"Yes Men," Integrity, and the Optimal Design of Incentive Contracts[1]

Christian Ewerhart[a,*], Patrick W. Schmitz[b,2]

[a] *The Boston Consulting Group, Westend Carree, Grüneburgweg 18, D-60322 Frankfurt a. M., Germany*

[b] *University of Bonn, Wirtschaftspolitische Abteilung, Adenauerallee 24–42, D-53113 Bonn, Germany*

Abstract. In a pioneering approach towards the explanation of the phenomenon of "yes man" behavior in organizations, Prendergast (1993) argued that incentive contracts in employment relationships generally make a worker distort his privately acquired information. This would imply that there is a trade-off between inducing a worker to exert costly effort and inducing him to tell the truth. In contrast, we show that with optimally designed contracts, which we term *integrity contracts*, the worker will both exert effort and report his information truthfully, and that hence the first best can be achieved.

JEL classification: D20, J30

Keywords: Yes men; Incentive contracts; Integrity

[*] Corresponding author. E-mail: ewerhart.christian@bcg.com.

[1] We would like to thank an anonymous referee for very helpful comments on an earlier version of the paper. Schmitz gratefully acknowledges financial support by Deutsche Forschungsgemeinschaft, SFB 303 at the University of Bonn.

[2] E-mail: schmitz@wipol.uni-bonn.de.

1 Introduction

During the last two decades, the standard principal agent model has been fruitfully extended in a number of dimensions.[3] One of the more recent developments of this literature relates to incentive problems connected with the acquisition and transmission of information in organizations.[4] More specifically, in an approach towards the explanation of the phenomenon of "yes man" behavior in organizations, Prendergast (1993) has argued that contracts cannot induce a worker to both exert costly effort on an information gathering activity and subsequently reveal his privately acquired information. This implies that the first best cannot be achieved, even though the actors are risk-neutral and there are no rents due to wealth constraints or pre-contractual private information. The basic idea underlying the argument is the following. In order to induce the worker to exert effort, the manager will have to use an incentive scheme that is based on an (imperfect) measure of the effort level chosen by the worker. The manager will therefore compare her own information with the worker's report and pay him accordingly. Now, if the worker observes, in addition to his valuable information, a signal on the manager's information, then the incentive scheme induces the worker to behave as a "yes man", i.e., he will use his second signal (which contains no new information on the true parameter) in order to bias his report towards his estimate of the manager's signal.

In this paper we argue that the "yes man" phenomenon, rather than being a necessary feature of relationships between managers and subordinates, is a consequence of suboptimal contract design. We will show that there are simple contracts which make the worker both reveal his private information and exert the first-best level of effort. We find that this result is interesting for at least four reasons. First, we stay within the contractibility assumptions of Prendergast's (1993) original model.[5] Next, although the manager's signal

[3]Cf. Hart and Holmström (1987) and Salanié (1997).

[4]See Aghion and Tirole (1997), Crémer, Khalil, and Rochet (1998), Dewatripont and Tirole (1998), Levitt and Snyder (1995), Lewis and Sappington (1997), and Shavell (1994). Cf. also Hermalin (1998) and the literature cited there.

[5]This is in contrast to a series of recent papers that prove first-best results by modifying the set of assumptions in existing models. For instance, Chung (1991), Rogerson (1992),

is statistically sufficient for the worker's second signal, we argue that nevertheless the second signal is still valuable for the manager. This runs counter the intuition that it is sufficient to ask the worker to report his useful signal only (cf. Holmström, 1979). Moreover, our main result is in accordance with a recent strand of management literature that argues in favor of supporting ethically sound behavior (see e.g. Paine, 1994) via top management directives. Finally, the view taken in this paper is also supported by the fact that in the 90's an increasing number of businesses started to implement guiding principles for their work force focussing on values like honesty and integrity, both within the firm and in relation with customers.[6]

The remainder of the article is organized as follows. In the next section, the basic model is introduced. In section 3 we formally show how the first best can be achieved by an integrity contract. The results are further discussed in section 4. Section 5 concludes. Some technical arguments have been relegated to the appendix.

