RAND Journal of Economics
Vol. 26, No. 4, Winter 1995
pp. 675–689

Costly distortion of information in agency problems

Giovanni Maggi*

and

Andrés Rodríguez-Clare**

Anecdotal evidence suggests that agents often spend resources distorting information transmitted to principals. We present a model where costly information distortion emerges as equilibrium behavior. The information structure we focus on is intermediate between (and encompasses) the cases of private information and public information: the agent can falsify the privately observed state at some cost. Although the principal can design contracts that induce no falsification, these may involve excessive information rents: falsification can be beneficial in spite of the waste of resources involved, because it helps reduce information rents. We examine how optimal contract and equilibrium payoffs change as the information structure ranges from private to public information.

1. Introduction

■ Agency relationships are very often characterized by systematic distortion of the information transmitted to the principals: regulated firms spend considerable resources trying to convince the government that their costs are higher than they really are; taxpayers make efforts to distort the accounting information they report to tax collectors; insured agents systematically magnify losses and damages in order to elicit higher compensation from insurance firms. In a study of three large U.S. corporations, Schiff and Lewin (1970) found that

> Division managers built slack into their annual divisional budgets by understating revenues and overstating costs—inflating personnel requirements, proposing unneeded projects, and failing to report the adoption of cost-lowering process improvements. This slack amounted to an estimated 20% to 25% of the division's budgeted operating expenses.

* Princeton University.
** University of Chicago.

We thank Masahiko Aoki, Thomas Hellmann, John McMillan, and the participants in seminars at Stanford University and Bocconi University for helpful discussions and comments. We owe much to Editor Michael Riordan and three anonymous referees for very detailed comments on previous drafts of this article. Giovanni Maggi's research was supported by funds from Bocconi University, Milano. Andrés Rodríguez-Clare's research was supported by the Alfred P. Sloan Foundation.

The standard principal-agent model with private information[1] does not offer an explanation for these stylized facts. The revelation principle states that any mechanism is equivalent to a truthful direct mechanism. Although the observation of untruthful mechanisms is not inconsistent with the revelation principle (because the optimal allocation can be implemented by truthful as well as untruthful mechanisms), the standard model does not explain why a principal might have a strict preference for contracts that induce falsification,[2] and more importantly, it is inconsistent with the observation that real resources are spent in the falsification process.

The standard principal-agent model can be extended to account for these stylized facts by introducing a more general information structure, namely, one in which the agent can distort, at a cost, the signal received by the principal about the state of nature. This case lies between (and encompasses) the case of purely private information and that of purely public information: the standard private-information model obtains when distorting the signal is costless, and the public-information model obtains when distorting the signal is prohibitively costly.

We argue that when information is partially private in the sense just described, inducing falsification can help the principal reduce information rents: although the principal can design a contract that induces no falsification, this may involve excessive information rents. The basic intuition is simple. Consider two types of agent, θ' and θ'', and suppose the relevant incentive is for θ' to mimic θ''. By inducing agent θ'' to spend resources in falsification, the principal makes it more costly for type θ' to mimic type θ'', thereby reducing his information rents. In other words, falsification can provide a "countervailing incentive" (in the sense of Lewis and Sappington (1989a, 1989b)) that makes it easier for the principal to induce self-selection.

The model provides an explanation for both of the stylized facts mentioned at the beginning: it explains why the principal may have a strict preference for mechanisms that give rise to falsification and why resources are spent in falsification activities. The very fact that falsifying information is costly can explain why principals may prefer falsification-inducing contracts.[3]

Our model focuses on the design of contracts that involve one productive activity by the agent, when he has private information about the cost of production. The model offers insights into the impact of changes in the information structure (i.e., the cost of falsification) on the optimal production level and on the extent of falsification. When the cost of falsifying information is low enough, the optimal production is insensitive to the falsification cost: in this case, the predictions of the standard principal-agent model in terms of production distortions are robust to the introduction of a cost of falsification. But if the cost of falsification is relatively high (i.e., information structure is closer to the public-information benchmark), the optimal production level becomes sensitive to the information structure, approaching the first-best level as the cost of

[1] See, for example, Guesnerie and Laffont (1984), Baron and Myerson (1982), and Laffont and Tirole (1986).

[2] There are two notable exceptions. The model of Green and Laffont (1986) differs from the standard principal-agent model in that the agent can misreport his private information only within certain bounds. They show that under certain conditions the optimal contract may induce the agent to lie. Dye (1988) shows that if the agent's private information is multidimensional and he is unable to communicate all dimensions to the principal, information distortion may be optimal. In both of these models, distorting information is costless, and no waste of resources occurs in equilibrium.

[3] Bushman and Indjejikian (1993) offer another explanation for the occurrence of information distortion. They argue that if a manager performs several tasks and the shareholders receive information only about an aggregate measure of the manager's actions, it may be desirable to induce accounting distortions to improve the allocation of the manager's incentives across tasks. This explanation is different from ours, in that it relies on the multiple-task structure of the contract; we show that there may be a rationale for information distortions even if the agent is required to perform a single productive task.

falsification increases. As far as the agent's information rents are concerned, when the falsification cost is low, all types earn positive rents (except for the least-efficient type). When the falsification cost is high, a whole interval of less-efficient types are kept on their reservation utility, and the size of this interval increases as the falsification cost increases.

Interesting welfare implications emerge from the analysis. As information becomes more public, the principal's utility increases and the agent's utility decreases, but total welfare—defined as the sum of utilities—follows a nonmonotonic path: although total welfare is maximal under purely public information, increasing the publicness of information may decrease welfare. This result is reminiscent of some well-known second-best theorems. For example, in incomplete-market models, increasing the number of markets may be socially harmful, although the social optimum is achieved in a complete-market economy. Another welfare implication is that if the principal can invest in the quality of the accounting system to make it more costly for the agent to falsify information, the social value of this kind of investment falls short of the principal's private value and thus the investment will exceed the socially optimal level.

There are other models in the principal-agent literature in which costly distortion of information may arise in the optimal contract. Lacker and Weinberg (1989) study an optimal sharecropping contract in which a risk-averse agent observes the state (crop) after contracting and can falsify it (e.g., hide some of the crop) at a cost. Their main result is that under some conditions, inducing falsification improves risk sharing between the principal and the agent, and hence the optimal contract may entail falsification even though this involves a direct waste of resources. This explanation relies on the agent's risk aversion and so is different from the one proposed in our article, where falsification arises with a risk-neutral agent.

