
Using Cost Observation to Regulate Firms

Author(s): Jean-Jacques Laffont and Jean Tirole

Source: *The Journal of Political Economy*, Vol. 94, No. 3, Part 1 (Jun., 1986), pp. 614-641

Published by: The University of Chicago Press

Stable URL: <http://www.jstor.org/stable/1833051>

Accessed: 26/02/2009 07:17

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=ucpress>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



The University of Chicago Press is collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Political Economy*.

Using Cost Observation to Regulate Firms

Jean-Jacques Laffont

Université de Toulouse

Jean Tirole

Massachusetts Institute of Technology

The paper emphasizes the use of accounting data in regulatory or procurement contracts when the supplier (1) has superior information about the cost of the project and (2) invests in cost reduction. The main result states that, under risk neutrality, the supplier announces an expected cost and is given an incentive contract linear in cost overruns. This (optimal) contract moves toward a fixed-price contract as the announced cost decreases. An investment choice is then introduced and the use of a rate-of-return regulation is studied.

I. Introduction

The literature on the control of public firms or private monopolies can be divided into a literature studying the properties of given incentive schemes and a few recent papers designed to characterize optimal control mechanisms. The major interest of the earlier literature¹

Support from the Commissariat au Plan and the National Science Foundation is gratefully acknowledged. We thank two referees and P. Picard for helpful comments on an earlier draft.

¹ The literature on "Soviet bonus schemes" (see, e.g., Ellman 1973; Fan 1975; Bonin 1976; Weitzman 1976) and on the rate-of-return regulation (see, e.g., Crampes [1982] for a study with asymmetric information) considers the revelation of production possibilities. Another literature (Domar 1974; Tam 1979, 1981; Finsinger and Vogelsang 1981, 1982; Vogelsang 1983) assumes that the firm has superior information about demand. Although the latter literature is relevant in some cases, we would expect informational asymmetries about production possibilities to be more important in gen-

stems from the simplicity of the schemes studied, which are easily related to what can be observed in planned economies or in large corporations. However, they are ad hoc. There is clearly a need for a normative theory that will derive optimal incentive schemes, study the performance of these ad hoc schemes, and test the soundness of the intuition on which they are based.

Recently a normative theory has emerged from the nonlinear pricing literature and the more abstract incentive theory developed to deal with the free-rider problem. In this approach the regulator/planner is viewed as a Bayesian statistician who has prior knowledge about cost and demand conditions. The optimization problem of the regulator is to maximize the expected social welfare under the constraint of the decentralization of information. The outcome of the analysis is the characterization of optimal incentive schemes given the objective functions and the observations made by the regulator.

Papers in this tradition (Loeb and Magat 1979; Baron and Myerson 1982; Sappington 1982) study the control of a private monopoly when the demand function is common knowledge and the cost function can be parameterized by one real number.² The optimal incentive mechanism in general entails a welfare loss compared with what could be achieved under perfect information.

Costs are easy to observe, at least at the firm's level. The value of cost observation to the planner depends on what he attempts to control. If he monitors a single project in a multiproject firm, the firm can shift expenses to and from the particular project, at both real and accounting levels. In a first approximation it is reasonable to assume that the planner does not perfectly observe the firm's cost for the project.³ When the planner controls the entire firm, aggregate cost information becomes very valuable. If cost observability is introduced into the Baron-Myerson model, it is then possible to infer the true cost parameter and to reach the first-best with appropriate penalties.

eral (e.g., the regulator can run a consumer survey or use the firm's output and price data to estimate demand). Bergson (1978) has stressed the role of distributional aspects. Our use of a social cost of transfers can be viewed as a formalization of the concern for equity in the design of incentive schemes.

² The analysis can be conceptually generalized to any number of parameters. In particular, demand may also be parameterized in the same way as costs if the regulator does not know demand. But then the optimal schemes cannot be derived analytically. Also, the assumption that these functions can be parameterized does not appear restrictive. Actually, even the manager derives his information from a finite number of observations and can have only an approximation of the true cost and demand functions even in a stationary environment. See also Guesnerie and Laffont (1984) for an application to the control of labor-managed firms and some additional theoretical developments.

³ Formally, the planner can extract some information from aggregate cost observation. The point is that the high dimensionality of the characteristics space reduces the value of the information.

In this paper we introduce possibly noisy cost observability as well as an unobservable effort variable. Section II describes the model. A regulated firm⁴ produces a public good. The planner observes the firm's output and cost but not its efficiency parameter, its effort, and the cost disturbance. The firm knows its efficiency before contracting. After contracting, it chooses an output and a level of effort, which together with an additive uncertainty result in a cost level. Its reward depends on output and cost. (See Sec. IV for other interpretations of the model.) Both parties are risk neutral, and the firm can reject the contract if it is not guaranteed a minimum payoff. Section III gives a complete technical analysis of the firm's and the planner's optimization problems. We suggest that this section be skipped in a first reading by readers who are mainly interested in the regulatory implications of the model. Section IV, the main section of the paper, summarizes the properties of the optimal incentive scheme and of the firm's performance. The optimal scheme is linear in ex post cost: the planner pays a fixed sum (which can be determined at the date of contracting) and then reimburses a fraction of the costs. This fraction is inversely related to the fixed transfer and decreases with the firm's output (or efficiency; it increases with the firm's announced expected cost in another interpretation). Some implications are drawn about when the optimal scheme resembles cost-plus-fixed-fee or fixed-price contracts. In particular, it is shown that the more concerned about output the regulator is, the more the optimal contract resembles a fixed-price contract. Section IV also gives an alternative interpretation of the model that embodies the choice of a quality level. Section V introduces a choice of technology. The firm can trade off variable and fixed costs. Our assumption that accounting data are (at least partially) observable allows us to study the efficiency properties of rate-of-return regulations. In our model capital accumulation is insufficient when investment is not directly observable (i.e., only total cost is), but an Averch-Johnson rule does not increase welfare when investment is observable. Section VI discusses the case of a risk-averse firm. Section VII compares our work with related contributions and presents conclusions.

⁴ We exclude in this paper the solution proposed by Demsetz (1968) of designing an auction and giving the market to the best offer, by assuming that there is a single informed firm. One justification can be that huge increasing returns to scale do not make it worthwhile to set up several firms to benefit from their competition. A related reason, when the parties renegotiate the contract rather than set up a new relationship, comes from the advantages of sticking with the incumbent. For more details see Williamson (1976).

II. The Model

A firm produces a single output q at (monetary) cost $C = (\beta - e)q + \epsilon$. Variable $e \geq 0$ is a level of effort, which decreases the initial marginal cost β . The efficiency parameter β belongs to $[\underline{\beta}, \bar{\beta}]$ where $\underline{\beta} > 0$; ϵ is a random variable with zero mean and denotes an ex post cost disturbance. We will interpret ϵ as a *forecast error*, unknown to the firm when it chooses its output and effort levels, and we assume that ϵ is independent of the parameters and choice variables of the model. Alternatively, ϵ could denote an independent *observation (accounting) error on cost* with absolutely no change in our results.