2 The model

We present a slightly simplified version of Prendergast's (1993) model which is sufficient to make the point.[7] Assume that there are two risk-neutral individuals called manager and worker. The manager is responsible for a long-term project, the result of which depends on how precise her pre-project estimate of an uncertain parameter η will be. Specifically, it is assumed that the expected value of the project's long-term return accruing to the manager

Hermalin and Katz (1993), Aghion, Dewatripont and Rey (1994), Nöldeke and Schmidt (1995), Edlin and Reichelstein (1996) and Maskin and Tirole (1999) show that the first best can be achieved in the framework of Rogerson (1984) and Hart and Moore (1988).

[6]E.g., in 1996 the General Motors Board of Directors wrote "This company will have a successful and enduring life – earned by the value of its products and services, and the integrity of its people." Cf. General Motors (1999). See also Shell (1999): "Shell companies insist on honesty, integrity and fairness in all aspects of their business and expect the same in their relationships with all those with whom they do business."

[7]The simplification is that in our model the variance of the manager's signal will be exogenously given. Giving the manager the possibility to reduce this variance by exerting effort complicates the exposition but does not change the economic insights.

(after a suitable normalization) is equal to the negative of the variance of the manager's estimate of η. The joint prior of both manager and worker on η is taken to be identical to the actual distribution of the parameter, which is assumed to be normal with mean η_0 and variance $\sigma_0^2 > 0$. At the moment when the manager takes up her task she receives a verifiable signal[8]

$$\eta_m = \eta + \varepsilon_m,$$

where ε_m denotes a normally distributed error with mean 0 and variance $\sigma_m^2 > 0$. The manager is given the option to employ the worker for the task of gathering additional information. When being hired by the manager, the worker can generate a private signal

$$\eta_w = \eta + \varepsilon_w,$$

where ε_w is a normally distributed error term with mean 0. The precision of the worker's private signal is assumed to depend on how much effort he invests in the information gathering activity. For a worker exerting effort $e \geq 0$, let $C(e)$ denote his cost of effort, and $\sigma_w^2 = h(e)$ the resulting variance of the error term ε_w. Assume that $C(e)$ is continuously differentiable for nonnegative effort levels and that $C'(e)$ is nonnegative, strictly increasing, and unbounded from above with $C'(0) = 0$. Assume also that $h(e)$ is continuously differentiable for nonnegative effort levels and that $h'(e)$ is strictly negative and strictly increasing. Finally, it is assumed that the worker also privately observes a signal on what the manager has seen,

$$\eta_\lambda = \eta_m + \lambda,$$

where λ is a normally distributed error with mean 0 and variance σ_λ^2, where $0 < \sigma_\lambda^2 < \infty$.[9]

As a benchmark, consider the first-best solution, which requires $C(e) + Var[\eta|\eta_m, \eta_w]$ to be minimized. In a perfect world, the worker reports his

[8]This signal may be interpreted as a documentation that comprises the factors that eventually led to the initiation of the project. If the manager's signal were not verifiable, no effort could ever be induced.

[9]The random variables η, ε_m, ε_w, and λ are assumed to be uncorrelated.

private information η_w truthfully.[10] Given a certain effort level e, the conditional variance of the manager's posterior distribution of η given η_m and η_w reads

$$V^* = \frac{\sigma_w^2 \sigma_m^2 \sigma_0^2}{\sigma_m^2 \sigma_0^2 + \sigma_m^2 \sigma_w^2 + \sigma_0^2 \sigma_w^2}.$$

Thus, under the assumptions made, there exists an efficient effort level $e^* > 0$.[11] The necessary first-order condition is given by

$$C'(e^*) = -h'(e^*) \frac{[\sigma_0^2 \sigma_m^2]^2}{[\sigma_0^2 \sigma_m^2 + \sigma_m^2 \sigma_w^2 + \sigma_w^2 \sigma_0^2]^2}.$$

3 The first best is achievable

Under second-best conditions, i.e., if neither effort nor long-term return is contractible, the manager may want to induce the worker to exert effort by offering suitable monetary incentives. Prendergast (1993) assumes that the manager measures the worker's performance by comparing the information that is available to her with a verifiable message $\widehat{\eta}_w$ to be sent by the worker. He shows that if the worker is induced to exert strictly positive effort, then he does not report η_w honestly, yielding a variance of the manager's estimate of η which is strictly higher than V^*. This implies that either the induced effort level or the precision of the manager's estimate must be suboptimal.