Information distortion arises also in a model by Laffont and Tirole (1993). In their model, the agent can take an action to reduce the true value of the cost, as well as an action to inflate the cost signal observed by the principal (both actions being unobservable). The agent derives a direct benefit from cost padding, since it allows him to appropriate money from the company, but he can be detected by the principal with a certain probability. Also, the agent has limited liability, so the principal is limited in the severity of the penalty imposed on the agent in case of detection; hence, to reduce the incentive to engage in cost padding, she may have to reward the agent with extra rents in the contingency he is not detected. For certain parameter values, they find that the optimal contract entails cost padding. The logic of information distortion in Laffont and Tirole (1993) is quite different from that in our model: in their article, cost padding is directly beneficial to the agent, and the principal may prefer to let the agent engage in cost padding rather than pay the extra rent necessary to prevent it. In our model, distorting information is personally costly to the agent, and the principal induces it to create countervailing incentives to lie.[4]

Another strand of literature related to our model is the "costly state verification" approach, as for example in Townsend (1979) and Baron and Besanko (1984). In these models the principal has to incur a cost in order to observe a signal of the true state of nature, and the agent can take no action to manipulate this signal. The focus of these models, as well as their formal structure, is different from that of our model, in that

[4] Our article is also related to Lewis and Sappington (1989a, 1989b). They argue that the presence of countervailing incentives implies the optimality of "inflexible rules" (i.e., pooling). In contrast with this prediction, we find that under costly information distortion, the optimal contract tends to be fully separating. The reason Lewis and Sappington's (1989a) results do not apply here is that their model assumes the agent's utility is convex in the private parameter, whereas in our model the agent's utility is concave in the private parameter. This point is explained in Maggi and Rodríguez-Clare (forthcoming), who extend Lewis and Sappington's (1989a) model to more general structures of countervailing incentives.

the agent does not engage in wasteful activities to bias the information received by the principal, the phenomenon we focus on. This class of models also includes that of Dunne and Loewenstein (1995), which appears elsewhere in this issue. They analyze situations in which several agents compete for a principal's project and show that if the principal cannot commit to a monitoring policy *ex ante,* the winning bidder may be induced to claim higher costs than realized *ex post.* In contrast to our model, in Dunne and Loewenstein there is no costly distortion of the signal observed by the principal: in their model, if the principal monitors she finds out the true cost level.

The article is organized as follows. In Section 2 we present the basic model. In Section 3 we study the impact of the information structure on the optimal contract and welfare. In Section 4 we discuss some extensions of the model and conclude.

2. The model

■ Suppose a principal contracts with an agent to produce a certain amount of some good, q, and compensates the agent with a monetary transfer, y. A level q of output generates revenue $V(q)$ for the principal, with $V'(q) > 0$ and $V''(q) < 0$. The principal's von Neumann-Morgenstern utility is assumed to be linear in income:

$$U^P(q, y) = V(q) - y.$$

The cost for the agent of producing q is given by Θ_q, with θ denoting the agent's marginal cost.[5] The informational structure of the model is as follows. The marginal cost θ is a random variable distributed over the interval $[\theta_0, \theta_1] \equiv \Theta$ according to cumulative and density functions $F(\theta)$ and $f(\theta)$. Assume $F(\theta)/f(\theta)$ is increasing in θ. The agent privately observes the realization of θ before signing the contract, while the principal does not. The principal observes a signal of marginal cost, $s = \theta + e$, where $e \in [-\infty, +\infty]$ denotes an unobservable action that the agent can take to distort the signal, in either direction. The cost associated with action e—which we refer to as the cost of falsification—is given by $c(e)$. We make the following assumptions about the cost of falsification.

Assumption 1. $c(e) \geq 0$; $c(0) = 0$; $c(e)$ twice continuously differentiable; $c''(e) > 0$.

Note that these assumptions imply $c'(0) = 0$.[6]

The cost of falsification captures the degree of "publicness" of information: if this cost is zero, the model coincides with the standard private-information model; if it is infinite, the model coincides with the public-information model. In Section 4 we extend the analysis to allow for a noise in the signal received by the principal, and we argue that the main results remain unchanged in the case of noisy signal.

The agent's utility is assumed to be linear in income.[7] Noting that $e = s - \theta$, the agent's utility can be written as $U(q, s, y, \theta) = y - \theta q - c(s - \theta)$.

---

[5] We could assume a more general production-cost function $C(q, \theta)$; the critical requirement is that it satisfy the single-crossing property, i.e., that $C_{\theta q}$ not change sign.

[6] The assumption of globally convex and differentiable cost of effort, which will simplify the analysis considerably, is standard in the literature. This assumption is not essential to the main insight of the article; what is essential for the optimal contract to entail falsification is that the marginal cost at $e = 0$ is not too high, and that the cost is locally convex at $e = 0$. This point is made in a previous version of this article, and a similar result arises in Lacker and Weinberg (1989).

[7] As long as θ is observed before signing the contract, risk neutrality is not essential to the model. The same results would obtain if the agent's utility were of the form $g(y - \theta q - c(s - \theta))$, with $g' > 0$, $g'' < 0$.

We assume that the agent has a constant reservation utility, normalized to zero. This implies that the contract has to give all types a nonnegative utility (participation constraint).[8]

By the revelation principle we can derive the optimal allocation $\{q(\theta), s(\theta), y(\theta)\}$ as the solution to the following problem:

$$\max_{q(\theta),\, s(\theta),\, y(\theta)} E[U^P(q(\theta), y(\theta))], \tag{P}$$

subject to

$$\theta \in \operatorname*{argmax}_{\Theta r} U(q(\theta_r), s(\theta_r), y(\theta_r), \theta), \qquad \forall \theta \in \Theta \tag{IC}$$

$$U(\theta) \equiv U(q(\theta), s(\theta), y(\theta), \theta) \geq 0, \qquad \forall \theta \in \Theta. \tag{PC}$$

There are two types of mechanism that can implement the optimal allocation: a revelation scheme, which assigns q, s, and y as functions of the agent's message, or a "no-communication" contract, for example a contract that assigns production and transfer as functions of the cost observation, $\{q(s), y(s)\}$—a mechanism that seems more natural in this context. Melumad and Reichelstein (1989) derive conditions under which the optimal allocation can be (at least approximately) implemented by a no-communication contract: it can be checked that the case examined here satisfies those conditions.

To solve problem (P) we make use of the following lemma, the proof of which is in the Appendix.

Lemma 1. A necessary condition for (IC) to hold is

$$U'(\theta) = U_\theta(q, s, y, \theta) = -q(\theta) + c'(s(\theta) - \theta). \tag{N}$$

Condition (N), together with the monotonicity conditions $q'(\theta) \leq 0$ and $s'(\theta) \geq 0$, is sufficient for (IC).

The usual procedure to solve a problem like (P) (see, for example, Guesnerie and Laffont (1984)) involves (1) replacing the (IC) constraint with condition (N) in Lemma 1; (2) solving the problem ignoring the monotonicity conditions and conjecturing that the (PC) constraint binds only at one extreme; and (3) checking that the solution to this "relaxed" problem satisfies the monotonicity conditions and the (PC) constraint for all θ.