A more general form for the influence of effort on marginal cost could be assumed without much change (see n. 15). Also, the effort in principle could influence the fixed cost as well: $C = (\beta - e)q + \alpha - ke + \epsilon$. The technical analysis then becomes more complex, but the same qualitative results hold if one assumes that the optimal incentive scheme is differentiable and that the various second-order conditions are satisfied (properties that are proved in the simpler case in which effort influences the marginal cost only, the case considered in this paper; see Laffont and Tirole 1984).

The output is not marketed by the firm;⁵ it is, for example, a public good that provides a consumer surplus $S(q)$ ($S' > 0$, $S'' < 0$). The planner observes and reimburses the cost incurred by the firm and pays in addition a *net* monetary transfer t . The utility level of the firm's manager is then $U = Et - \psi(e)$, where $\psi(e)$ stands for the disutility of effort. We assume that $\psi'(e) > 0$ and $\psi''(e) > 0$ for any $e > 0$. In the whole paper, expectations are taken with respect to ϵ .

The gross payment made by the planner to the firm is $(t + C)$. We assume that the planner can raise this amount only through a distortionary mechanism (excise taxes, e.g.) so that the social cost of one unit raised is $(1 + \lambda)$.⁶

⁵ Our analysis is almost unchanged if the good is a private good. It suffices to replace $\{S(q)\}$ by $\{\tilde{S}(q) \equiv S(q) + \lambda S'(q)q\}$ in the regulator's objective function, where λ is the shadow cost of public funds (see below). This change reflects the fact that, because of the cost of public funds, the firm's revenue is valuable. In an earlier draft (Laffont and Tirole 1984), we solved for the optimal regulatory policy for a private good-producing firm. We did not find much support for the average cost pricing rule in our setup (one way of formalizing average cost pricing corresponds to the following pricing requirement: $S'[q]q = s + C$). There exist finer ways to use cost and output (or price) observations than the average cost pricing rule to avoid monopoly pricing and to induce effort. Indeed, this rule imposes a rigid incentive and pricing structure. In particular, it is not sensible enough to the level of fixed cost, if any, and to the structure of information. See Freixas and Laffont (1985) for an analysis that compares only marginal cost pricing and average cost pricing in a framework with moral hazard.

⁶ For a discussion of this formalism and its (close) relationship to a weighted social welfare function, see Caillaud et al. (1985).

Consumer's welfare resulting from the activity of the firm is then $S(q) - (1 + \lambda)E(t + C)$.

If a utilitarian planner were able to observe the parameters of the cost function as well as the level of effort, he would solve

$$\max_{(q, e, t)} \{S(q) - (1 + \lambda)E(t + C) + U\} = \{S(q) - (1 + \lambda)[\psi(e) + (\beta - e)q] - \lambda U\} \quad (1)$$

subject to

$$U \geq 0. \quad (2)$$

The constraint (2), called the individual rationality constraint, says that the utility level of the firm's manager must be positive to obtain his participation. Procurement in which the principal is a private firm would lead to a different objective function of the type profit minus transfer. None of our results would be affected; the important feature is that the principal dislikes transfers.

The first-order conditions of problem (1) are

$$U = 0, \quad (3)$$

$$S'(q) = (1 + \lambda)(\beta - e), \quad (4)$$

$$\psi'(e) = q. \quad (5)$$

The individual rationality constraint is binding. The marginal utility of the commodity $S'(q)$ is equated to its social marginal cost $(1 + \lambda)(\beta - e)$. The marginal disutility of effort $\psi'(e)$ is equated to its marginal utility, that is, the marginal decrease in cost q .

We now make an assumption that ensures that the full-information solution exists and is unique.

ASSUMPTION 1. (i) $S'(0) > (1 + \lambda)[\bar{\beta} - \psi'^{-1}(0)]$. (ii) $\psi'(\underline{\beta}) > \bar{q}$, where \bar{q} is defined by $S'(\bar{q}) = 0$. (iii) $S''\psi'' + (1 + \lambda) < 0$.

Part i of assumption 1 says that the marginal surplus at no production is not too small. Part ii says that it is too costly (in terms of effort) to reduce marginal cost to zero, whatever the initial marginal cost. Part iii requires enough convexity in the full-information problem.

The task of this paper is to characterize and study the control mechanisms based on the observability of the output level q and the total cost C . The planner does not know β and cannot observe the level of effort e . He has a uniform prior on the range $[\underline{\beta}, \bar{\beta}]$ of β ; moreover, he knows the objective function of the firm (a more general distribution could be assumed; the uniform distribution saves notation and simplifies the technical analysis since it satisfies the monotonic hazard rate property, which prevents bunching).

III. The Optimal Incentive Scheme

The firm chooses output and effort. Once cost is realized and observed, the planner rewards the firm according to the two observables q and C . Equivalently (from the revelation principle), the planner can ask the firm to reveal its true productivity parameter, denoted $\hat{\beta}$. The reward then depends on the announcement β and the ex post cost, $t(\beta, C)$, and output is imposed by the planner, $q(\beta)$. As is well known, we can restrict ourselves to a truth-telling mechanism so that the firm's optimal strategy includes $\beta = \hat{\beta}$. Let $e(\beta)$ denote the optimal effort function for the truthful mechanism $\{q(\beta), t(\beta, C)\}$. We will characterize implementable allocations, that is, allocations that induce the firm to tell the truth such that the level of effort is (voluntarily) chosen by the firm. We will then treat the effort as a control variable for the regulator and check that one can find a transfer function $t(\beta, C)$ that leads the firm to choose the corresponding level of effort. Note that we restrict attention to deterministic mechanisms. It can be shown that, if ψ'' is nonnegative, random mechanisms are not optimal. Let $C(\beta) \equiv [\beta - e(\beta)]q(\beta)$ be the resulting expected cost and let $s(\beta) \equiv Et[\beta, C(\beta) + \epsilon]$ denote the expected net transfer (the expectation is taken with respect to the disturbance term ϵ).

A. The Firm's Optimization Problem

In equilibrium it must be the case that the firm's decision variables $[\beta = \hat{\beta}, e = e(\hat{\beta})]$ maximize $\{Et[\beta, (\hat{\beta} - e)q(\beta) + \epsilon] - \psi(e)\}$.

For the moment let us consider only a restricted class of possible deviations from the optimal strategy $[\hat{\beta}, e(\hat{\beta})]$. We will show that ruling out deviations in this class completely determines the output and effort functions. We will then exhibit a mechanism that implements this allocation; in particular, other types of deviations are not optimal for the firm when it faces this mechanism. Last, we will argue that this mechanism is optimal for any distribution of the disturbance.

