the boss is indifferent about hearing it. In case (3) it would change the decision of the untalented boss, so he would welcome hearing the report but the talented boss would be indifferent.

Profitability and Career Concerns

Now let us return career concerns to the boss's objective function. Let us continue to assume that the agent reports truthfully. There are three possible equilibria, differing in whether a boss who has heard the signal *good* and adopted the project chooses *drop* after hearing the agent report *bad*.

A modest equilibrium: a pooling equilibrium in which both types of boss drop the project after hearing the agent report *bad* (and neither type ever chooses *hedge*).

A prideful equilibrium: a pooling equilibrium in which both types of boss continue the project after hearing the agent report *bad* (and one or both might choose *hedge*, depending on the signal accuracies).

A separating equilibrium: a separating equilibrium in which the talented boss continues the project after hearing the agent report *bad* (though he would choose *hedge* if the agent's report is accurate enough) but the untalented boss drops it.

There will not exist a separating equilibrium in which the talented boss drops and the untalented boss continues, because the talented boss's success posterior is higher than the untalented's, giving him expectation of higher profits and final career estimate from continuing after his initial signal of *good* and the agent report of *bad*.

Proposition 2: If the agent always reports his signal truthfully, the equilibrium can be characterized as follows based on the importance of profits to the boss, x, and the accuracy of the agent's signal, γ. Figure 1 illustrates.

(1) If and only if $\gamma \geq \gamma^{**}$ a modest equilibrium exists, with *hedge* never chosen.

(2) If x is small enough, a prideful equilibrium exists. If $\gamma > \gamma^{*}$ the untalented boss will *hedge*; if $\gamma > \gamma^{**}$ the talented boss will too.

(3) For no set of parameters does there exist both a prideful equilibrium and a separating equilibrium, regardless of x.

(4) If x is large enough, there exist values $\gamma' \in (\gamma^{*}, \gamma^{**})$ and $\gamma'' \in (\gamma^{**}, 1)$ such that: (a) if $\gamma \leq \gamma'$ a prideful equilibrium exists in which the untalented agent chooses *hedge* if $\gamma > \gamma^{*}$; (b) if $\gamma \in [\gamma', \gamma'']$ a separating equilibrium exists in which the talented agent chooses *hedge* if $\gamma > \gamma^{**}$; and if $\gamma > \gamma^{*}$ a modest equilibrium exists with *hedge* never chosen.

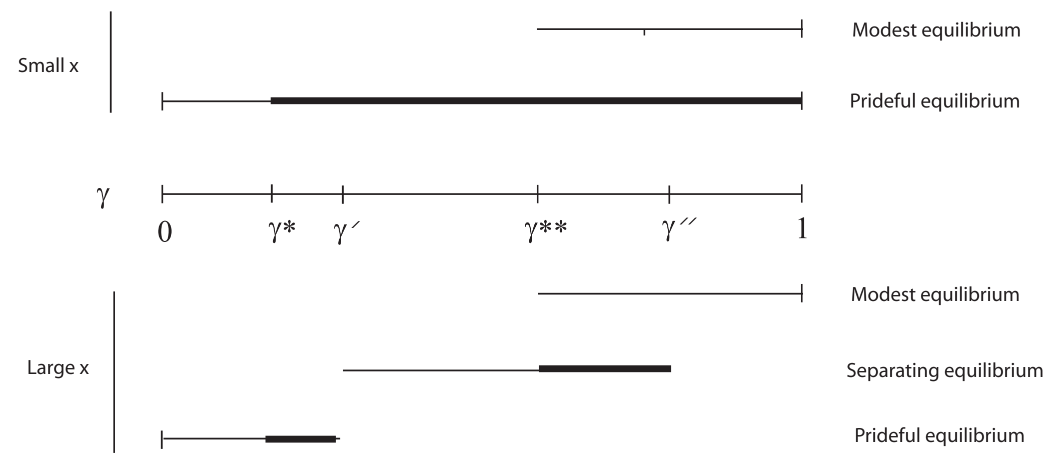

Figure 1: Equilibria with an Honest Agent
(the thick line indicates inefficiency)

Proof. The proof is divided into the parts of the proposition.

Part (1). The out-of-equilibrium belief most likely to allow the modest equilibrium to exist is that which most punishes a deviation to *continue* given the No Perverse Beliefs refinements. That belief is $Pr(talented|continue) = Pr(talented) = \theta$, passive conjectures. Under passive conjectures, the interim reputation of a deviator remains unchanged. Thus, only profits and the final reputation matter to a deviation, and both of these are highest with *drop* if and only if $\gamma \geq \gamma^{**}$, as we showed in defining γ^{**} in the section on profit maximization. Thus, neither type of boss will deviate from *drop* under passive conjectures for those values of γ, but they will under any other values.