This procedure works if U_θ does not change sign, but this is not necessarily true in our case. As a consequence, it is possible that the solution to the relaxed problem violates the (PC) constraint for some θ. For this reason, we have to introduce the (PC) constraint explicitly in the optimization problem. Observing that the principal's utility can be rewritten as $U^P = V(q) - \theta q - c(s - \theta) - U$, the problem reduces to the following:

$$\max_{q(\theta),\, s(\theta),\, U(\theta)} E[V(q) - \theta q - c(s - \theta) - U], \tag{P'}$$

subject to

[8] We assume that the revenue from production is high enough relative to the cost θq, for all θ, so that the principal will induce all types to produce a strictly positive amount. This assumption allows us to avoid problems of optimal cutoff.

$$U'(\theta) = -q(\theta) + c'(s(\theta) - \theta), \qquad \forall \theta \in \Theta \tag{IC}$$

$$U(\theta) \geq 0, \qquad \forall \theta \in \Theta. \tag{PC}$$

We treat s and q as the control variables and U as the state variable. The Hamiltonian for this problem is given by

$$H = [V(q) - \theta q - c(s - \theta) - U]f(\theta) - \mu(q - c'(s - \theta)) + \tau U,$$

where μ is the costate variable and τ is the multiplier of the constraint $U \geq 0$.

To solve the problem, we make use of a set of sufficient conditions due to Seierstad and Sydsaeter (1987) for an optimal control problem with pure state constraints.[9] The first-order conditions for the maximization of the Hamiltonian with respect to q and s are

$$(V'(q) - \theta)f(\theta) - \mu = 0 \tag{1}$$

$$-c'(s - \theta)f(\theta) + \mu c''(s - \theta) = 0. \tag{2}$$

If we assume that H is concave in (q, s), conditions (1) and (2) are sufficient for the maximization of the Hamiltonian.[10] The other conditions are

$$\mu'(\theta) = -\partial H/\partial U = f(\theta) - \tau(\theta) \qquad \text{(costate equation)} \tag{3}$$

$$U'(\theta) = -q(\theta) + c'(s(\theta) - \theta) \qquad \text{(state equation)} \tag{4}$$

$$\tau(\theta)U(\theta) = 0, \quad \tau(\theta) \geq 0, \quad U(\theta) \geq 0 \qquad \text{(complementary slackness)} \tag{5}$$

$$\mu(\theta_0)U(\theta_0) = 0, \quad \mu(\theta_1)U(\theta_1) = 0, \quad \mu(\theta_0) \leq 0, \quad \mu(\theta_1) \geq 0 \qquad \text{(transversality conditions).}^{11} \tag{6}$$

Since H is concave in U, conditions (1) through (6) are sufficient for an optimum. Our strategy to solve the problem is to conjecture a solution and verify that it satisfies these conditions. The key step is to conjecture the costate variable $\mu(\theta)$. To this end, we need some additional notation. Let $q^*(\mu, \theta)$ and $s^*(\mu, \theta)$ denote respectively the optimal levels of q and s as functions of μ and θ, derived from conditions (1) and (2), and let $\mu_0(\theta)$ be the solution in μ to $-q^*(\mu, \theta) + c'(s^*(\mu, \theta) - \theta) = 0$.

The schedule $\mu_0(\theta)$ just defined will play a key role in conjecturing the costate variable. In words, $\mu_0(\theta)$ is the value of the costate variable such that the agent's utility is constant ($U'(\theta) = 0$). This means that if the (PC) constraint is to be binding on a nondegenerate level, $\mu(\theta)$ must be equal to $\mu_0(\theta)$ on that interval.

The structure of the optimal contract depends critically on whether or not $\mu_0(\theta)$ crosses the distribution function $F(\theta)$. In the Appendix we show that $\mu_0(\theta) > 0$ and that $\mu_0(\theta)$ cannot cross $F(\theta)$ from below (given the assumption that F/f is increasing). Consequently, only two cases are possible: (i) $\mu_0(\theta)$ lies above $F(\theta)$ (or equivalently,

[9] To be precise, the conditions stated below are more restrictive than the ones in Seierstad and Sydsaeter (1987). It is easy to check that if the conditions below are satisfied, Seierstad and Sydsaeter's sufficient conditions are satisfied as well.

[10] A sufficient condition for the concavity of H is $-c''f + \mu c''' \leq 0$. The solution we will conjecture entails $0 \leq \mu(\theta) \leq 1$ for all θ, so it suffices to assume that $c''' \leq c''f$, a condition that c''' be "not too big."

[11] In the problem studied by Seierstad and Sydsaeter (1987) the state variable is fixed at the initial point and free at the final point, whereas in the case considered here the state variable is free at both extremes. One can derive the transversality conditions (6) by a straightforward extension of Seierstad and Sydsaeter's result.

$\mu_0(\theta_1) \geq 1$), and (ii) $\mu_0(\theta)$ crosses $F(\theta)$ once from above (or equivalently, $\mu_0(\theta_1) < 1$). We analyze these two cases separately.

□ **Case (i): $\mu_0(\theta_1) \geq 1$.** If $\mu_0(\theta_1) \geq 1$, then at the optimal contract, the costate variable is given by $\mu(\theta) = F(\theta)$, the agent's rents $U(\theta)$ are decreasing for all θ, and the (PC) constraint is binding only at θ_1. To see this, let $q^1(\theta) \equiv q^*(F(\theta), \theta)$ and $s^1(\theta) \equiv s^*(F(\theta), \theta)$. Clearly, $q^1(\theta)$ and $s^1(\theta)$ represent the solution to problem (P′) when the (PC) constraint is imposed only at θ_1. Since $\mu_0(\theta) \geq F(\theta)$ for all θ, then the slope of the information rents induced by $q^1(\theta)$ and $s^1(\theta)$ is negative for all θ: $U'(\theta) = -q^1(\theta) + c'(s^1(\theta) - \theta) \leq 0$.[12] Therefore, in this case the standard procedure of imposing only the terminal condition $U(\theta_1) = 0$ gives the solution to the problem.

From condition (2), and the fact that $\mu(\theta) = F(\theta)$, one can see that all types (except θ_0) exaggerate their cost: $s(\theta) > \theta$. On the other hand, from condition (1), the optimal production level $q(\theta)$ is lower than the first-best level (except at θ_0), as in the costless-falsification case. Also, it can be checked that $q(\theta)$ is nonincreasing and $s(\theta)$ is nondecreasing, so the sufficient conditions for incentive compatibility are satisfied, and the solution to problem (P′) represents the solution to the original problem (P).