[10] Note that η_m is statistically sufficient for η_λ, so that the worker does not need to report η_λ in a first-best world.

[11] To see this, note that total costs are continuous on $[0; \infty[$ and that

$$\lim_{e \to \infty} C(e) + V^*(e) = \infty,$$

by the assumptions made on $C(.)$ and by $V^*(.) > 0$. Notice that in general the second-order condition, which requires convexity of the total cost function $C(e) + V^*(e)$ with respect to e, need *not* be satisfied. In fact, $V^*(.)$ is convex in e if and only if

$$h''(e) \geq \frac{2[h'(e)]^2}{c + h(e)},$$

for all $e > 0$, where $c = \sigma_0^2 \sigma_m^2 / (\sigma_0^2 + \sigma_m^2)$. Hence, if e.g. $h(e) = \overline{\sigma}_w^2 \exp(-e)$, then for $c < \overline{\sigma}_w^2$, the function $V^*(.)$ is strictly convex in the non-empty interval $]|\ln(c/\overline{\sigma}_w^2)|; \infty[$, but strictly concave in the non-empty interval $]0; |\ln(c/\overline{\sigma}_w^2)|[$.

We shall now describe a contract that implements the first best. The contract says that the worker has to make a report consisting of *two* parts, $\widetilde{\eta}_w$ and $\widetilde{\eta}_m$. In the first part of the report, the worker is asked to reveal his valuable signal η_w. In the second part of the report, the worker is asked to announce his best estimate for the manager's information η_m. Then the worker is paid solely on the basis of the second part of his report. Specifically, the worker receives an amount w_1 if $|\widetilde{\eta}_m - \eta_m| < k$ and w_0 otherwise, for some constant k. We will refer to such contracts as *integrity contracts.* Note that an integrity contract is completely specified by the tupel (w_0, w_1, k).

While explicit integrity contracts may not be frequently observed in the real world in a literal sense, they may actually be quite common in implicit contracting. For example, it is common practice for political and economic advisors to structure their written reports for political decision makers in one part that contains the facts and another part in which conclusions and recommendations are derived. Although both parts of the report are required, the political leader typically assesses the advisor by how close the conclusions are to her gut feeling rather than by the details of the analysis. An advisor coming up with conclusions similar to the ones the political leader had derived alone thereby may gain a higher probability of being employed again in the future.[12]

The following simple observation is the key to the main result of this paper.

Observation: *A worker working under the terms of an integrity contract will truthfully reveal his acquired information in the first part of the report.*

To see this, recall that in the given setting the worker does not have an intrinsic interest in the decision to be made by the manager. Rather, it is the worker's objective to maximize his expected wage. Since the contractual wage in an integrity contract does not depend on the first part of the report (but regularly on the second), the worker is indifferent between reporting the truth and lying when announcing $\tilde{\eta}_w$, so that it is rational for him to

[12] In an alternative interpretation, a supervisor, rather than asking his subordinate for his opinion on a specific issue, would ask him two questions: "What do you think is my opinion on this matter" and afterwards "And, be honest, what is your own opinion?".

truthfully reveal η_w.[13]

The fact that integrity contracts are consistent with honest revelation of the worker's valuable information allows to resolve the tension between inducing the worker to exert effort and encouraging him to reveal his acquired information. This is the content of the following result.

Theorem 1 *There exists an integrity contract which implements the first best.*

Proof. In order to prove the theorem, it is sufficient to show that there always exists an integrity contract (w_0, w_1, k) inducing the worker to exert effort $e^* > 0$. If such a contract exists, then the first best can be achieved, because the worker has no incentive to distort the information in the first part of his report, as has been argued above.