In such an equilibrium, the boss is already maximizing expected profits by choosing *drop*. Choosing *hedge* would merely reduce the scale of profits and losses, thus reducing expected profits.

Part (2). If x is small enough, then the boss seeks to maximize the interim and final reputations, using profits only when near-indifferent about the talent estimates. That the boss does care at least slightly about profits means that Proposition 1's reasoning applies and the boss will start by following his signal. We will continue to assume without loss of generality that the boss's signal is *good*.

This equilibrium is most easily supported by the out-of-equilibrium belief $Pr(talented|drop) = 0$. This belief is permitted by the No Perverse Beliefs refinement, since it is the untalented type with his lower success posterior whose expected payoff falls least from deviating to *drop*, whether the agent's report was *good* or *bad*. Given that $Pr(talented|drop) = 0$, the interim reputation will be 0 from a deviation to drop. So will the final reputation, since with such an extreme out-of-equilibrium belief, the further information of $failure$ would not affect the posterior belief. The audience would believe that the deviator had simply been lucky in dropping the project.

Since the interim and final reputations will fall to 0 from a deviation to *drop*, if profits are sufficiently unimportant in the payoff function neither type of boss will deviate.

The boss does still wish to maximize profits to the extent that it does not hurt his reputation. Thus, if $\gamma > \gamma^*$ the untalented boss will *hedge* and if $\gamma > \gamma^{**}$ the talented boss will too. For those parameters, *continue* results in negative expected profits, and *hedge* cuts the magnitude of those expected profits in half.

Part (3). A separating equilibrium requires that the talented boss continue the project and the untalented boss drop it after hearing the agent report

bad. We can write the incentive compatibility conditions as:
$\pi_t^s(continue) > \pi_t^s(drop)$
$\pi_u^s(continue) < \pi_u^s(drop)$

A prideful equilibrium is most likely to exist with the most extreme out-of-equilibrium belief: $Pr(talented|drop) = 0$, which we will assume. Its incentive compatibility conditions are
$\pi_t^p(continue) > \pi_t^p(drop)$
$\pi_u^p(continue) > \pi_u^p(drop)$

But $\pi_u^s(drop) = \pi_u^p(drop) = Z$, since the profits are the same and the beliefs are both that the boss must be untalented with probability one. That implies $\pi_u^s(continue) < Z$ and $\pi_u^p(continue) > Z$, so $\pi_u^p(continue) > \pi_u^s(continue)$. But it is also true that the profits from *continue* are the same in in both equilibria, so the only difference would be in reputation. Both reputations are higher in the separating equilibrium than in the pooling equilibrium, implying $\pi_u^p(continue) < \pi_u^s(continue)$, yielding a contradiction. Hence, the incentive compatibility constraints for the two equilibria are inconsistent. For given parameter values, only one of the two equilibria can exist.

Part (4). In this case, profits are important in the objective function, but the interim and final reputations still have positive weight. Earlier, we found that both types of boss continuing with the project maximizes profit for $\gamma \in [0, \gamma^*]$ but that the untalented boss dropping the project maximizes profit for $\gamma \in [\gamma^*, \gamma^{**}]$. If γ^* were the lower limit for the separating equilibrium to exist, however, then if γ were slighly higher an untalented boss would reduce profit only slightly by deviating to *continue* while increasing his interim and final reputations to 1 instead of 0. Hence, the lower limit for a separating equilibrium to exist must be some $\gamma' > \gamma^*$. For lower γ, the combination of higher profit (for $\gamma < \gamma^*$) and the higher reputation from *continue* means that the only equilibrium is the prideful equilibrium. If x is large enough, however, the value of γ' cannot be greater than γ^{**}, because x can be chosen to make profit so important in the objective function that γ' can be made arbitrarily close to γ^*, even though not equal to it.

Similarly, if γ takes a value slightly higher than γ^{**} then profit maximization demands that even the talented boss drop the project, but he will continue it in a separating equilibrium because to drop it would cause his reputation to fall to 0. If x is big enough, the upper boundary γ'' for the separating equilibrium can be made arbitrarily close to γ^{**} but cannot be made to equal it.

The lower boundary for the modest pooling equilibrium is γ^{**} because under the passive conjectures out-of-equilibrium belief, the one most helpful for its existence under the No Perverse Beliefs refinement, the talent estimates are unaffected by deviation to *continue* except through the success posterior. Thus, profits and the final reputation are both maximized by the profit-maximizing action of *drop*, while the interim reputation is unaffected.