□ **Case (ii): $\mu_0(\theta_1) < 1$.** As we mentioned above, $\mu_0(\theta_1) < 1$ implies that $\mu_0(\theta)$ crosses $F(\theta)$ once and from above. So for θ close to θ_1, we have $\mu_0(\theta) < F(\theta)$. This implies that the optimal costate variable cannot be equal to $F(\theta)$ for all θ: if we set $\mu(\theta) = F(\theta)$, rents would be increasing ($U'(\theta) > 0$) for θ close to θ_1, and since $U(\theta_1) = 0$ (from the transversality conditions), the (PC) constraint would be violated. It can be checked that the sufficient conditions for an optimum are satisfied when the costate variable is given by

$$\mu(\theta) = \begin{cases} F(\theta) & \text{for } \theta_0 \leq \theta \leq \theta^* \\ \mu_0(\theta) & \text{for } \theta^* < \theta \leq \theta_1, \end{cases} \tag{7}$$

where θ^* is the value of θ such that $\mu_0(\theta) = F(\theta)$. From (3), and the fact that $\mu_0'(\theta) < f(\theta)$ for $\theta \in [\theta^*, \theta_1]$, it follows that $\tau(\theta) > 0$ for $\theta \in [\theta^*, \theta_1]$; that is, the (PC) constraint is binding (rents are zero) on this interval. The optimal contract is readily derived from the optimal costate variable. The production and falsification levels are given by $s(\theta) = s^*(\mu(\theta), \theta)$ and $q(\theta) = q^*(\mu(\theta), \theta)$, and $U(\theta)$ is determined by the (IC) constraint and the endpoint condition $U(\theta^*) = 0$. In the Appendix we check that $q(\theta)$ and $s(\theta)$ are respectively nonincreasing and nondecreasing in θ, provided c''' is not too high.[13]

The following proposition summarizes the structure of the optimal contract. Let $q^F(\theta)$ denote the first-best level of output, defined implicitly by $V'(q) = \theta$.

Proposition 1. At the optimal contract, all types except θ_0 exaggerate their cost ($s(\theta) > \theta$) and produce less than the first-best amount ($q(\theta) < q^F(\theta)$). Furthermore, in case (i) ($\mu_0(\theta_1) \geq 1$), the (PC) constraint is binding only at θ_1, and the optimal $q(\theta)$ and $s(\theta)$ are determined by equations (1) and (2) with $\mu(\theta) = F(\theta)$.

[12] To prove this, note that by definition of $\mu_0(\theta)$ we have $-q^*(\mu_0(\theta), \theta) + c'(s^*(\mu_0(\theta), \theta) - \theta) = 0$ for all θ. But we also have $\partial q^*/\partial\mu < 0$ and $\partial s^*/\partial\mu > 0$. Hence, for $\mu_0(\theta) > F(\theta)$, we have

$$q^1(\theta) = q^*(F(\theta), \theta) > q^*(\mu_0(\theta), \theta)$$

and $s^1(\theta) = s^*(F(\theta), \theta) < s^*(\mu_0(\theta), \theta)$. This implies that $-q^1(\theta) + c'(s^1(\theta) - \theta) < 0$ for all θ.

[13] In particular, a sufficient condition for monotonicity is $c''' \leq -fV''(c'')^2$. This condition is satisfied, for example, in the case of quadratic cost of falsification.

In case (ii) ($\mu_0(\theta_1) < 1$), the (PC) constraint is binding on the interval $[\theta^*, \theta_1]$, and on the interval $[\theta_0, \theta^*]$, $q(\theta)$ and $s(\theta)$ are determined by equations (1) and (2) with $\mu(\theta) = F(\theta)$. On the interval $[\theta^*, \theta_1]$, $q(\theta)$ and $s(\theta)$ are determined by equations (1) and (2) with $\mu(\theta) = \mu_0(\theta)$, or equivalently by

$$V'(q) - \theta = c'(s - \theta)/c''(s - \theta) \tag{8}$$

$$q - c'(s - \theta) = 0. \tag{9}$$

Although the principal can design a contract that induces no falsification ($s(\theta) = \theta$ for all θ), Proposition 1 suggests that this kind of contract can be improved upon by a falsification-inducing contract. The intuition for this basic result is quite simple. Consider first a situation in which there are only two types, θ' and θ'', with $\theta'' > \theta'$. By inducing agent θ'' to spend resources in falsification, the principal makes it more costly for agent θ' to mimic type θ'', thereby reducing his information rents. Inducing falsification introduces a "countervailing incentive" to lie, in the sense of Lewis and Sappington (1989a, 1989b). Introducing this countervailing incentive has a cost for the principal because type θ'' must be compensated for the additional cost of falsification (since the (PC) constraint is binding for θ''). But since the marginal cost of falsification is zero at $s = \theta$, inducing a small amount of falsification involves a second-order increase in transfers to θ'', whereas the savings in information rents are a first-order effect. Inducing falsification is therefore beneficial for the principal.

Coming back to the continuous-type formulation, the slope of the information rents—which is also a measure of the incentive to lie—is given by

$$U'(\theta) = -q(\theta) + c'(s(\theta) - \theta).$$

By inducing $s(\theta) > \theta$, which implies $c'(s(\theta) - \theta) > 0$, the principal can push the incentive to lie (i.e., $|U'(\theta)|$) closer to zero. Inducing the less efficient types to falsify, and paying them more, is functional to saving rents on the more efficient types. These savings outweigh the increase in compensation to the less efficient types.

Next we comment on the differences between case (i) and case (ii), but first we offer an interpretation of the sign of $\mu_0(\theta_1) - 1$, which determines the relevant regime.

Roughly speaking, μ_0 measures the size of the distortion needed to neutralize information rents ($U'(\theta) = 0$): when μ_0 is small, a small distortion is sufficient to neutralize information rents. Under public information ($c(\cdot) = \infty$), μ_0 is equal to zero; under private information ($c(\cdot) = 0$), μ_0 is high, possibly infinite. In other words, μ_0 can be interpreted as a rough measure of the publicness of information, which in turn determines the strength of countervailing incentives. Case (ii) obtains when the information structure is closer to the public-information benchmark, and hence countervailing incentives are stronger. This interpretation of the two cases will become more clear in the next section, where we parameterize the publicness of information in a very simple way, by assuming a quadratic cost of falsification.