Consider the following class of deviations from equilibrium $[\beta, e(\beta)]$ for firm $\hat{\beta}$: it announces β and makes effort $\bar{e}(\beta|\hat{\beta}) \equiv e(\beta) + \hat{\beta} - \beta$. The set of such deviations $[\beta, \bar{e}(\beta|\hat{\beta})]$ will be called the *concealment set* for firm $\hat{\beta}$. Note that, when there is no uncertainty, any deviations outside the concealment set can be detected by the planner. Note also that the concealment set includes $[\hat{\beta}, e(\hat{\beta})]$ and that, if firm $\hat{\beta}$ announces β and makes effort $\bar{e}(\beta|\hat{\beta})$, the cost distribution is the same as for firm β and therefore the expected transfer is $s(\beta)$. So ruling out deviations in the concealment set amounts to requiring that

$$\hat{\beta} \text{ maximizes } U(\beta|\hat{\beta}) \equiv s(\beta) - \psi[\bar{e}(\beta|\hat{\beta})]. \quad (6)$$

Appendix A shows that, if s and \bar{e} are such that (6) is satisfied, then these two functions as well as the effort function e are differentiable almost everywhere. So the first-order condition is

$$\dot{s}(\beta) - \psi'[\bar{e}(\beta|\hat{\beta})]\dot{\bar{e}}(\beta|\hat{\beta}) = 0 \quad (7)$$

almost everywhere (ae), where a dot denotes a derivative with respect to β . Using the definition of \bar{e} and truth telling, we obtain (we delete the qualifier "ae" from now on for notational simplicity)

$$\dot{s}(\beta) - \psi'[e(\beta)][\dot{e}(\beta) - 1] = 0. \quad (8)$$

The local second-order condition can be written using the first-order condition

$$\frac{\partial^2 U}{\partial \beta^2}(\beta|\hat{\beta}) \Big|_{\beta=\hat{\beta}} = - \frac{\partial^2 U}{\partial \beta \partial \hat{\beta}}(\beta|\hat{\beta}) \Big|_{\beta=\hat{\beta}} \leq 0 \quad \text{for any } \hat{\beta} \quad (9)$$

or

$$\dot{e}(\beta) \leq 1 \quad \text{for any } \beta. \quad (10)$$

Note that (10) can be given a simple interpretation: the firm's average cost is decreasing. Appendix B shows that, if the local second-order condition is satisfied, then the global second-order condition is also satisfied.

Last, letting $U(\beta) (= s(\beta) - \psi[e(\beta)])$ denote firm β 's (equilibrium) utility, we notice that the first-order condition (8) is equivalent to

$$\dot{U}(\beta) = -\psi'[e(\beta)]. \quad (11)$$

In other words, the increase in the firm's utility for a unit decrease in "intrinsic cost" of β is equal to the marginal disutility of effort (since the firm can reduce its effort by an amount equal to its increase in efficiency). We summarize these results in the following proposition.

PROPOSITION 1. Firm's Optimization Problem. If deviations in the firm's concealment set are not profitable, then the effort, transfer, and utility functions are differentiable almost everywhere. The first-order incentive compatibility constraint is given by (11). This necessary condition is also sufficient if the effort function satisfies (10).

We now turn to the planner's optimization problem. We will first assume that deviations in the concealment set are the only possible deviations, so that (10) and (11) are sufficient conditions for incentive compatibility. So we solve a subconstrained optimization problem for the principal. We later show that the solution makes deviations outside the concealment set also unprofitable for the firm. Thus we are justified to consider the simpler optimization problem.

B. The Planner's Problem

We assumed that the planner has uniform beliefs on $[\underline{\beta}, \bar{\beta}]$. His optimization problem is then (using the definitions of U and C and ignoring the uniform density):

$$(P) \quad \max E \int_{\underline{\beta}}^{\bar{\beta}} (S[q(\beta)] - (1 + \lambda)\{\psi[e(\beta)] + [\beta - e(\beta)]q(\beta) + \epsilon\} - \lambda U(\beta))d\beta \quad (12)$$

subject to

$$\dot{U}(\beta) = -\psi'[e(\beta)], \quad \text{ae} \quad (11)$$

$$\dot{e}(\beta) \leq 1, \quad \text{ae} \quad (10)$$

$$U(\beta) \geq 0, \quad \forall \beta. \quad (13)$$

Equation (13), the individual rationality constraint, says that the firm is willing to participate. As mentioned earlier, P is a subconstrained problem. We make it even less constrained by ignoring the second-order condition (10). Naturally we will have to check that the two types of ignored constraints are indeed satisfied by the solution of the less constrained problem.

Note that, from (11), U is a decreasing function of β , so (13) is satisfied if and only if $U(\bar{\beta}) \geq 0$. As social welfare decreases with U , we can also replace (13) by $U(\bar{\beta}) = 0$, so we study the simplified program

$$(P') \quad \max E \int_{\underline{\beta}}^{\bar{\beta}} (S[q(\beta)] - (1 + \lambda)\{\psi[e(\beta)] + [\beta - e(\beta)]q(\beta) + \epsilon\} - \lambda U(\beta))d\beta \quad (12)$$

subject to

$$\dot{U}(\beta) = -\psi'[e(\beta)], \quad \text{ae} \quad (11)$$

$$U(\bar{\beta}) = 0. \quad (14)$$

We treat P' as an optimal control problem with state variable U and control variables e and q . Appendix C studies this program and can be summarized by the following proposition.

PROPOSITION 2. The necessary conditions for an interior optimum of P' are

$$U(\bar{\beta}) = 0, \quad (15)$$

$$\dot{U}(\beta) = -\psi'[e(\beta)], \quad (11)$$

$$S'(q) = (1 + \lambda)(\beta - e), \quad (16)$$

$$\psi'(e) = q - \frac{\lambda}{1 + \lambda} (\beta - \underline{\beta})\psi''(e). \quad (17)$$

We make the following assumption.

ASSUMPTION 2. There exists a unique interior optimum of P' .

It is easily seen that assumption 2 is satisfied as long as either λ or $(\bar{\beta} - \beta)$ is "not too big." Appendix C shows that the first-order necessary conditions for P' are then sufficient.

C. Implementation

Under assumption 1, (16) and (17) determine the levels of output and effort $q^*(\beta)$ and $e^*(\beta)$; (11) and (15) then determine the firm's utility $U^*(\beta) = \int_{\beta}^{\bar{\beta}} \psi'[e^*(\delta)]d\delta$. The expected transfer is then given by $s^*(\beta) = U^*(\beta) + \psi[e^*(\beta)]$. As we mentioned earlier, (1) we still must check that the second-order condition (10) for the firm's maximization program is satisfied, so that the solution to P' is also the solution to P , and (2) we can find a transfer function $t(\beta, C)$ that implements the optimum of P (in particular, it should induce the right effort level and should not induce the firm to deviate outside the concealment set either).