The modest pooling equilibrium already maximizes profits by choosing *drop*, so *hedge* would reduce profits. If $\gamma \in [\gamma^*, \gamma']$, the equilibrium is prideful and the untalented type chooses *continue* even though expected profit is negative, so that type will choose *hedge* to reduce the loss while leaving his reputation unaffected. If $\gamma \in [\gamma^{**}, \gamma'']$, the equilibrium is separating and the talented type chooses *continue* even though expected profit is negative, so that type will choose *hedge* to reduce the loss while leaving his reputation unaffected. ■

Parts (1), (2), and (3) of Proposition 2 together imply that if profits matter sufficiently little to the boss compared with career concerns, then the prideful equilibrium exists, the separating equilibrium does not, and the modest equilibrium exists if and only if the agent signal is sufficiently accurate. Thus, there is a possibility (when agent signals are accurate) or a certainty (when agent signals are inaccurate) of an equilibrium in which the boss ignores the agent's information even when it could help him maximize profit.

Part (4) of Proposition 2 says that if profits matter enough to the boss and the agent reports truthfully that his signal was *bad* then if the agent's signal is inaccurate there exists only the prideful pooling equilibrium, if it is more accurate there exists only the separating equilibrium, if it is still more accurate there exists both the separating and the modest pooling equilibrium (for $\gamma \in [\gamma^{**}, \gamma''])$, and if it is accurate enough there exists only the modest pooling equilibrium. The critical values, however, are not the same as we found for profit maximization, generating a corollary.

Corollary: The modest equilibrium is profit-maximizing, but in the prideful and separating equilibria both types of boss might continue with projects they believe are unprofitable after hearing the signal *bad* from the agent.

Proof. Part (1) of Proposition 2 says that the modest equilibrium drops the project for the same values of γ for which dropping it maximizes profit. Part (2) says that even if γ is such that one or both types of boss should *drop* the project to maximize profits, if x is small enough the prideful equilibrium exists and the project is continued. Part (4) says that even if x is large enough, the prideful equilibrium will exist for $\gamma \in [\gamma^*, \gamma']$, an interval for which the untalented type would drop the project to maximize profits. Parts (3) and (4) together imply that a separating equilibrium exists only if x is large enough, and that it then exists and the talented boss will continue the project even if $\gamma \in [\gamma^{**}, \gamma'']$, an interval for which the talented boss should drop the project to maximize profits. ■

If career concerns are important enough, the boss will refuse to change his initial decision, even though he might regret having made it. This would be true regardless of whether the audience sees the agent's report, since in the prideful equilibrium the *drop* decision hurts the boss's reputation regardless of what the agent reported. If the agent's report is public, however, the interim reputation is immediately affected, either rising if the report is *good* or falling if it is *bad*. The best the boss can do if he regrets his initial decision but fears to choose the public decision *drop* is to choose the private decision *hedge* and make expected profits less negative.

Thus, one reason the separating and prideful equilibria fail to maximize profits because in some circumstances the boss will fail to drop the project after hearing bad news from the agent despite wishing he had chosen to reject the project initially. Going outside the model, two other inefficiencies might result. First, any effort the agent exerts is wasted if it cannot affect the boss's decision. Second, but more ambiguously, only the separating equilibrium conveys information about the boss's talent to the outside world, information which might be important in allocating the boss to the appropriate job or in helping others know whether they should rely on his decisions. For values of γ in $[\gamma^*, \gamma']$, profit-maximizing separation is replaced by prideful pooling in equilibrium, so information fails to reach the marketplace. On the other hand, for values of γ in $[\gamma^{**}, \gamma'']$ modest pooling is profit-maximizing, but career concerns can result in separation in equilibrium, because the talented type of boss finds it worthwhile to sacrifice profits by continuing. This separation reduces profits, but it also increases the amount of information reaching the market.

The prideful equilibrium depends on out-of-equilibrium beliefs. The beliefs do not have to be as extreme as the belief $Pr(talented|drop) = 0$ used in the proof, but the equilibrium at least requires that $Pr(talented|drop) < Pr(talented = \theta$. The No Perverse Beliefs refinement also permits the passive conjectures belief $Pr(talented|drop) = Pr(talented = \theta$, but under that belief, the equilibrium would break down because for some values of α_2 and γ the untalented boss's posterior success probability $Pr(success|good, bad)$ falls below the profitability threshold p^* and he would deviate to *drop* if there is no reputational penalty.

Although it has taken considerable work to get to this point, the interesting aspect of Proposition 2– the idea that managers may be reluctant to change their decisions in light of new information– is not new. Dur (2001) uses the idea in an electoral model to explain why a political leader would be reluctant to change policies and admit poor judgement to voters. Prendergast & Stole (1996) construct a model in which bosses with varying quality of information make decisions based on career concerns and profitability. They find that young bosses will exaggerate the profitability of projects because they wish to signal that they have good information on profitability, but old bosses are unwilling to react to unfavorable new information because changing their mind would signal that their older information was of low quality.