In case (i), the optimal production and the optimal falsification levels are determined independently, respectively by conditions (1) and (2) with $\mu = F$. These conditions require that the distortions in the production and falsification activities be equal to each other and to the hazard rate: $V' - \theta = c'/c'' = F/f$. Changes in the production value function $V(\cdot)$ do not affect the optimal falsification,[14] and the optimal level of production is insensitive to changes in the cost of falsification $c(\cdot)$. Production is the

[14] If the production cost were represented by a more general function, $C(q, \theta)$, changes in the cost function would have no impact on the optimal falsification either.

same as in the costless-falsification benchmark and is given by $q^*(F(\theta), \theta)$. In this case, the predictions of the standard private-information model in terms of output distortions are robust to the introduction of a cost of falsification. This is because inducing falsification reduces information rents but does not affect the relationship between production and information rents.

In case (ii), the contract has a different structure in the two intervals $[\theta_0, \theta^*]$ and $[\theta^*, \theta_1]$. On the interval $[\theta_0, \theta^*]$, the optimal production is the same as in the case of purely private information and is determined independently of the falsification technology, as in case (i). On the interval $[\theta^*, \theta_1]$, conditions (8) and (9) imply that $q(\theta)$ and $s(\theta)$ are no longer independent, so that, for example, changes in the cost of falsification do affect the optimal level of output. Intuitively, inducing falsification makes the participation constraint binding, and through this channel it affects the relationship between production and information rents. Condition (8) requires that the distortions in q and s be equalized, as in case (i). Condition (9) requires that information rents be constant ($U'(\theta) = 0$), which is necessary for (PC) to be binding: this condition replaces the condition that distortions be equal to the hazard rate.[15]

Notice that in case (ii) the optimal contract is closer to the first-best contract than it is in case (i): for θ in $[\theta^*, \theta_1]$, $q(\theta)$ and $s(\theta)$ are closer to their first-best levels,[16] and information rents are zero. This is intuitive, since in case (ii) countervailing incentives are stronger than in case (i). A final remark: on the interval $[\theta^*, \theta_1]$ the optimal output and falsification are independent of the distribution $F(\theta)$, an uncommon feature in the optimal contract literature.

The basic model analyzed so far admits an interesting extension. In many situations, the agent may be constrained to display a state that lies within the support of the density function, $[\theta^*, \theta_1]$. For example, in the sharecropping model of Lacker and Weinberg (1989), θ represents the total crop, and it is unfeasible for the agent to display a negative crop. Following the cost interpretation of θ, it might be impossible for the agent to convince the principal that his costs are higher than a certain level. Thus it might be required that the displayed cost s lie between θ_0 and θ_1 (but it is easy to see that only the constraint $s \leq \theta_1$ is binding). Maggi and Rodríguez-Clare (1994) analyze in detail this version of the model, which we call the "carrier constraint" case. Here we summarize the relevant results.

At the optimal contract, all types except θ_0 and θ_1 overrepresent their cost ($s > \theta$), with a right interval of types displaying the highest cost, θ_1. Information rents are higher in the presence than in the absence of a carrier constraint. This is intuitive, because in the presence of a carrier constraint the principal cannot induce as much falsification as she would like to reduce the slope of information rents. When the falsification cost is sufficiently high, information rents are nonmonotonic: less efficient types may be better off than more efficient types. In particular, if the falsification cost is very high, an interval of middle types earns no rents, while types in the right tail of the distribution (except θ_1) earn positive rents. This is a consequence of the fact that the carrier constraint is "more binding" for the less efficient types, so falsification is less effective as a countervailing-incentive device for these types.

In the next section we specialize to a quadratic cost of falsification, in order to analyze the impact of changes in the information structure on the optimal contract and on the principal's and agent's welfare.

[15] The feature that the participation constraint binds for intermediate types appears also in other principal-agent models with countervailing incentives, for example, Lewis and Sappington (1989b). Maggi and Rodríguez-Clare (1995) investigate the common mathematical structure of countervailing-incentives models, deriving conditions under which the participation constraint binds for interior types.

[16] To see this, note that for θ in $[\theta^*, \theta_1]$ the optimal costate variable is closer to zero in case (ii) ($\mu < F$) than in case (i) ($\mu = F$), and the size of the distortion is increasing in μ.

3. The impact of the information structure on the optimal contract and welfare

■ Suppose the cost of falsification is given by $c(s - \theta) = \alpha(s - \theta)^2$. In this case, the parameter of $\alpha \geq 0$ captures the degree of publicness of information. If $\alpha = 0$, falsification is costless, therefore information is purely private. As α increases, it becomes more costly to falsify information about θ, and for $\alpha = +\infty$ the public-information model obtains. Our concern in this section is to study how changes in α affect the optimal allocation, the payoffs of the agent and the principal, and the total welfare.

We assume that θ is uniformly distributed over $\Theta = [0, 1]$, and specialize to a quadratic value function for output: $V(q) = bq - (d/2)q^2$, with $b > 0$ and $d > 0$. To ensure interior solutions, we assume $b > 2$.

The first step is to derive $\mu_0(\theta)$ and compare it with $F(\theta) = \theta$. Simple algebra reveals that $\mu_0(\theta) = (b - \theta)/(1 + 2\alpha d)$. There is a critical level of α, namely $\alpha^* = (b - 2)/2d$, such that $\mu_0(\theta)$ crosses $F(\theta)$ (once) if and only if $\alpha > \alpha^*$. Next we analyze separately the cases $\alpha < \alpha^*$ and $\alpha > \alpha^*$.

□ **Case (i): $\alpha < \alpha^*$.** In this case, plugging the quadratic specifications in conditions (1) and (2) and imposing $\mu(\theta) = F(\theta) = \theta$, we get the optimal production and falsification levels:

$$q(\theta) = (b - 2\theta)/d \tag{10}$$

$$s(\theta) - \theta = \theta. \tag{11}$$

Notice that both the production and the falsification levels are independent of the publicness of information (α). Also note that the optimal contract is discontinuous at $\alpha = 0$: the extent of falsification does not tend to zero as information becomes purely private. A small cost of falsification can determine a relatively big amount of falsification in equilibrium.

To derive the impact of α on the agent's utility, observe that the slope of the agent's utility in equilibrium is given by $U'(\theta) = -q(\theta) + 2\alpha(s(\theta) - \theta)$. Since $q(\theta)$ and $s(\theta)$ are independent of α, and $s(\theta) \geq \theta$, $U'(\theta)$ becomes less negative as α increases. Therefore, since $U(\theta_1) = 0$, we conclude that the agent is worse off for each θ (except for θ_1) as α increases. As far as the principal's welfare is concerned, straightforward algebra reveals that she is better off as α increases. Total welfare (for a given θ) is equal to $W(\theta) \equiv V(q(\theta)) - \theta q(\theta) - c(s(\theta) - \theta)$. Since $q(\theta)$ and $s(\theta)$ do not depend on α, and the total waste of resources is increasing in α, total welfare decreases as information becomes more public.