First, to see whether the second-order condition (10) is satisfied, we must solve (16) and (17) and check that $\dot{e}^*(\beta) \leq 1$. Let us show that, under our assumptions, $\dot{e}^*(\beta) < 0$; from (16) and (17) we have

$$S''\dot{q} = (1 + \lambda) - (1 + \lambda)\dot{e} \quad (18)$$

and

$$\left[\psi'' + \frac{\lambda}{1 + \lambda} (\beta - \underline{\beta})\psi''' \right] \dot{e} - \dot{q} = - \frac{\lambda}{1 + \lambda} \psi'' \quad (19)$$

or

$$\left[\psi'' + \frac{1 + \lambda}{S''} + \frac{\lambda}{1 + \lambda} (\beta - \underline{\beta})\psi''' \right] \dot{e} = \frac{1 + \lambda}{S''} - \frac{\lambda}{1 + \lambda} \psi''. \quad (20)$$

If assumption 2 is satisfied, the local concavity of the Hamiltonian in the control variables implies that the coefficient of \dot{e} in (20) is positive and therefore that $\dot{e}^*(\beta) < 0$.

PROPOSITION 4. Under assumptions 1 and 2, the firm's effort increases with its efficiency. Therefore, the firm's second-order condition is satisfied, and the solution to P' is also the solution to the more constrained problem P .

Second, let us study the implementation problem. As before, let $\{e^*(\beta), q^*(\beta), U^*(\beta)\}$ denote the solution to P' , and let $s^*(\beta)$ and $C^*(\beta)$ denote the corresponding expected transfer and expected cost.

The answer to the implementability question is trivial in the case of no disturbance ($\epsilon \equiv 0$). As we noticed earlier, only deviations within the concealment set can then go undetected. So the solution to P is also the solution to the global problem. To implement it, it suffices for

the planner (i) to ask the firm to announce its characteristic β , (ii) to choose output $q^*(\beta)$, and (iii) to give transfer $s^*(\beta)$ if $C = C^*(\beta)$, and $-\infty$ otherwise. This simple "knife-edge" mechanism, however, is not robust to the introduction of any disturbance; if there is any noise, the probability of incurring an extreme penalty becomes positive and makes the firm unwilling to participate.

Let us now turn to the general case of cost disturbance. To solve the problem completely, we must find a transfer function $t(\beta, C)$ such that $\{\hat{\beta}, e^*(\hat{\beta})\}$ is optimal for the firm:

$$\{\hat{\beta}, e^*(\hat{\beta})\} \text{ maximizes } E\{t[\beta, (\hat{\beta} - e)q(\beta) + \epsilon] - \psi(e)\}$$

and

$$Et\{\hat{\beta}, [\hat{\beta} - e^*(\hat{\beta})]q^*(\hat{\beta}) + \epsilon\} = s(\hat{\beta}).$$

Imagine that the planner gives the firm the following transfer function (linear in observed cost):

$$t(\beta, C) = s^*(\beta) + K^*(\beta)[C^*(\beta) - C], \quad (21)$$

where

$$K^*(\beta) = \frac{\psi'[e^*(\beta)]}{q^*(\beta)}. \quad (22)$$

Remember that $C^*(\beta) \equiv [\beta - e^*(\beta)]q^*(\beta)$ and that $s^*(\beta) \equiv \psi[e^*(\beta)] + \int_{\hat{\beta}}^{\beta} \psi'[e^*(\delta)]d\delta$. Then firm β solves

$$\max_{\{\beta, e\}} E(K^*(\beta)\{C^*(\beta) - [(\hat{\beta} - e)q^*(\beta) + \epsilon]\} + s^*(\beta) - \psi(e)) \quad (23)$$

or

$$\max_{\{\beta, e\}} \{s^*(\beta) - \psi'[e^*(\beta)]\bar{e}^*(\beta|\hat{\beta}) + \psi'[e^*(\beta)]e - \psi(e)\}. \quad (24)$$

Optimization with respect to e clearly leads to

$$e = e^*(\beta). \quad (25)$$

When (25) and (8) are used, the optimization with respect to β gives

$$\beta = \hat{\beta}. \quad (26)$$

So this linear incentive scheme implements the optimal allocation if the firm's second-order condition for (24) is satisfied. Straightforward computations show that this condition boils down to

$$\dot{e}^*(\beta) \leq 0. \quad (27)$$

We draw the following conclusion.

PROPOSITION 5. Under assumptions 1 and 2, the optimal allocation

can be implemented by an incentive scheme that is linear in cost: $t(\beta, C) = s^*(\beta) + K^*(\beta)[C^*(\beta) - C]$.

The second-order condition (27) is more stringent than (10). This deserves some comment. Equation (27) corresponds to one way of implementing the optimal solution, which requires the transfer to be linear in cost. If (27) is satisfied (as is the case under our assumptions), then the linear scheme is a perfectly legitimate way of implementing the solution. If (27) is not satisfied, the linear scheme is not optimal since it imposes too stringent a second-order condition. This point is best explained in the no-uncertainty case ($\epsilon \equiv 0$). As we have seen, the knife-edge mechanism is an alternative way to implement the optimal allocation. This mechanism gives the most lenient second-order condition, (8), since its extreme penalties for cost overruns restrict the set of possible deviations to the concealment set. To the contrary, the linear scheme defined by (21) allows more deviations, and its linearity in cost restricts the possible punishments for deviations out of the concealment set. Thus the second-order condition is unsurprisingly more stringent.

The linear scheme implements the optimal allocation. Furthermore, it has a very appealing property. Notice that the optimal allocation is independent of the distribution of cost uncertainty. Intuitively, the linear scheme is the only scheme that implements the optimal allocation whatever the distribution of cost uncertainty.

PROPOSITION 6. Under assumptions 1 and 2, the linear scheme $\{t(\beta, C) = s^*(\beta) + K^*(\beta)[C^*(\beta) - C]\}$ implements the optimal allocation for any cost uncertainty (with zero mean). It is the only scheme having this property.

The proof of the last part of proposition 6 can be found in Appendix D.

D. Summary of the Technical Analysis

We studied the simple program P' , which maximizes expected social welfare under the individual rationality constraint for the least efficient firm and the first-order incentive constraint. We then looked at the firm's and the planner's second-order condition, and we showed how one can implement the optimal allocation. The linear scheme is the optimal scheme (for any distribution of the cost disturbance).

For the study of the optimal scheme, we will sometimes need a further assumption.

ASSUMPTION 3. $\{\psi''/\psi'\}$ is nonincreasing.

Assumption 3 puts a (positive) upper bound on the third derivative of the cost function. An example of a cost function satisfying assumption 3 is the quadratic cost function $\psi(e) = e^2/2$.

REMARK. Equations 10 (and a fortiori assumption 2) and (16) imply that the output is nonincreasing in β . This remark will prove useful in the interpretation of the optimal scheme.

E. The Optimal Scheme under Cost Unobservability (Baron-Myerson)

In this subsection we want to compare the solution derived in Section IIIB with the inferior solution that would obtain if the planner were unable to observe cost. The latter situation has been extensively studied in the literature (see, e.g., Baron and Myerson 1982; Guesnerie and Laffont 1984).