In Prendergast & Stole (1996), the boss's new information is private and exogenous. If he does not change his decision, nobody knows his new information. Here, the boss's new information comes in the form of an agent report known to the audience. Thus, we can ask the question of whether the boss finds it worthwhile to receive the report, or whether he would wish to hire only yes-men whose reports are uninformative.

Will the Boss Hire a Yes-Man?

We have so far assumed that the agent reports honestly, but the second of our two questions is whether he will really do so. Let us start by assuming that the boss's decision about whether to hire an honest agent or a yes-man is unobservable by the audience. They do not know the personality traits of the agent nor the details of his compensation contract. In this case, the boss must trade off the desirable features of an honest agent– useful information that could be used to make a *drop* or *hedge* decision — against the possibly undesirable feature of help or harm to his reputation.

Proposition 3: If the boss's weight on profits, x, is low enough, a unique equilibrium exists in which the boss hires a yes-man, continues the project, and never chooses *hedge*.

Proposition 3: If the boss's weight on profits, x, is low enough, a unique equilibrium exists in which the boss hires a yes-man, chooses *continue*, and never chooses *hedge*.

Proof. When x is low, what matters most to the payoff are the interim and final reputations, not the profit. Hence, if we can show that the boss's reputation is better when a yes-man is hired, we know that the payoff is higher even though profits fall because the *continue* and *no hedge* decisions are made without useful information.

In any equilibrium in which either type of boss hired an honest agent, the signal *bad* would reduce the interim reputation and the signal *good* would increase it, from equation (7). If the audience believes some type of boss has hired an honest agent, that type could increase his payoff by deviating to hiring a yes-man who always reports *good.* That would increase his interim reputation whenever an honest agent would have reported *bad.* Hiring a yes-man will not affect his final reputation, because as established in Proposition 2 the boss will choose *continue* regardless of the report and so will have the same probability of success. This shows there can be no equilibrium in which some type of boss hires an honest agent.

An equilibrium in which both types of boss hire a yes-man does exist. Such an equilibrium requires us to specify out-of-equilibrium beliefs even though the audience does not observe the hiring decision, since in equilibrium the agent never reports *bad.*

It can be supported by using the same belief that Proposition 2 established would support the prideful equilibrium when the agent is honest, from equation (7). That belief can be interpreted as the result of passive conjectures- no inference about type following the discovery that the agent cannot be a yes-man— but Bayesian updating from observing the informative report of *bad*. (It cannot be supported by the simpler belief that $Pr(talented|adopt, bad) = 0$ because that would violate the No Perverse Beliefs refinement).

The boss chooses between hiring an honest agent and using Proposition 2's strategy, or hiring a yes-man and always picking *continue*. Proposition 2's strategy is also *continue*, so the only difference is the superior interim reputation gained by hiring a yes-man, so that is what the boss will do.

If both types of boss do hire a yes-man, then his signal is uninformative, and the boss's success posterior remains unchanged, so his initial decision to *adopt* remains profit-maximizing; either *drop* or *hedge* would reduce expected profits.

This has shown that both types of boss will hire a yes-man. Must the equilibrium in the resulting subgame be the prideful one of *continue* by both types, or could it be separating or a modest *drop* pooling equilibrium?

There exists no separating equilibrium because Proposition 2 established that a separating equilibrium and a prideful pooling equilibrium cannot coexist. Existence of a separating equilibrium shows that the untalented type does not care enough about reputation to sacrifice the profit-maximizing decision to get a perfect reputation.

Could there be a modest pooling equilibrium in which the boss chooses an honest agent, and then always chooses *drop* following the report *bad*? Such an equilibrium is most strongly supported, given the No Perverse Beliefs refinement, by passive conjectures: $Pr(talented|adopt, continue) = \theta$. Under that belief, the audience believes a deviation to *continue* is equally likely for either type of boss. In forming the final reputation belief, the audience will therefore use $Pr(talented|adopt, bad, success)$ and $Pr(talented|adopt, bad, failure)$, which will not change if conditioned on the *drop* decision. Thus, although the *drop/continue* decision affects expected profit, it does not affect the final reputation in the modest equilibrium. (It would under the out-of-equilibrium belief $Pr(talented|adopt, continue) < \theta$ that also might support the equilibrium, but the effect of *drop* on final reputation would be negative, so deviation is still less tempting.) Knowing the signal *bad* was received by the agent is therefore only of use for increasing profits, not final reputation.