□ **Case (ii): $\alpha > \alpha^*$.** In this case, $\mu_0(\theta)$ crosses $F(\theta)$ at $\theta^* \equiv b/2(1 + \alpha d) < 1$. Hence, the optimal costate variable coincides with $F(\theta)$ for $\theta < \theta^*$, and with $\mu_0(\theta)$ for $\theta > \theta^*$. From the first-order conditions (1) and (2) we find that the optimal contract entails

$$q(\theta) = \begin{cases} (b - 2\theta)/d & \text{for } \theta < \theta^* \\ 2\alpha(b - \theta)/(1 + 2\alpha d) & \text{for } \theta \geq \theta^* \end{cases} \tag{12}$$

$$s(\theta) - \theta = \begin{cases} \theta & \text{for } \theta < \theta^* \\ (b - \theta)/(1 + 2\alpha d) & \text{for } \theta \geq \theta^*. \end{cases} \tag{13}$$

In contrast with the previous case, now α does affect the optimal production and falsification. The optimal contract becomes sensitive to the cost of falsification only when this cost is high enough: there is a "band of inertia" around the private-information case ($\alpha = 0$), within which the optimal contract does not change. Notice that as α tends to $+\infty$, θ^* converges to zero and $(q(\theta), s(\theta))$ converges to $(q^F(\theta), \theta)$: as we approach the public-information benchmark, the optimal production and falsification both tend to their first-best levels.

As far as the equilibrium payoffs are concerned, the agent's rents clearly decrease with α for all θ, converging to zero as α tends to $+\infty$, and the principal's expected utility can be shown to increase with α. The impact of α on total expected welfare is more interesting. Changes in α have opposite effects for different values of θ. First, for $\theta > \theta^*$, an increase in α brings the allocation closer to the first-best, and welfare is enhanced; some algebra shows that $\partial W(\theta)/\partial\alpha = 1/(1 + 2\alpha d)^2 > 0$. Second, for $\theta < \theta^*$, the allocation (q, s) is insensitive to α, therefore, just as in the previous section, we have $\partial W(\theta)/\partial\alpha < 0$. When α increases sufficiently, however, θ^* approaches zero, thus the second effect comes to dominate the first one. In the limit as α approaches $+\infty$, total welfare approaches the first-best level for all θ, given by $W^F(\theta) = V(q^F(\theta)) - \theta q^F(\theta)$.[17]

This last result bears emphasis. Total welfare, both contingent on θ and in *ex ante* terms, follows a nonmonotonic path as the information structure changes from purely private to purely public information. Although total welfare is maximum under purely public information, increasing the publicness of information may be socially harmful. This result is reminiscent of some well-known second-best theorems: for example, in the incomplete-markets literature, increasing the number of markets may be socially harmful, although the complete-market economy leads to the social optimum. The reason for the nonmonotonicity of welfare lies in the "band of inertia" around $\alpha = 0$: here the optimal allocation (production and falsification) fails to get closer to the first-best allocation as information becomes more public, while the waste of resources in falsification increases (because the cost of falsification α increases). In Figure 1 the principal's utility, the agent's utility, and the total welfare are plotted against α.

Another implication of the welfare analysis concerns the case in which the publicness of information is endogenous. Suppose the principal can choose the level of α, incurring a cost $\hat{C}(\alpha)$. An interpretation, for example, is that the principal can improve the quality of the accounting system, making it more difficult for the agent to falsify information. First note that from the social point of view, the optimal level of α is either zero or some value higher than α^*: a value lower than α^* can never be socially optimal. The principal will choose a level of α that equates its marginal cost $\hat{C}'(\alpha)$ with the marginal increase in her expected utility (assuming an interior solution). In general the principal will overinvest in the accounting system, because the social marginal benefit of the investment, $dEW/d\alpha$, is lower than her private marginal benefit, $dEU^P/d\alpha$. This follows from the fact that $dEU^A/d\alpha < 0$, which implies that

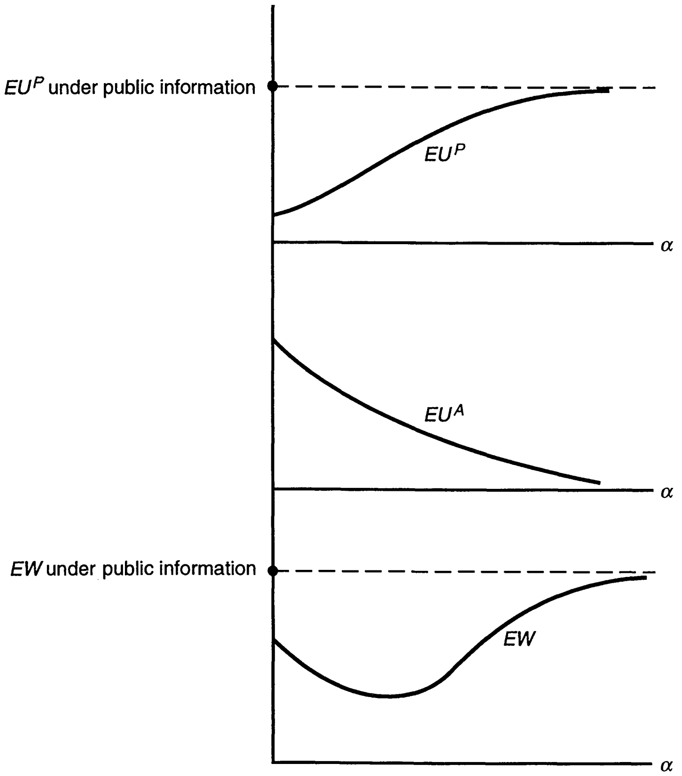
$$dEW/d\alpha = dEU^P/d\alpha + dEU^A/d\alpha < dEU^P/d\alpha.$$

4. Agency problems under partially private information

■ Thus far, this article has made two related but distinct points. First, we have argued that the presence of a cost of falsifying information provides an explanation for the observed phenomenon that resources are spent distorting information in agency relationships. Second, we have argued that costly information distortion is an interesting

[17] To prove this, one must prove that the resources spent in falsification, given by $\alpha(s(\theta) - \theta)^2$, converge to zero as α goes to $+\infty$. Plugging in the expression for the optimal $s(\theta)$, the waste of resources is equal to $\alpha(b - \theta)^2/(1 + 2\alpha d)^2$, which tends to zero as α goes to $+\infty$.

FIGURE 1



way to model situations that lie between the extremes of purely private and purely public information. Here we focus exclusively on this second point. In particular, we extend our results to a more general model of partially private information and argue that the basic model studied in the previous sections captures the essential elements of this more general model.