In this subsection only, $s(\beta)$ will denote the *gross* transfer to the firm when it announces characteristic β . A net transfer does not make sense since C is not observed.

We derive the Baron-Myerson results for our model. For ease of exposition, we will ignore second-order conditions in the presentation. The firm's program is

$$U(\hat{\beta}) = \max_{\{\beta, e\}} \{s(\beta) - (\hat{\beta} - e)q(\beta) - \psi(e)\}. \quad (28)$$

The firm's first-order conditions are

$$\dot{s}(\beta) = (\beta - e)\dot{q}(\beta), \quad (29)$$

$$\psi'[e(\beta)] = q(\beta). \quad (30)$$

Equation (30) shows that effort is socially optimal conditional on output. This is intuitive since the cost, which is unobservable by the planner, is fully borne by the firm. Note also that the incentive constraints imply that

$$\dot{U}(\beta) = -q(\beta). \quad (31)$$

So the planner's subconstrained program is

$$\max \int_{\underline{\beta}}^{\bar{\beta}} (S[q(\beta)] - (1 + \lambda)\{\psi[e(\beta)] + [\beta - e(\beta)]q(\beta)\} - \lambda U(\beta))d\beta \quad (32)$$

subject to

$$\dot{U}(\beta) = -q(\beta), \quad (33)$$

$$U(\bar{\beta}) = 0. \quad (34)$$

It is easily shown that the necessary conditions for this program are

$$S'(q) = (1 + \lambda)(\beta - e) + \lambda(\beta - \underline{\beta}), \quad (35)$$

$$\psi'(e) = q. \quad (36)$$

The role of cost observability will be studied in Section IVB.

IV. The Optimal Allocation and Incentive Scheme

This section draws the economic implications of the previous technical analysis (assuming assumptions 1 and 2).

A. Comparison with the Full-Information Allocation

PROPOSITION 7. The asymmetry in information implies for all β (except β) a lower output and a lower effort.

Proof. Compare $\{(4), (5)\}$ and $\{(16), (17)\}$ and apply assumption 1. Q.E.D.

The intuition behind proposition 7 is simple. Under moral hazard, the regulator cannot reimburse the totality of the firm's cost. However, it does not want to adopt a fixed-price contract (which it would be forced to do under cost unobservability). Under such a contract no moral hazard problem arises. But the firm, bearing the full cost, has a tendency to understate its efficiency to be allocated a low output and thus incur a low cost (as shown by Baron and Myerson [1982]). This underproduction can be avoided if the firm is made the residual claimant for social welfare, that is, if it is rewarded $\{S(q)/(1 + \lambda)\}$ (up to a constant). But making the firm the residual claimant is too costly under incomplete information about the firm's productivity. Reimbursing part of the firm's cost helps alleviate this issue by making the firm less concerned about cost and therefore less conservative in its output decision. Indeed, if there were no moral hazard, the optimal contract would be cost-plus. Clearly the trade-off between inducing revelation (cost-plus contract) and inducing effort (fixed-price contract) results in an "incentive contract" (partial sharing of cost), as shown by the optimal incentive scheme.

Given that the firm's cost is partially reimbursed, effort is suboptimal. Hence marginal cost is excessive, and output is therefore suboptimal, as shown by proposition 7.

B. The Role of Cost Observability

The optimal scheme under cost unobservability has been studied by Baron and Myerson (1982). For our model, the comparison between the two cases is given by $\{(16), (17)\}$ and $\{(35), (36)\}$. As explained above, Baron and Myerson's fixed-price contract implies no effort distortion for a given output contrary to the optimal "incentive contract" derived for cost observability. This effort distortion is more than offset from a welfare point of view by the lower price distortion $\{S'(q) - (1 + \lambda)(\beta - e)\}$ for the incentive contract. Indeed, the fixed-price contract, which ignores the cost information, is also available to

the planner under cost observability but is not optimal because some cost reimbursement eases revelation of the technological information.

C. Efficiency and the Choice of Output and Effort

In Section III we saw that the optimal levels of effort, output, and expected costs were all decreasing with the marginal cost parameter β . This monotonicity property is not surprising for the output and expected cost variables; it is explained below for the effort variable.

D. The Optimal Incentive Scheme

Let us rephrase proposition 6 by assuming (more realistically) that the net transfer depends on output and observed cost (from the revelation principle the two approaches are equivalent). Since output is a monotonic function of β (see Sec. IIID), we have the following proposition.

PROPOSITION 8. For any distribution of the cost uncertainty, the optimal allocation is implemented through a linear scheme

$$\bar{u}(q, C) = \bar{s}(q) + \bar{K}(q)[\bar{C}(q) - C],$$

where $\bar{C}(q)$ is the optimal expected cost given q and $0 < \bar{K}(q) \leq 1$. Furthermore, (i) $\bar{s}(q)$ is an increasing function of q , (ii) $\bar{K}(q)$ is an increasing function of q if assumption 3 is satisfied, and (iii) \bar{K} converges to one (fixed-price contract) when uncertainty becomes small— $(\bar{\beta} - \beta) \rightarrow 0$.

Proof. The functions $\{\bar{s}, \bar{K}, \bar{C}\}$ are derived from $\{s^*, K^*, C^*\}$ by substituting q for β (since q is monotonic in β). Monotonicity of \bar{s} results from its definition and the second-order condition. Part iii results from the definition of \bar{K} and (17). Last, differentiating (22) and using (17) gives

$$\dot{\bar{K}} \alpha \frac{q - \psi'}{\psi''} [(\psi'')^2 - \psi' \psi'''] \dot{e} - \frac{\lambda}{1 + \lambda} \psi' \psi''.$$

Assumptions 2 and 3 and (17) then imply that $\dot{\bar{K}} < 0$. Q.E.D.

Proposition 8 has several important implications for regulation. First, in the context of our model, the optimal allocation can be implemented by a particularly simple incentive scheme. Furthermore, the knowledge of the distribution of the cost disturbance around zero is not required to build this scheme. The contract is an incentive contract. It can be decomposed into a fixed-price contract $\bar{s}(q)$ and a partial cost reimbursement. After agreeing on an output, the planner gives a first reward $\bar{s}(q)$, which increases with output. Then, after

observing the final cost, he gives a penalty or a bonus that is proportional to cost overruns.