The honest agent's report also affects the interim reputation, however. If the audience believes that the agent is honest, a *good* report increases the interim reputation and a *bad* report reduces it. Thus, the boss would deviate to hiring a yes-man who would always increase his reputation. If he hires a yes-man, however, he will never hear the report *bad* and so will never choose *drop*. ■

Proposition 3 says that if the boss puts enough weight on career concerns then not only would he fail to take action in response to bad news from the agent, he will not even hear the bad news. The boss will hire a yes-man to avoid reports of bad news which by raising doubt about his own decision would lead the audience to downgrade his reputation. This is self-defeating in that it also means that the audience will not react to an agent report of *good*, which is more likely than *bad* given that the boss's adoption decision is more often right than wrong. The yes-man's report cannot hurt the boss's reputation, but it cannot help it either. Nonetheless, no equilibrium can exist in which the audience believes the agent report, because the boss could then deviate to hiring a yes-man and always get the reputation-enhancing *good* report.

Surprisingly, the modest pooling equilibrium now fails to exist, even for parameters that in Proposition 2 supported it. In Proposition 2, the report *bad* hurt the boss's interim reputation, but all the boss could do was to make the best of it by using the valuable information to drop the project. In Proposition 3, he can avoid the harm to his interim reputation by hiring a yes-man, at the cost of making a less-informed decision with lower profits. The boss would change his action in response to the report *bad* from a truthful agent, but it is not worth the damage to his reputation.

Proposition 4. If x is large enough, there exist values $(\gamma_1, \gamma_2, \gamma_3, \gamma_4)$ such that $(\gamma^* < \gamma_1 \leq \gamma_2 < \gamma^{**})$ and $(\gamma^{**} < \gamma_3 \leq \gamma_4 < 1)$ such that:
(1) If $\gamma \in [\gamma^*, \gamma_1]$ a prideful equilibrium exists in which the boss hires a yes-man, always chooses *continue*, and never chooses *hedge*.

(2) If $\gamma \in [\gamma_1, \gamma_2]$ a prideful equilibrium exists in which the boss hires an honest agent, always chooses *continue*, and chooses *hedge* following the report *bad*.
(3) If $\gamma \in [\gamma_2, \gamma_3]$ a separating equilibrium exists in which the untalented boss hires an honest agent, the talented boss hires a yes-man, and neither chooses *hedge*.
(4) If $\gamma \in [\gamma_3, \gamma_4]$, there exists a separating equilibrium in which both types of boss hires an honest agent and the talented boss chooses *hedge* following the report *bad*.
(5) If $\gamma \geq \gamma^*$ a modest equilibrium exists in which the boss hires an honest agent and *hedge* is never chosen.

Proof. For any parameter γ, the boss has a choice between the strategies of Proposition 2, when the agent was honest, and a strategy that does not depend on the truthfulness of the report.

(1) For $\gamma \in [0, \gamma^*]$, the boss would never choose *drop* or *hedge* because *continue* maximizes profit when the agent's report *bad* is that inaccurate. The honest agent's effect on the boss's payoff is purely negative, because a report of *bad* hurts his interim reputation without changing any other decision. For ϵ arbitrarily small, at $\gamma^* + \epsilon$, the boss would choose *continue* and *hedge* if the agent is honest. There, hiring a yes-man instead would reduce profit by a negligible amount because *hedge* could no longer be used, but it would cause a discontinuous rise in the expected interim reputation because the report *bad* would disappear. If γ rises enough above γ^*, however (and x is high, as here) we know from Proposition 2 the untalented boss will be willing to sacrifice reputation for profit and deviate to *drop*. Thus, there is an interval of $\gamma \in [0, \gamma_1]$ for $\gamma_1 \in (\gamma^*, \gamma^{**}$ for which the boss will choose *continue* and a yes-man and not hedge.

(2) Part (1) showed that for γ above but near enough to γ^* the boss's equilibrium strategy will be to hire a yes-man, *continue*, and not hedge. For large enough values of γ, Proposition 2 tells us that as γ becomes larger, the equilibrium will become separating, as the profit from *drop* comes to outweigh the loss in reputation from *drop* for the untalented agent. Before that critical level of γ, however, a third strategy may be best for both types of boss: : *continue*, but hiring an honest agent and choosing *hedge* after a *bad* report from him.

Could we have an equilibrium in which both types of boss choose *continue*, but hire an honest agent and hedge? If there is a deviation to a yes-man, that changes the payoffs only when the report is *bad*. Then loss of hedging reduces profit, by an amount inversely proportional to θ, the extent to which hedging moves profit to zero. It improves the interim reputation, since the report *bad* is no longer observed. It does not affect the final reputation, since we have postulated that the boss chooses *continue* in either case. If θ is small enough, the reduction in profit outweighs the improvment in interim reputation, and the deviation increases the boss's payoff. If θ is large enough, the deviation is not profitable. Thus, the interval of γ values for which this equilibrium exists, $[\gamma_1, \gamma_2]$, might or might not be empty, depending on θ.