The more general information structure that we have in mind allows for a noise in the principal's observation of the signal and a cost for the principal to acquire the signal. Formally, the signal observed by the principal can be written as $S = \theta + e + \epsilon$, where ϵ is a noise with cumulative distribution $G(\epsilon)$. The principal has to incur a cost β in order to observe the signal. This setup coincides with the one analyzed in the previous sections if $\beta = 0$ and if $G(\epsilon)$ is degenerate at zero (no noise). On the other hand, if the agent's cost of falsification is infinite (falsification is not possible) the model coincides with the costly-state-verification model of Baron and Besanko (1984).

Consider first the case of noisy signal, but retain the assumption of costless signal observation ($\beta = 0$). Let $(q^*(\theta), e^*(\theta))$ denote the optimal allocation under no noise, as derived in the previous sections. Caillaud, Guesnerie, and Rey (1992) show that under conditions that are satisfied in this model, the optimal allocation under noisy observation coincides with that under noiseless observation.[18] Thus, the optimal allocation derived in the previous sections is not affected by the presence of noise in the principal's observation of the signal.

[18] In this context, a contract specifies a transfer as a function of the observables q and S, and of the report θ_r: $y(q, S, \theta_r)$. Caillaud, Guesnerie, and Rey (1992) show that there always exists a transfer schedule linear in the observables that implements the allocation $(q^*(\theta), e^*(\theta))$ for all distributions of ϵ (what they call "universal implementation").

Results change when it is costly for the principal to observe the signal ($\beta > 0$), but not in a very surprising way. The essential modification is that if β assumes intermediate values, the principal will choose not to observe the signal for certain values of the report θ_r, namely when it is lower than some critical level θ_a.[19] This allows the principal to decrease information rents in two ways. First, observing the signal for the high-θ types introduces countervailing incentives for these types. Second, this decreases the incentive of types below θ_a to mimic types above θ_a. For $\theta > \theta_a$, the optimal allocation coincides with that under costless signal observability, that is, $(q^*(\theta), e^*(\theta))$, and for $\theta < \theta_a$ it coincides with the optimal allocation under signal unobservability (as, for example, in Baron and Myerson (1982)).

If $c(e)$ is quadratic and given by $c(e) = \alpha e^2$, as in Section 3, the parameters α and β summarize the information structure: as α increases or β decreases, information becomes more public. One can examine how changes in α and β affect the optimal contract. First, as we saw in the previous sections, as α increases, countervailing incentives become stronger and information rents decrease; this is relevant for the region $\theta > \theta_a$, where the principal observes the signal. Second, as β falls, the critical level θ_a decreases and countervailing incentives are active for more types, decreasing information rents even further.

To summarize, there is a tight relationship between the degree of publicness of information and the strength of countervailing incentives that the principal can "create." Thus, the design of optimal contracts under partially private information can largely be viewed as the optimal design of countervailing incentives for the agent.

Appendix

■ *Proof of Lemma 1.* Let $\tilde{U}(\theta_r, \theta) \equiv U(q(\theta_r), s(\theta_r), y(\theta_r), \theta)$. The (IC) condition can be written as $\theta = \text{argmax}_{\Theta}\tilde{U}(\theta_r, \theta)$. A necessary condition for this to hold is $\tilde{U}_1(\theta, \theta) = 0$ for all θ. But this implies $(d/d\theta)\tilde{U}(\theta, \theta) = \tilde{U}_1(\theta, \theta) + \tilde{U}_2(\theta, \theta) = \tilde{U}_2(\theta, \theta)$. This equality can be written as $U'(\theta) = U_\theta(q, s, \theta)$, which is condition (N).

To prove sufficiency, note that $\tilde{U}_{12}(\theta_r, \theta) = -q'(\theta_r) + c''(s(\theta_r) - \theta)s'(\theta_r)$. Since $c'' > 0$, the monotonicity conditions $q' \le 0$ and $s' \ge 0$ imply $\tilde{U}_{12}(\theta_r, \theta) \ge 0$, for all (θ_r, θ). This implies that

$$\tilde{U}_1(\theta_r, \theta) \le \tilde{U}_1(\theta_r, \theta_r) = \tilde{U}_1(\theta, \theta) = 0$$

for $\theta_r > \theta$, and hence for $\theta_r > \theta$ we have

$$\tilde{U}(\theta_r, \theta) - \tilde{U}(\theta, \theta) = \int_\theta^{\theta_r} \tilde{U}_1(s, \theta)\, ds \le \int_\theta^{\theta_r} \tilde{U}_1(s, s)\, ds = 0,$$

which implies $\tilde{U}(\theta_r, \theta) \le \tilde{U}(\theta, \theta)$ for $\theta_r > \theta$.

Applying the same argument for $\theta_r < \theta$, sufficiency is established. *Q.E.D.*

Proof that $\mu_o(\theta) > 0$ and that $\mu_0(\theta)$ cannot cross $F(\theta)$ from below. As a first step, it is useful to derive $\partial q^*/\partial\theta$, $\partial q^*/\partial\mu$, $\partial s^*/\partial\theta$, and $\partial s^*/\partial\mu$. Differentiating condition (1), we get $\partial q^*/\partial\theta = (f^2 - \mu f')/f^2V''$, $\partial q^*/\partial\mu = 1/fV''$. Differentiating (2), we get $\partial s^*/\partial\theta = 1 - c'f'/A$, $\partial s^*/\partial\mu = c''/A$, where $A \equiv fc'' - \mu c'''$. If we assume strict concavity of H, we have $A > 0$.

To show $\mu_0(\theta) > 0$, note that $U_\theta = -q^*(\mu, \theta) + c'(s^*(\mu, \theta) - \theta)$, hence $dU_\theta/d\mu = -\partial q^*/\partial\mu + (\partial s^*/\partial\mu)c''$. Given the results above, we have $dU_\theta/d\mu > 0$. Since $s^*(0, \theta) = \theta$ and $q^*(0, \theta) = q^F(\theta)$, at $\mu = 0$ we have $U_\theta = -q^F(\theta) < 0$. We can then conclude that $U_\theta = 0$ requires $\mu > 0$. By definition of μ_0, this implies $\mu_0 > 0$.

To show that $\mu_0(\theta)$ cannot cross $F(\theta)$ from below, we show that $\mu_0'(\theta) < f(\theta)$ when $\mu_0(\theta) < F(\theta)$. Differentiate the equality that defines μ_0, $-q^*(\mu, \theta) + c'(s^*(\mu, \theta) - \theta) = 0$, with respect to μ and θ, to obtain $\mu_0'(\theta) = -[\partial q^*/\partial\theta + (1 - \partial s^*/\partial\theta)c'']/[\partial q^*/\partial\mu - (\partial s^*/\partial\mu)c'']$.