We solve the worker's remaining decision problem (the choice of the effort level and the second part of the report) in two steps. First, assume that effort e has been chosen and η_w and η_λ are realized. Then the worker maximizes his interim net expected payoff

$$w_0 + [\text{prob}(|\widetilde{\eta}_m - \eta_m| < k|\eta_w, \eta_\lambda)](w_1 - w_0)$$

by choice of $\widetilde{\eta}_m = \widetilde{\eta}_m(e, \eta_w, \eta_\lambda)$. One can check that the worker's interim belief over the random variable η_m is normally distributed with mean $M = \mu_0\eta_w + \mu_1\eta_0 + \mu_2\eta_\lambda$, where the parameters are defined as

$$\mu_0 = \frac{\sigma_0^2\sigma_\lambda^2}{\sigma_0^2\sigma_w^2 + \sigma_0^2\sigma_m^2 + \sigma_w^2\sigma_m^2 + \sigma_\lambda^2[\sigma_w^2 + \sigma_0^2]}$$

$$\mu_1 = \frac{\sigma_w^2\sigma_\lambda^2}{\sigma_0^2\sigma_w^2 + \sigma_0^2\sigma_m^2 + \sigma_w^2\sigma_m^2 + \sigma_\lambda^2[\sigma_w^2 + \sigma_0^2]}$$

$$\mu_2 = \frac{\sigma_0^2\sigma_w^2 + \sigma_0^2\sigma_m^2 + \sigma_w^2\sigma_m^2}{\sigma_0^2\sigma_w^2 + \sigma_0^2\sigma_m^2 + \sigma_w^2\sigma_m^2 + \sigma_\lambda^2[\sigma_w^2 + \sigma_0^2]}.$$

[13] We hence follow Prendergast (1993, p. 764) and the main-stream principal-agent literature (see, e.g., Grossman and Hart 1983, p. 22) in assuming that the worker, when indifferent between any two optimal choices, chooses the one that is preferred by the principal. If the worker has arbitrary small but positive costs of lying (e.g., due to scruples or to faking pieces of evidence), he even strictly prefers to tell the truth.

For a given $k > 0$, the term prob$(|\widetilde{\eta}_m - \eta_m| < k|\eta_w, \eta_\lambda)$ (which will in general depend on e) is the integral of the density function f of the worker's interim belief on η_m over the interval $[\widetilde{\eta}_m - k, \widetilde{\eta}_m + k]$. Hence, as the Gauss density f is symmetric with respect to its mean M, strictly increasing left from M, and strictly decreasing right from M, the term prob$(|\widetilde{\eta}_m - \eta_m| < k|\eta_w, \eta_\lambda)$ is strictly increasing for $\widetilde{\eta}_m < M$ and strictly decreasing for $\widetilde{\eta}_m > M$. Thus, given a positive wage difference $w_1 - w_0$, the worker chooses $\widetilde{\eta}_m(e, \eta_w, \eta_\lambda) = M$.

Next, consider the worker's choice of the effort level. He maximizes

$$w_0 + [\text{prob}(|M - \eta_m| < k)](w_1 - w_0) - C(e).$$

Note that prob$(|M - \eta_m| < k)$ is a continuously differentiable function of σ_w^2. Hence the necessary first-order condition for the worker is given by

$$h'(e)\frac{\partial \text{prob}(|M - \eta_m| < k)}{\partial \sigma_w^2}(w_1 - w_0) = C'(e).$$

The first-order condition becomes sufficient if the associated second-order condition

$$\left[h''(e)\frac{\partial \text{prob}(|M - \eta_m| < k)}{\partial \sigma_w^2} \right. \tag{1}$$
$$\left. +[h'(e)]^2\frac{\partial^2 \text{prob}(|M - \eta_m| < k)}{\partial \sigma_w^2 \partial \sigma_w^2}\right](w_1 - w_0) - C''(e) < 0$$

holds. We show in Appendix A that there is a k such that the second-order condition holds for any positive wage differential $w_1 - w_0$, and for any effort level $e > 0$. Then, since $C'(e^*) > 0$ and $h'(e^*) < 0$ by assumption, and since

$$\frac{\partial \text{prob}(|M - \eta_m| < k)}{\partial \sigma_w^2} = -\frac{k}{\sqrt{2\pi}\sigma^3}\exp(-\frac{k^2}{2\sigma^2})\mu_0^2 < 0 \tag{2}$$