(3) For $\gamma \in [\gamma^*, \gamma^{**}]$ the profit maximizing strategy is to hire an honest agent and for the untalented boss to *drop* after a bad report and the talented boss to *continue*. We have seen in parts (1) and (2) that for γ close to γ^* in this interval, even if x is large if it does not equal 1 then career concerns will cause the untalented boss to choose *continue* instead. In the remainder of the interval, $[\gamma'', \gamma^{**}]$, a separating equilibrium clearly exists, since both types are maximizing profit, we have defined the value γ'' as the value at which profit outweighs reputation for the untalented agent, and the talented agent would hurt both his reputation and profit by deviating to *drop*.

For small ϵ, however, there is a separating equilibrium in which the talented boss would choose *continue* even for $\gamma^{**} + \epsilon$, because the gain in profit from *drop* would also be small, but the loss in reputation from moving to $Pr(talented|drop) = 0$ would be large. Thus, the separating equilibrium will exist up to some value $\gamma = \gamma_3$ that is strictly greater than γ^{**}.

In this equilibrium the talented boss would choose *continue* after the report *bad*. *Continue* would have higher profits than *hedge* for $\gamma < \gamma^{**}$, so the agent's information would be useless and the talented boss would prefer a yes-man to avoid the report *bad* hurting his interim reputation. For $\gamma = \gamma^{**} + \epsilon$, *continue, hedge* would have higher profit than *continue, not hedge*, but the gain in profit would be small and the loss in reputation from *bad* would be a discontinuous increase. Thus, the talented boss would choose a yes-man up to a value strictly greater than γ^{**}.

(4) Proof that there may exist $\gamma \in [\gamma_3, \gamma_4]$ for which the equilibrium is to choose an honest agent and *continue, hedge* after the report *bad* is similar

to the proof in part (2). If θ is small, then *hedge* reduces the profit loss from *continue* for $\gamma > gamma^{**}$ to a small value. Fixing that small value, for arbitrarily small ϵ it is not worth sacrificing reputation to that small value. For larger values of γ it will be, however, and the talented agent would prefer an honest agent and *continue*, *hedge* to either a yes-man and *continue*, *not hedge* or *drop*.

(5) If $\gamma \geq \gamma^{**}$, profit is maximized for by types of boss by hiring an honest agent and choosing *drop* after the report *bad*. The only reason a boss would deviate would be if *continue* would improve his reputation, and that will not happen if we use passive conjectures as the out-of-equilibrium belief so $Pr(talented|drop = \theta$. This belief thus supports the pooling equilibrium . ■

Proposition 4 tells us that that if the boss can choose whether to hire a yes-man, unobserved by the audience whose opinion he values, then even if he places a higher value on profit than on reputation, if he places any value on reputation he will often end up choosing yes-man. One class of parameters for which he will hire a yes-man is when a truthful report would not change any of his decisions anyway, either because his decision maximizes profits or because a change in decision would harm his reputation too much. The second class of parameters for which he will hire a yes-man is when when a truthful report would change his unobservable decision to hedge the project, but this is less important than the hurt to his reputation when the audience sees his agent disagree with him.

We have been assuming, however, that the audience cannot observe whether the boss hires a yes-man or not, which leads to the conclusion that they can deduce that he will. What if he can prove that he has hired an honest agent?

Proposition 5. If the boss's decision to hire a yes-man is observable by the audience, the equilibria have the following policies for hiring the yes-man and for choosing to *continue* or *drop* the project if the agent reports *bad*.

(1) If $\gamma \in [0, \gamma^*]$, there is an equilibrium in which both types of boss hire a yes-man and one in which they do not.

(2) If x is small enough, then if $\gamma \geq \gamma^*]$:

(2a) If $\gamma \geq \gamma^*]$, in one equilibrium the boss hires an honest agent and in another a yes-man. In both he chooses *continue*.

(2b) If $\gamma \geq \gamma^{**}]$, there is a third equilibrium in which the boss hires an honest agent and chooses *drop*.

(3) If x is large enough, then if $\gamma \geq \gamma^*]$:

(3a) If $\gamma \in [\gamma^*, \gamma']$, for $\gamma' > \gamma^*$ the same value as in Proposition 2, the boss hires a yes-man and chooses *continue*.

(3a) If $\gamma \in [\gamma^*, \gamma'']$, for $\gamma'' > \gamma^{**}$ the same value as in Proposition 2, the boss hires an honest agent, the talented boss chooses *continue* and the untalented boss chooses *drop*.

(3b) If $\gamma \geq \gamma'']$, the boss hires an honest agent and chooses *drop*.

Proof.