Plugging in the expressions we derived for $\partial q^*/\partial\theta$, $\partial q^*/\partial\mu$, $\partial s^*/\partial\theta$, and $\partial s^*/\partial\mu$, the inequality $\mu_0' < f$ can be written as

[19] If β is low enough, the principal will observe the signal for all values of θ_r; if β is high enough, she will never observe the signal.

$$[(c'')^2 f - c'c''f']/A > (2 - \mu_0 f'/f^2)/V''. \tag{A1}$$

To show that (A1) is satisfied when $\mu_0 < F$, we show that when $\mu_0 < F$ the left-hand side is positive and the right-hand side is negative. The assumption that F/f is increasing implies $1 - Ff'/f^2 > 0$. Therefore, $2 - Ff'/f^2 > 0$, and hence for $\mu_0 < F$ the right-hand side is negative. Now note that $1 - Ff'/f^2 > 0$ implies $1 - \mu_0 f'/f^2 > 0$ (for $\mu_0 < F$), hence $f^2 - c'ff'/c'' > 0$ (using condition (2)). Therefore, we have

$$(c'')^2 f - c'c''f' > 0.$$

But this implies that the left-hand side is positive. *Q.E.D.*

Proof of monotonicity for Proposition 1. To check whether the monotonicity conditions are satisfied, note first that $dq^*/d\theta = \partial q^*/\partial\theta + (\partial q^*/\partial\mu)\mu'(\theta)$ and $ds^*/d\theta = \partial s^*/\partial\theta + (\partial s^*/\partial\mu)\mu'(\theta)$. We need only worry about case (ii), since case (i) can be seen as a special case of case (ii) with $\theta^* = \theta_1$.

Let us examine $dq^*/d\theta$ first. On the interval $[\theta_0, \theta^*]$ we have $\mu'(\theta) = f(\theta)$. Plugging in the expressions for $\partial q^*/\partial\theta$ and $\partial q^*/\partial\mu$ derived earlier, we obtain $dq^*/d\theta = (2 - Ff'/f^2)/V''$. Using the assumption that the hazard rate is increasing, we get $dq^*/d\theta > 0$. On the interval $[\theta^*, \theta_1]$ we have $\mu(\theta) = \mu_0(\theta)$. Using the expression for $\mu_0'(\theta)$ derived earlier in this Appendix and some algebraic manipulation, we obtain

$$dq^*/d\theta = -f(c'')^2/[A - fV''(c'')^2] < 0.$$

As far as $ds^*/d\theta$ is concerned, on the interval $[\theta_0, \theta^*]$ we have $ds^*/d\theta = [fA + (f^2 - Ff')c'']/fA > 0$. On the interval $[\theta^*, \theta_1]$ we have $ds^*/d\theta = [-\mu c''' - fV''(c'')^2]/fABV''$, where $B \equiv \partial q^*/\partial\mu - c''\partial s^*/\partial\mu < 0$. A sufficient condition for $ds^*/d\theta > 0$ is that $\mu c''' + fV''(c'')^2 < 0$, which is satisfied if $c''' < -fV''(c'')^2$. *Q.E.D.*

References

BARON, D. AND BESANKO, D. "Regulation, Asymmetric Information and Auditing." *RAND Journal of Economics,* Vol. 15 (1984), pp. 447–470.

——— AND MYERSON, R. "Regulating a Monopolist with Unknown Costs." *Econometrica,* Vol. 50 (1982), pp. 911–930.

BUSHMAN, R.M. AND INDJEJIKIAN, R.J. "Stewardship Value of 'Distorted' Accounting Disclosures." *Accounting Review,* Vol. 68 (1993), pp. 765–782.

CAILLAUD, B., GUESNERIE, R., AND REY, P. "Noisy Observation in Adverse Selection Models." *Review of Economic Studies,* Vol. 59 (1992), pp. 595–615.

DUNNE, S.A. AND LOEWENSTEIN, M.A. "Costly Verification of Cost Performance and the Competition for Incentive Contracts." *RAND Journal of Economics,* Vol. 26 (1995), pp. 690–703.

DYE, R.A. "Earnings Management in an Overlapping Generations Model." *Journal of Accounting Research,* Vol. 26 (1988), pp. 195–235.

ERICSON, R. "The Classical Soviet-Type Economy: Nature of the System and Implication for Reform." *Journal of Economic Perspectives,* Vol. 5 (1991), pp. 11–27.

GREEN, J.R. AND LAFFONT, J.-J. "Partially Verifiable Information and Mechanism Design." *Review of Economic Studies,* Vol. 53 (1986), pp. 447–456.

GUESNERIE, R. AND LAFFONT, J.-J. "A Complete Solution to a Class of Principal-Agent Problems, with an Application to the Control of a Self-Managed Firm." *Journal of Public Economics,* Vol. 25 (1984), pp. 329–369.

LACKER, J.M. AND WEINBERG, J.A. "Optimal Contracts Under Costly State Falsification." *Journal of Political Economy,* Vol. 97 (1989), pp. 1345–1363.

LAFFONT, J.-J. *The Economics of Uncertainty and Information.* Cambridge, Mass.: MIT Press, 1989.

——— AND TIROLE, J. "Using Cost Observation to Regulate Firms." *Journal of Political Economy,* Vol. 94 (1986), pp. 614–641.

——— AND ———. "Cost Padding, Auditing and Collusion." In J.-J. Laffont and J. Tirole, eds., *A Theory of Incentives in Procurement and Regulation.* Cambridge, Mass.: MIT Press, 1993.

LEWIS, T. AND SAPPINGTON, D. "Countervailing Incentives in Agency Problems." *Journal of Economic Theory,* Vol. 49 (1989a), pp. 294–313.

——— AND ———. "Inflexible Rules in Incentive Problems." *American Economic Review,* Vol. 79 (1989b), pp. 69–84.

MAGGI, G. AND RODRÍGUEZ-CLARE, A. "Costly Distortion of Information in Agency Problems." Mimeo, 1994.

——— AND ———. "On Countervailing Incentives." *Journal of Economic Theory,* Vol. 66 (1995), pp. 238–263.

MCAFEE, R. AND MCMILLAN, J. "Organizational Diseconomies of Scale." Mimeo, California Institute of Technology, 1990.

MELUMAD, N.D. AND REICHELSTEIN, S. "Value of Communication in Agencies." *Journal of Economic Theory,* Vol. 47 (1989), pp. 334–368.

SCHIFF, M. AND LEWIN, A. "The Impact of People on Budgets." *Accounting Review,* Vol. 45 (1970), pp. 259–268.

SEIERSTAD, A. AND SYDSAETER, K. *Optimal Control Theory with Applications.* New York: North-Holland, 1987.

TOWNSEND, R.M. "Optimal Contracts and Competitive Markets with Costly State Verification." *Journal of Economic Theory,* Vol. 21 (1979), pp. 265–293.