(see Appendix B), where $\sigma^2 =$Var$(M - \eta_m)$, there exists a (unique) positive wage difference $w_1 - w_0$ such that the first-order condition holds for the efficient effort level e^*. Hence, there is an integrity contract (w_0, w_1, k) that induces the worker to choose e^*, which proves the theorem. ■

To understand how efficiency can be guaranteed, recall first that the manager receives the full information from the worker under an integrity contract.

Thus, when using integrity contracts, the manager's marginal revenue from implementing a higher effort level is the same as if the information of the worker was publicly observable. Also, as the worker does not earn any rent and is solely reimbursed for his cost, the marginal costs of implementing a higher effort level for the principal are equal to the social costs. Thus, since the principal earns the social revenue and carries the social costs from a given effort level, he will prefer the first-best effort level e^* over any other implementable effort level.

While some technical work is necessary to provide a formal proof, it is intuitively clear that the worker can be induced to exert the efficient effort e^*. To see why, note that under an integrity contract the worker has an interest to improve his information about the manager's signal. Yet, the only way to improve this information is by exerting effort directed towards an improvement of his information about the true state of the world. Hence, by a suitable choice of incentives (i.e. of $w_1 - w_0$) the worker can be induced, via this external effect, to 'abuse' his technology to produce a higher accuracy of the valuable information.[14]

From a contract-theoretic point of view, it is interesting to note that asking the worker to make two announcements can be welfare-improving, even though the manager's own signal is statistically sufficient for the worker's second signal. The reason is that the second part of the report is needed as a basis for the worker's compensation which must be designed in order to provide effort incentives, while the first part is needed in order to elicit the worker's unbiased opinion.[15] In contrast, Prendergast (1993) considers contracts which ask the seller to make only one announcement. As the

[14] It is interesting to note that although the worker's payment is not directly contingent on his announcement $\widetilde{\eta}_w$, it may nevertheless affect equilibrium behavior. The reason is that, since the information transmission becomes efficient, the additional announcement changes the principal's revenue resulting from the worker's effort, so that another effort level is implemented.

[15] In order to implement the first best, it is important that the worker's expected payment depends only on the second but not on the first part of his report. Otherwise, the worker might find it in his interest to misrepresent the information in the first part of his report. If the principal is then unable to track back what the true information must have been, information transmission is inefficient so that the parties fail to achieve the first best.

implementation of the first-best level of effort $e^* > 0$ would require strict incentives (meaning $w_1 > w_0$) to induce the worker to exert positive effort, the same incentives would make the worker distort his information. By separating the information channel enabling incentives for effort exertion from that transmitting the valuable information, the dilemma can be resolved.[16]

4 Discussion

Four features of the present framework are essential to derive the efficiency result.

Absence of insider concerns: The worker in the model is not interested in the outcome of the decision to be made by the manager. Relaxing this assumption must be expected to have serious consequences on efficiency since it will become more difficult to elicit the information of the worker. Consider, e.g., a division head in a business enterprise reporting to an officer. If the officer's discretion encompasses, say, the budget of the division, it will be more likely that the division head distorts information that will affect the officer's budgetary decisions.

Exclusion of rent-seeking possibilities: The first best may not be reached if the parties can spend socially spurious effort in order to increase their income. For example, the worker could have the possibility to improve the precision of his second signal, η_λ, by investing a positive amount in a second effort variable e_λ.[17] In that case, the worker may be encouraged to first search for and then conform to the supervisor's views.[18] If the cost of a marginal effort for searching for the manager's signal is sufficiently low, then the worker will invest $e_\lambda > 0$ which is clearly inefficient as this effort generates an additional cost without revealing any valuable information.

[16]Note that while our result is basically positive since it shows that the first best can be achieved, it can also be interpreted to be negative because it contradicts Prendergast's (1993, p. 764) explanation of why in some circumstances firms may prefer not to offer incentive pay to workers. In the present model, low-powered (flat) wage schemes can never be optimal if the class of contracts is not restricted.