Second, the coefficient of proportionality depends on the scale of the project. Indeed, under assumption 3, the fraction $(1 - \bar{K})$ of costs that are reimbursed decreases with output. There are two reasons for this. The first reason is associated with a scale effect. We know that low-cost firms produce more. For those firms marginal cost reduction is more valuable, so effort should be particularly encouraged. This suggests that \bar{K} ought to be higher. However, the marginal incentive to exert effort is not \bar{K} , but $\bar{K}q$, and we know that q is already higher for a low-cost firm. So to conclude that \bar{K} is also higher, we made the sufficient assumption 3. The second reason is associated with the limitation of the firm's rent. We know that the firm's utility and, therefore, transfer are obtained by imposing the individual rationality constraint $U(\bar{\beta}) = 0$ and integrating backward the incentive compatibility constraint $\dot{U}(\beta) = -\psi'[e(\beta)]$. A high level of effort for an inefficient firm (β close to $\bar{\beta}$) is thus reflected in a higher utility for almost all types of firms. So effort should be encouraged more for more efficient firms (indeed, at the optimum, there is no effort distortion for $\beta = \bar{\beta}$). In particular, for a fixed-size project, $(1 - \bar{K})$ always decreases with β (regardless of assumption 3).

Third, when the uncertainty becomes small, reimbursing the cost to induce efficient revelation of information becomes valueless. Only the moral hazard problem remains relevant, and, under risk neutrality, the contract converges to a fixed-price contract. This phenomenon to some extent was observed by Ponssard and de Pouvourville (1982, p. 55) in the dynamic evolution of contracts in the French weapons industry. They observed that, as a project evolves over time, the contract resembles more and more a fixed-price contract. This may be explained by the fact that the government acquires information about the firm's cost function.⁷

An even more familiar way of interpreting proposition 6 uses the fact that the expected average cost $c^*(\beta)$ is increasing in β (see Sec. IIIA). Imagine that the firm, instead of announcing its efficiency parameter β , announces an expected average cost c^a ; $\bar{q}(c^a)$ units are then ordered and the firm is rewarded ex post according to $\bar{t}(c^a, c)$, where c is the ex post average cost. We have the following proposition.

PROPOSITION 9. The optimal allocation can be implemented by asking the firm to announce an expected average cost c^a and by making

⁷ Of course, our model here is a static one. Ponssard and de Pouvourville's observations are vindicated by our model if both the planner and the firm take a myopic perspective in a dynamic context. When the parties take a dynamic perspective, the study should be completed by a dynamic analysis of the corresponding ratchet effect (see Laffont and Tirole 1985).

the transfer depend on the expected and realized average costs: $\bar{t}(c^a, c) = \bar{s}(c^a) + \bar{K}(c^a)(c^a - c)$; $\bar{s}(c^a)$, $\bar{q}(c^a)$, and (under assumption 3) $\bar{K}(c^a)$ are decreasing functions, and $0 < \bar{K} \leq 1$.

The ex ante reward (\bar{s}) and the slope of the ex post bonus scheme (\bar{K}) decrease with the announced cost. We can relate this result to evidence on actual incentive schemes. Contracts usually specify a higher transfer if the firm is willing to increase its share of cost overruns or underruns (see, e.g., Scherer 1964, p. 260). This practice is given a normative justification by proposition 9: the latter shows that the transfer (\bar{s}) and the coefficient of cost sharing (\bar{K}) are positively correlated.

E. Influence of Demand on the Optimal Contract

Let us briefly study how the sharing coefficient \bar{K} varies with the demand function. Let us posit that the consumer's surplus depends on a parameter Θ : $S(q, \Theta)$. A way of formalizing the idea that the output becomes (marginally) more valuable when Θ increases is to assume that $S_{q\Theta} > 0$. In this case demand grows with the parameter Θ (e.g., for linear demand, Θ can represent the intercept or minus the slope of the demand curve).

PROPOSITION 10. Under assumption 3, the optimal contract resembles more a fixed-price contract when the demand for output increases.

Proof. Differentiate (16) and (17) and use assumptions 2 and 3 to obtain $\partial \bar{K} / \partial \Theta \propto S_{q\Theta}$. Q.E.D.

The intuition behind proposition 10 is the following. A higher demand leads to higher output. So cost reduction through effort becomes more valuable. It then makes sense to have the firm bear a higher fraction of its cost overruns (not surprisingly, use is made of assumption 3, which also plays a role in showing that \bar{K} must grow with q).

F. Contracting on Quality

Suppose that the scale variable to be determined is the quality of the output rather than its level (which we can take to be one). The model and its conclusions are unchanged if q denotes a quality parameter instead of a quantity as long as quality is observable ex post. In particular, under asymmetric information, (i) there is underprovision of effort and quality, and (ii) the sharing rate of the optimal linear scheme is positively correlated with quality.

Proposition 10, in the quality interpretation, tells us that the more concerned about quality the regulator is, the more the optimal con-

tract resembles a fixed-price contract because marginal cost reductions must be encouraged more when higher qualities are chosen. Alternative models may lead to the opposite conclusion. Imagine, for instance, that quality is observable but not verifiable so that the contract cannot be made contingent on quality. The firm must then trade off immediate cost savings (low quality) and reputation. A way to encourage the firm to choose a higher quality is then to share a higher fraction of cost. Similarly, even if no reputation is involved, the possibility of bankruptcy may also move the optimal contract toward a cost-plus contract. These models may fit better the casual observation (e.g., for defense and building contracts) that contracts resembling cost-plus ones are often used when the level of quality matters much to the planner.

G. Indivisible Project

Let us illustrate graphically the solution in the simple case of an indivisible project ($q = 1$).

Suppose that the project is worth undertaking for any value of $\beta \in [\underline{\beta}, \bar{\beta}]$. When $\epsilon \equiv 0$, the utility function of the manager can be written as $U(s, C, \beta) = s - \psi(\beta - C)$. The problem can then be viewed as a classic adverse selection problem with two observables, s and C . The Spence-Mirrlees condition

$$\frac{\partial}{\partial \beta} \left(\frac{\partial U / \partial C}{\partial U / \partial s} \right) = \psi'' > 0$$

is satisfied. The optimal contract can be represented by a nonlinear price $s(C)$ (AB in fig. 1), and each manager β chooses his best contract on this curve. Under assumptions 1 and 2, AB (because it is convex) can be replaced without modifying the equilibrium by the family of straight lines tangent to AB , which corresponds to our family of linear incentive contracts with varying coefficients. From risk neutrality these contracts lead to the same levels of effort when costs are random.

When q is a continuous variable, the same interpretation holds when C is replaced by average cost.