(1)

(2a) In the first equilibrium, both types choose an honest agent and *continue* and the out-of-equilibrium belief is $Pr(talented|yes - man) = 0$. No Perverse Beliefs allows this, because a deviation hurts the reputations of both types of boss equally, but helps the profits of the untalented boss more. An untalented boss is more likely to hear a *bad* report and increase his profits by choosing *drop* instead of *continue*. Since x is small, though, reputation is most important in the payoff function and neither type will wish to deviate under the given belief.

Could there be an equilibrium in which both types of boss hire a yes-man and choose *continue*, with an out-of-equilibrium belief $Pr(talented|honest) = \theta - \Delta$ for some $\Delta > 0$? Yes, because the talented boss loses the least from a deviation, because he is more likely to hear a report of *good* from the honest agent. Even so, however, if Δ is big enough, his expected interim reputation will be worse than θ. In the extreme case that $\Delta = \theta$, even a report of *good* would not raise the interim reputation above 0, and by continuity of the interim reputation posterior slightly smaller values of Δ will still leave the expected interim reputation less than θ. Since we are assuming x large, no amount of improved profits from hiring an honest agent can make up for the loss in reputation.

(2b) Small x and $\gamma \geq \gamma^{**}]$ implies existence of an additional equilibrium in which the boss hires an honest agent and chooses *drop*. This can be supported by the out-of-equilibrium belief $Pr(talented|yes-man) = \theta - Delta$ for some $\Delta > 0$. For Δ sufficiently far from θ this is permitted by No Perverse Beliefs because the untalented boss will be hurt less by deviation because in equilibrium he hears the report *bad* more often, yet his reputation will be nonetheless be hurt by deviation.

(3a) Large x and $\gamma \in (\gamma^*, \gamma']$ implies an equilibrium in which the boss hires a yes-man and chooses *continue*. Consider the out-of-equilibrium belief $Pr(talented|honest) = 0$. This satisfies No Perverse Beliefs because the untalented agent is hurt less by deviating, since his profits rise from an informative signal. As shown in Proposition 2, however, the benefit from higher profits is less than the cost from a worse reputation for this γ.

There does not exist an equilibrium in which just one type of boss hires a yes-man, because then the other type could improve his reputation from 0 to 1 by deviating to also hire a yes-man.

There does not exist an equilibrium in which both types hire an honest agent and choose *continue*. Consider the out-of-equilibrium belief $Pr(talented|yes-man) = \theta - \Delta$ for suitable $\Delta \geq 0$ depending on γ. If the untalented boss hires a yes-man, then his reputation is hurt less than talented boss's because

the agent report *bad*, whereas the talented agent hurts his reputation, by Lemma 5a below. The untalented agent, unlike the talented agent, hurts his profit by hiring a yes-man. If b γ is close to γ^* the damage to profit is small compared to the gain in reputation for smallΔ, but as γ increases, Δ can also increase and

Thus, No Perverse Beliefs is satisfied, and

loses less by deviating because the improvement to his reputation from stopping the report *bad* is greater than for the talented agent, so this satisfies No Perverse Beliefs. If Δ is small enough, this outweighs his

That violates No Perverse Beliefs because the untalented agent is hurt less by deviating, since his

(3b) Large x and $\gamma \in [\gamma', \gamma'']$. Let the out-of-equilibrium belief be $Pr(talented|yes-man) = \theta$. The talented boss will not deviate to hiring a yes-man because of Lemma 5a below. The untalented boss will not deviate because from Propsosition 2's proof that after hiring an honest agent he will prefer *drop* to *continue*, we know that he values the improvement in reputation from never hearing *bad* less than the loss in profit. If both types hire an honest agent, then Proposition 2 tells us that the equilibrium will have the talented boss choosing *continue* and the untalented *drop* for these values of γ.

Lemma 5a: Suppose the audience holds the belief$Pr(talented|hire\ honest\ agent) = \theta$. A talented boss's expected interim reputation will be greater than θ if he hires an honest agent.

Proof of Lemma: The talented boss's expected interim reputation depends on the probabilities of the reports *good* and *bad* from the agent and on how the interim reputation changes after each kind of report:

$$\begin{aligned} E_{talented\ hires} Pr(talented|adopt, agent\ report) = \; & Pr(talented|adopt, bad) Pr(adopt, bad|talented) \\ & + Pr(talented|adopt, good) Pr(adopt, good|talent \end{aligned} \tag{12}$$

This probability is a weighted average of the low $Pr(talented|adopt, bad)$ and the high $Pr(talented|adopt, good)$, since $Pr(adopt, bad|talented) + Pr(adopt, good|talented) = 1$.