[17]See also Prendergast (1993, p. 760).

[18]This is a well-known phenomenon in the literature on so-called leader-member exchanges. See Deluga and Perry (1994, p. 81).

Absence of rents for the worker: In the current setting, the worker earns no rent in a contractual relationship compared to his outside option. The first best may not be attainable if the worker earns a positive rent. For example, if the worker has limited liability (i.e., $w_0 \geq 0$), his contractual rent may vary with the effort level. Hence, for the manager the marginal cost of implementing a given effort level may be different with limited liability when compared to the original setting. Thus, the manager may find it in her interest to implement an effort level different from e^*, so that a social loss may result. Similar arguments apply when the worker earns a rent because of risk-aversion, where the worker has to be reimbursed for the risk he faces under a variable wage contract. Finally, the worker may earn rents from missing markets at the stage of contracting (e.g. due to asymmetric information), again potentially with negative effects on welfare.

Complete contracts: When compared to Prendergast's (1993) inefficiency result, Theorem 1 shows that "yes man" behavior may be a consequence of exogenous restrictions on the set of feasible contracts. One of the reasons for the exclusion of integrity contract could be the potentially higher contracting costs (e.g., due to a higher complexity). However, the exact nature of such restrictions remains unclear, so that further research on these issues would be desirable.[19]

Given our result, it seems natural to ask to what extent the "yes man" problem can be mitigated by more sophisticated contract design. Beside the type of contract discussed in this paper we give three, as we think, important examples for organizational design that can improve efficiency of information acquisition and transmission in organizations.

Delegation: Inefficiencies in information acquisition and transmission may be reduced by the delegation of decisions. However, the gain from avoiding inefficiencies from the distortion of information (and the gain from possibly better incentives to search for information on the part of the worker) must be traded for the loss from a potentially suboptimal decision.[20]

Advocacy: To avoid the "yes man" problem, supervisors may want to ac-

[19] For a recent discussion of incomplete contracting, see Tirole (1999). Cf. also Grossman and Hart (1986).

[20] For a thorough discussion, cf. Aghion and Tirole (1997).

tively reward open inquiry and meticulous evaluation of their own proposals. Subordinates could be required to criticize a supervisor's proposed course of action by identifying potential flaws, presenting worst case outcomes, and suggesting alternative ideas.[21] In one-dimensional conflicts, the manager may even want to assign certain roles to her subordinates, who are then rewarded for evidence that supports only one extreme position.[22]

Third Parties: The involvement of an independent agent is a natural measure for mitigating "yes man" behavior. In the officer – division head example, typically part of the information would be acquired by a third party (such as a controlling department or an independent advisor) that may possess less information than the division head but that can be given incentives to report truthfully.

5 Conclusion

The model of Prendergast (1993) was designed to pinpoint inefficiencies in organizations resulting from the misrepresentation of information under incentive contracts. In this paper we have shown that these inefficiencies can be avoided and that the first best can be achieved by using simple *integrity contracts* that give agents a chance to be honest without hurting themselves.

One essential assumption in Prendergast's original framework that allows to derive this result is that the worker has no personal interest in the decision to be made by the manager. Since we would expect that a worker in an organization is not completely indifferent about the decisions made by his supervisor (in particular if these decisions concern his professional career), we find it more convincing to interpret Prendergast's "worker" as an outsider to the organization such as an economic or political advisor who is asked to write an independent report.

Fruitful future research on this topic may hence include the analysis of "yes man" behavior of a subordinate who has non-trivial preferences regarding the decisions to be made by his supervisor.

[21] Cf. Deluga and Perry (1994) and the literature cited there.

[22] See Dewatripont and Tirole (1998).