V. Choice of Technology and Rate-of-Return Regulations

The same approach can be applied to the case in which the firm has, ex ante, the possibility of choosing between various technologies that involve different splittings of cost between fixed costs and marginal costs. Let

$$C = (\hat{\beta} + \beta_1 - e)q + \alpha(\beta_1) + \epsilon,$$

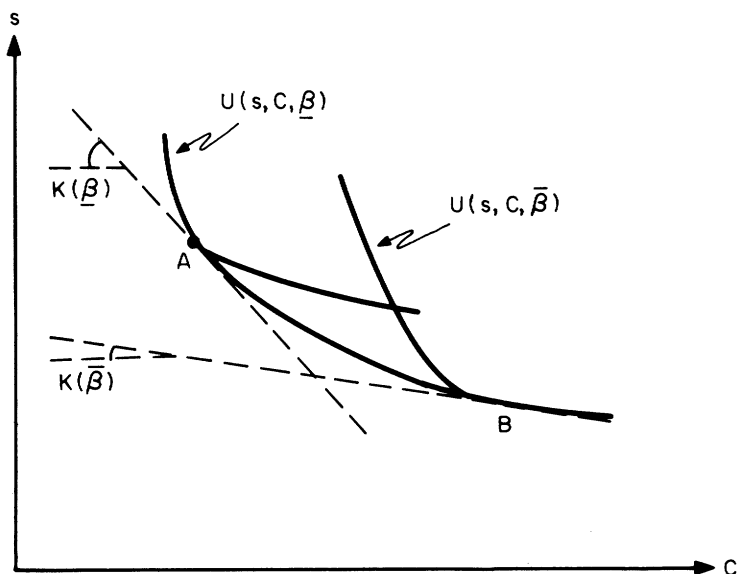


FIG. 1

where $\alpha(\beta_1)$ is the firm's fixed costs and ϵ is a cost disturbance with zero mean; $\underline{\beta}$ is given. By increasing β_1 , the firm decreases its fixed cost— $\alpha' < 0$, $\alpha'' > 0$, $\alpha'(0) = -\infty$ —and increases its variable cost. In a first step neither the choice of β_1 , nor the level of effort, nor the particular value of $\underline{\beta}$ can be observed by the regulator.

Let us briefly argue that technological choices between fixed and variable costs and possibly their unobservability by the regulator may be relevant features of real-world procurement situations. For example, a power company may choose between high-fixed-cost technologies (e.g., nuclear plants) and high-variable-cost ones (e.g., coal). Similarly increasing overhead within a plant (supervisors, foremen, engineers) increases the fixed costs while reducing variable costs (associated with mistakes, delays, low effort, etc.). The latter example suggests that it is sometimes fairly hard for public accountants to split the total cost they observe into fixed and variable costs (for an example of how a firm can manipulate this accounting procedure, see Peck and Scherer [1962, p. 518]).

The analysis of Section III is hardly modified by the introduction of the extra choice variable β_1 . It is easily seen that the optimal allocation must satisfy

$$S'(q) = (1 + \lambda)(\beta + \beta_1 - e), \quad (37)$$

$$\psi'(e) = q - \frac{\lambda}{1 + \lambda} (\beta - \underline{\beta})\psi''(e), \quad (38)$$

$$\alpha'(\beta_1) = -q. \quad (39)$$

From (37) and (38) and assuming a unique full-information allocation (analogue of assumption 1), we can easily see that q must be smaller than at the optimum under perfect information. Then from (39) we can conclude that there is a bias toward less fixed costs due to imperfect information.⁸ The intuition behind this result is simple: imperfect information about β leads to suboptimal quantities. Thus marginal cost reductions (through β_1) are effective on a lower number of units of output than in the perfect information case. Therefore, there is an incentive to keep marginal cost high and fixed costs low. Let us notice, however, that the firm makes the right technological decision given its output because, in the optimal contract, part of the cost is borne by the firm.

Let us now observe that this bias toward low fixed costs is also a local welfare property around the incomplete information optimum. To this purpose, let us examine how the social welfare changes with β_1 in the neighborhood of the second-best solution when investment is not observable by the regulator:

$$\begin{aligned} \frac{d}{d\beta_1} \{S'(q) - (1 + \lambda)[EC + \psi(e)] - \lambda U\} \\ = S'(q)\dot{q} - (1 + \lambda)[- \dot{q}e + (\beta - e)\dot{q} + \psi'(e)\dot{e}] \quad (40) \\ = \frac{\lambda(\beta - \bar{\beta})\psi''(e)}{q} \dot{e}, \end{aligned}$$

where (37), (38), and (39) are used in two ways: the partial derivatives of C and U with respect to β_1 are zero.

So, under assumption 2,

$$\frac{d}{d\beta_1} \{S(q) - (1 + \lambda)[C + \psi(e)] - \lambda U\} \leq 0. \quad (41)$$

This means that welfare would increase if the firm decreased β_1 slightly below the cost minimization level. In other words, the regulator would like to force the firm to *overinvest* a bit in *fixed costs*. Let us now imagine that the level of investment $\alpha(\beta_1)$ (and therefore β_1) is *observable* by the regulator. One may wonder whether a rate-of-return regulation associated with a transfer function $t(q, C)$ cannot implement the new second-best allocation (i.e., under investment observability). By rate-of-return regulation we mean a constraint on, for example, transfers to the firm (t) per unit of capital invested (α) (where the regulated rate of return could depend on the firm's output). It is

⁸ This underinvestment property is closely related to that in Tirole (1986), who gives a general result for incomplete contracts (the result obtained here assumes a complete contract).

well known that rate-of-return regulation induces an upward bias in capital accumulation. So, a priori, such a rule may improve the previous allocation while letting the firm choose its investment.

It turns out that, in our model, the optimal allocation under investment observability is the same as without observability. This result is contingent on the separable form we assumed for the cost function. Let us give some intuition for it. The firm, when free to choose its investment, has a common incentive with the regulator to minimize cost (see eq. [39]). So incentives may differ only if the choice of β_1 has an influence on the incentive compatibility constraint. But the latter, $\dot{U} = -\psi'(e)$, is unaffected by the observability of β_1 : knowing $q(\beta)$ and knowing that the firm minimizes cost with respect to β_1 , the planner infers the term $[\beta_1 q(\beta) + \alpha(\beta_1)]$, and therefore the "concealment set" (defined in Sec. III) $\{\beta - e = \beta - e(\beta)\}$ is not affected by investment observability. So for our specification there is nothing more that the regulator can do if he happens to be able to observe investment.⁹ In particular, imposing a (binding) rate-of-return regulation would be detrimental since it would destroy cost minimization.¹⁰

VI. Risk Aversion

Let us briefly explore the consequences of risk aversion on the firm's behavior and on the incentive scheme. Let us assume that the manager has the following expected utility function (which can be justified by an approximation argument):

$$U = Et - \gamma \text{ var } t - \psi(e). \quad (42)$$

The cost function is $C = (\beta - e)q + \epsilon$, where ϵ is a random variable with mean zero and variance σ^2 .

The analysis under risk aversion becomes complex. We will not derive the optimal mechanism but simply study how the coefficient K^* of the linear mechanism $t(\beta, C) = s^*(\beta) + K^*(\beta)[C^*(\beta) - C]$ (which is optimal under risk neutrality) must be changed because of risk aversion. The derivation of the optimal *linear* incentive scheme for small γ (see Laffont and Tirole 1984) leads to the following (unsurprising) conclusion.

PROPOSITION 11. Assume that assumption 1 is satisfied, that $\psi''' \geq 0$, and that the coefficient of risk aversion γ in (42) is "small." The

⁹ One may then wonder why a decrease in β_1 can increase social welfare. The answer is that the decrease suggested by eq. (41) is hypothetical in that it does not take into account the change in incentives required to bring it forth.