Under the audience's belief, both types of boss hire an honest agent. Then, the audience would use the unconditional weighting probabilities $Pr(adopt, bad)$ and $Pr(adopt, good)$ in the equation:

$$\begin{aligned} E_{both\ hire} Pr(talented|adopt, agent\ report) = \; & Pr(talented|adopt, bad) Pr(adopt, bad) \\ & + Pr(talented|adopt, good) Pr(adopt, good) = \theta, \end{aligned} \tag{13}$$

where the value is θ because the expected value of the posterior after hearing the agent report must be the same as the prior.

Because the talented boss is less likely to hear a agent report *bad*, $Pr(adopt, bad) > Pr(adopt, bad|talented)$ and $Pr(adopt, good) < Pr(adopt, good|talented)$, so equation (12) is putting greater weight on the lower number $Pr(talented|adopt, bad)$

and less on the higher $Pr(talented|adopt, good)$ than is equation (13). Thus, the expected interim reputation in (12) is higher than in (13): $E_{talented\ hires} Pr(talented|adopt, agent$ θ. Consequently, the talented boss's expected interim reputation rises if he hires an honest agent. ■

(3c) Large x and $\gamma \in [\gamma'', 1]$. Let the out-of-equilibrium belief be $Pr(talented|yes - man) = \theta$. We know from Lemma 5a that the talented agent's expected interim reputation is higher if he hires an honest agent, and for this γ value his profit is too. The untalented agent could improve his expected interim reputation by deviating to hiring a yes-man, but at the cost of lower profit, which combined would hurt his payoff if x is large. Thus, both would hire an honest agent, and maximize their payoffs as in Proposition 2 by choosing *drop*.

There cannot be an equilibrium in which only one type of boss hires the honest agent and the other does not. Suppose only the talented boss hires the honest agent. By deviating to also hire him, the untalented boss would immediately raise his interim reputation from 0 to 1. Later, the agent report might be *bad*, but that would not alter the interim talent expectation, since using Bayes's Rule a belief of 1 is never changed by new information. Suppose instead that only the untalented boss hires the honest agent. By the same reasoning, the talented boss will deviate to hire the honest agent and raise his reputation from 0 to 1.

References

Why do Policy Makers Stick to Inefficient Decisions? Robert A.J. Dur. Public Choice. Leiden: Jun 2001. Vol. 107, Iss. 3-4; p. 221

D Hirshleifer & AV Thakor Managerial conservatism, project choice, and debt Rev. Financ. Stud. 5: 437-470. Not mentioned yet.

Corporate Conservatism and Relative Compensation Jeffrey Zwiebel The Journal of Political Economy, Vol. 103, No. 1. (Feb., 1995), pp. 1-25.Not mentioned yet.

Yes men, integrity, and the optimal design of incentive contracts*1 Christian Ewerhart & Patrick W. Schmitz Journal of Economic Behavior & Organization Volume 43, Issue 1, September 2000, Pages 115-125

Herding over the career Christopher N. Avery & Judith A. Chevalier Economics Letters Volume 63, Issue 3, June 1999, Pages 327-333

Holmstrm, Bengt (1982) "Managerial Incentive Problems - A Dynamic Perspective", in Essays in Economics and Management in Honor of Lars Wahlbeck. Helsinki: Swedish School of Economics, 1982.

Holmstrm, Bengt (1999) Managerial Incentive Problems: A Dynamic Perspective Review of Economic Studies, Volume 66, Issue 1, Page 169-182, Jan 1999.

Milbourn, Todd T., Richard L. Shockley & Anjan V. Thakor (2001) Managerial Career Concerns and Investments in Information The RAND Journal of Economics, Vol. 32, No. 2. (Summer, 2001), pp. 334-351.

Impetuous Youngsters and Jaded Old-Timers: Acquiring a Reputation for Learning Canice Prendergast & Lars Stole The Journal of Political Economy, Vol. 104, No. 6. (Dec., 1996), pp. 1105-1134.

Information Control, Career Concerns and Corporate Governance Journal of Finance, No. 4, August 2006 Fenghua Song & Anjan V. ThakorNot mentioned yet.

A Theory of "Yes Men" Canice Prendergast The American Economic Review, Vol. 83, No. 4. (Sep., 1993), pp. 757-770.

Security Analysts' Career Concerns and Herding of Earnings Forecasts Harrison Hong, Jeffrey D. Kubik & Amit Solomon (2000) The RAND Journal of Economics, Vol. 31, No. 1. (Spring, 2000), pp. 121-144.

Scharfstein, D.S.& Stein, J.C. (1990) Herd behavior and investment. American Economic Review 80, 465-479.

The New Palgrave Dictionary of Economics, 2nd Edition Srihari Govindan & Robert Wilson "Refinements of Nash equilibrium" Not mentioned yet.