6 Appendix

A. There exists a k such that the second-order condition (1) holds.
We show that for a sufficiently high k, the term in square brackets in the second-order condition (1) is negative for any $e > 0$. Since $h''(e) > 0$ by assumption and $\frac{\partial \text{prob}(|M-\eta_m|<k)}{\partial \sigma_w^2} < 0$ (see Appendix B), it suffices to show that $\text{prob}(|M - \eta_m| < k)$ is strictly concave with respect to σ_w^2. We will analyze the term

$$\begin{aligned} & \frac{\partial^2 \text{prob}(|M-\eta_m|<k)}{\partial \sigma_w^2 \partial \sigma_w^2} \\ = \; & \frac{k}{2\sqrt{2\pi}\sigma^7 N} \exp(-\frac{k^2}{2\sigma^2})\mu_0^2 \left[(3\sigma^2 - k^2)\mu_0^2 N + 4\sigma^4(\sigma_0^2 + \sigma_m^2 + \sigma_\lambda^2)\right] \end{aligned}$$

as a function of σ_w^2, where $N = \sigma_0^2\sigma_w^2 + \sigma_0^2\sigma_m^2 + \sigma_w^2\sigma_m^2 + \sigma_\lambda^2[\sigma_w^2 + \sigma_0^2]$. Note first that

$$\sigma^2 = \mu_1^2\sigma_0^2 + \mu_0^2\sigma_w^2 + (\mu_2 - 1)^2\sigma_m^2 + \mu_2^2\sigma_\lambda^2.$$

Clearly, for a given triple of positive $\sigma_0^2, \sigma_m^2, \sigma_\lambda^2 < \infty$, σ^2 is bounded from above since $\sigma_w^2 = h(e) < h(0) < \infty$, and since the parameters μ_0, μ_1, and μ_2 lie in $]0,1[$. On the other hand, the term $\mu_0^2 N$ is bounded from below since σ_w^2 is bounded from above. Hence, for any sufficiently high k, the function $\text{prob}(|M - \eta_m| < k)$ is strictly concave with respect to σ_w^2.

B. Proof of (2).
We calculate the first derivative of $\text{prob}(|M - \eta_m| < k)$ with respect to σ_w^2. Recall that $M = \mu_0\eta_w + \mu_1\eta_0 + \mu_2\eta_\lambda$. Hence $\xi = M - \eta_m$ is normally distributed with mean 0 and variance

$$\sigma^2 = \mu_1^2\sigma_0^2 + \mu_0^2\sigma_w^2 + (\mu_2 - 1)^2\sigma_m^2 + \mu_2^2\sigma_\lambda^2.$$

A short calculation shows that then $\partial\sigma^2/\partial\sigma_w^2 = \mu_0^2$. Hence, by the definition of the derivative and by Leibnitz's rule,

$$\begin{aligned} & \frac{\partial \text{prob}(|M-\eta_m|<k)}{\partial \sigma_w^2} = \mu_0^2 \frac{\partial \text{prob}(|\xi|<k)}{\partial \sigma^2} \\ = \; & \frac{\mu_0^2}{\sigma^2} \frac{\partial \text{prob}(|\xi|<k)}{\partial \ln \sigma^2} \end{aligned}$$

$$
\begin{aligned}
&= \frac{\mu_0^2}{\sigma^2} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left\{ \int_{-k}^{+k} \frac{1}{\sqrt{2\pi\sigma^2 \exp(\varepsilon)}} \exp\left(-\frac{x^2}{2\sigma^2 \exp(\varepsilon)} \right) dx \right. \\
&\quad \left. - \int_{-k}^{+k} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2} \right) dx \right\} \\
&= \frac{\mu_0^2}{\sigma^2} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left\{ \int_{-k/\sqrt{\exp(\varepsilon)}}^{+k/\sqrt{\exp(\varepsilon)}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\widetilde{x}^2}{2\sigma^2} \right) d\widetilde{x} \right. \\
&\quad \left. - \int_{-k}^{+k} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2} \right) dx \right\} \\
&= -\frac{2\mu_0^2}{\sigma^2} \lim_{\varepsilon \to 0} \frac{1}{\varepsilon} \left\{ \int_{k/\sqrt{\exp(\varepsilon)}}^{k} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2} \right) dx \right\} \\
&= \frac{2\mu_0^2}{\sqrt{2\pi}\sigma^3} \exp\left(-\frac{k^2}{2\sigma^2} \right) \left(\frac{\partial}{\partial \varepsilon} \left(k \exp(-\frac{\varepsilon}{2}) \right) \right)_{\varepsilon=0} \\
&= -\frac{k\mu_0^2}{\sqrt{2\pi}\sigma^3} \exp\left(-\frac{k^2}{2\sigma^2} \right).
\end{aligned}
$$

7 References

Aghion, P., Dewatripont, M., and Rey, P., 1994, Renegotiation Design with Unverifiable Information, Econometrica 62, 257–282.

Aghion, P., and Tirole, J., 1997, Formal and Real Authority in Organizations, Journal of Political Economy 105, 1–29.

Chung, T.-Y., 1991, Incomplete Contracts, Specific Investments, and Risk Sharing, Review of Economic Studies 58, 1031–1042.

Crémer, J., Khalil, F., and Rochet, J.-C., 1998, Contracts and Productive Information Gathering, Games and Economic Behavior 25, 174–193.

Deluga, R.J., and Perry, J.T., 1994, The role of subordinate performance and ingratiation in leader-member exchanges, Group and Organization Management 19 (1), 67–86.

Dewatripont, M., and Tirole, J., 1998, Advocates, forthcoming in Journal of Political Economy.

Edlin, A. S. and Reichelstein, S., 1996, Holdups, Standard Breach Remedies and Optimal Investment, American Economic Review 86, 478–501.

General Motors, 1999, Message from the GM Board of Directors. Official web site: http://www.gm.com/about/investor/1996/

Grossman, S. J. and Hart, O. D., 1983, An Analysis of the Principal-Agent Problem, Econometrica 51, 7–45.

Grossman, S. J. and Hart, O. D., 1986, The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration, Journal of Political Economy 94, 691–719.

Hart, O. D., and Holmström, B., 1987, The Theory of Contracts, in: Bewley, T. F., ed., Advances in Economic Theory, Cambridge, 71–155.

Hart, O. D., and Moore, J., 1988, Incomplete Contracts and Renegotiation, Econometrica 56, 755–85.

Hermalin, B. E., 1998, Toward an Economic Theory of Leadership: Leading by Example, American Economic Review 88, 1188–1206.

Hermalin, B. E. and Katz, M. L., 1993, Judicial Modifications of Contracts Between Sophisticated Parties: A More Complete View of Incomplete Contracts and Their Breach, Journal of Law, Economics, and Organization 9, 230–255.

Holmström, B., 1979, Moral Hazard and Observability, Bell Journal of Economics 10, 74–91.

Levitt, S. D., and Snyder, C. M., 1997, Is No News Bad News? Information Transmission and the Role of "Early Warning" in the Principal-Agent Model, Rand Journal of Economics 28(4), 641–61.

Lewis, T. R. and Sappington, D. E. M., 1997, Information Management in Incentive Problems, Journal of Political Economy 105, 796–821.

Maskin, E. and Tirole, J., 1999, Unforeseen Contingencies, Property Rights, and Incomplete Contracts, forthcoming in Review of Economic Studies.

Nöldeke, G. and Schmidt, K. M., 1995, Option Contracts and Renegotiation: A Solution to the Hold-Up Problem, Rand Journal of Economics 26, 163–179.

Paine, L. S., 1994, Managing for Organizational Integrity, Harvard Business Review, March-April, 106-117.

Prendergast, C., 1993, A Theory of "Yes Men", American Economic Review 83, 757–770.

Rogerson, W. P., 1984, Efficient Reliance and Damage Measure for Breach of Contract, Rand Journal of Economics 15, 39–53.

Rogerson, W. P., 1992, Contractual Solutions to the Hold-Up Problem, Review of Economic Studies 59, 774–794.

Salanié, B., 1997, The Economics of Contracts: A Primer, Cambridge and London.

Shavell, S., 1994, Acquisition and disclosure of information prior to sale, Rand Journal of Economics 25, 20–36.

Shell, 1999, Our General Business Principles. Official web site: http://www.shell.com/values/

Tirole, J., 1999, Incomplete Contracts: Where Do We Stand?, Econometrica, forthcoming.