¹⁰ For more complex cost functions, we conjecture that a rate-of-return regulation may increase or decrease welfare relative to the now suboptimal scheme $t(q, C)$, depending on the effect of the investment on the concealment set.

fraction of cost that is reimbursed in the best linear scheme increases with the coefficient of risk aversion.

VII. Related Work and Conclusions

The desirability of an incentive contract is reminiscent of the moral hazard literature.¹¹ The main difference between this literature and our work is that we added private information at the contracting date.¹² This accounts for the fact that, even under risk neutrality, incentive contracts (of a linear form in our model) are desirable. More important, it allowed us to show how the sharing coefficient must vary with the fixed fee or with the firm's intrinsic efficiency.

The paper most related to our work is Baron and Besanko's (1984). They consider a procurement situation analogous to ours. The planner does not know the marginal cost β . Ex post, he observes a variable correlated with the firm's cost, so there is an observation error. The firm's only decision variable is the announcement β , so there is no moral hazard. The authors assume that the planner is constrained to impose ex post a penalty in some interval $[0, \bar{N}]$. They show that it is optimal to impose the penalty \bar{N} if the observed cost is "low" and zero otherwise (what "low" means depends on the announcement β). They also show that, under some conditions, the price (or quantity) policy $q(\beta)$ is independent of the possibility of observing cost. In other words, ex post auditing is only a way to reduce the transfer to the firm (separability property).

The idea behind these results is the following. If no cost observation is available, the model boils down to Baron and Myerson's (1982), and the problem is simply to elicit the firm's marginal cost. As the cost is fully borne by the firm, a low-marginal-cost firm will tend to announce a high marginal cost to be allocated a low quantity to produce. Costly transfers are then required to prevent the firm from lying. Introducing cost observation does not affect the firm's real cost since there is no moral hazard, but it gives some information about the firm's marginal cost. To further prevent the firm from announcing high marginal costs, one puts penalties on low-cost observation (if auditing is costly and therefore is not done systematically, this statement must be qualified by the fact that high announced costs are

¹¹ See, e.g., Mirrlees (1974, 1976), Harris and Raviv (1979), Holmström (1979), Shavell (1979), and Grossman and Hart (1983). We should also mention the literature on the use of ex post observations in insurance markets and optimal taxation (e.g., Mirrlees 1974, 1976; Polinsky and Shavell 1979; Landsberger and Chazan 1983). There it is shown that penalties based on, e.g., the occurrence of accident can help reduce moral hazard.

¹² Baron (1982) studies a model of the demand for investment banking advising under adverse selection, moral hazard, uncertainty, and risk neutrality.

more likely to be audited). The Baron-Besanko story might be that of an agency that applies for a yearly budget. If it has not used up its budget at the end of the year, it is punished for its excessive greed in the following year by being allocated a lower budget.¹³

Our conclusions differ strikingly from Baron and Besanko's. First, under moral hazard, the planner cannot reward high costs. Otherwise the firm could always manage to increase expenditures. Indeed, we find that only a fraction of costs is reimbursed. Second, our pricing policy relies heavily on the possibility of observing the firm's cost (see Sec. IV). Cost observability reduces the distortion between price and (social) marginal cost.

Let us summarize our conclusions. (i) We gave a complete characterization of the firm's and planner's problems. (ii) Under moral hazard and total cost observability, the firm's effort is suboptimal, and its price is too high compared with perfect information. (iii) The planner can use a reward function that is linear in cost. The same linear function can be used for any distribution of the cost disturbance. (iv) The fraction of realized cost that is reimbursed to the firm is not a constant but decreases with the firm's output or increases with the firm's announced cost. This results from the fact that the different types of firm self-select when signing the contract with the regulator. The most efficient firm chooses a fixed-price contract. The less efficient firms opt for an incentive contract. The regulator agrees to reimburse a higher fraction of costs, the less efficient the firm is. Furthermore, the fixed transfer increases with the fraction of total cost that the firm is willing to share. (v) The optimal contract moves toward a fixed-price contract when demand increases. (vi) Cost observability improves welfare. It has a tendency to distort the effort decision, but it allows more control over the pricing policy. (vii) The linear reward function deals with increasing risk aversion in the best way by increasing the fraction of reimbursed costs. (viii) If the firm makes an unobservable technological choice between fixed and variable costs, there is a bias toward low fixed costs and high variable costs. A rate-of-return regulation may not improve welfare in spite of insufficient capital accumulation.

Conclusions iii and iv are the main conclusions of the paper. It is well known that optimal incentive contracts under moral hazard and risk aversion are complex and nonrobust.¹⁴ Moving the focus of atten-

¹³ Still it is very hard to avoid moral hazard. See the well-known stories of trucks driving around the barracks yard at Christmas time to use up their gas endowment.

¹⁴ An interesting exception to the setup is considered by Holmström and Milgrom (1984), who derive a linear optimal scheme. The linearity there is obtained for a very different reason from the one developed in this paper. The conclusion is reached in a pure moral hazard context and results from the richness in the action space and an exponential form for the agent's utility function.

tion from risk aversion to adverse selection enables us to obtain optimal contracts that are linear. And, precisely because they are linear, these contracts are robust to changes in the distributions of the cost accounting and forecast errors. We should emphasize that the irrelevance of noise is not a trivial consequence of risk neutrality. The noise garbles the auditing process, and only after having shown that the optimal allocation can be implemented by offering a menu of linear contracts to the firm can one conclude that the optimal scheme is robust.¹⁵ We should also mention that the possibility of implementing the optimal solution through a menu of linear contracts holds under much more general circumstances than those considered in this paper, for instance, for more general cost functions,¹⁶ although the characterization of the complete set of environments such that this is indeed possible is out of the scope of this paper.

Appendix A

Differentiability of the Effort, Transfer, and Utility Functions

LEMMA 1. $\beta < \hat{\beta} \rightarrow \bar{e}(\beta|\hat{\beta}) \geq \bar{e}(\hat{\beta}|\hat{\beta})$.

Lemma 1 says that a firm with cost $\hat{\beta}$ must make a higher effort when it announces a cost lower than the true one.

Proof. From the incentive compatibility constraints, we know that

$$\begin{aligned} s(\hat{\beta}) - \psi[\bar{e}(\hat{\beta}|\hat{\beta})] &\geq s(\beta) - \psi[\bar{e}(\beta|\hat{\beta})], \\ s(\beta) - \psi[\bar{e}(\beta|\beta)] &\geq s(\hat{\beta}) - \psi[\bar{e}(\hat{\beta}|\beta)]. \end{aligned} \quad (A1)$$

¹⁵ This point is most easily made in the indivisible project case. Suppose that, contrary to the case of a uniform distribution, the distribution of β does not satisfy the monotone hazard rate property. For a quadratic ψ , e.g., $e(\beta)$ must increase over some interval (which, in passing, may lead to the reintroduction of the firm's second-order condition $\bar{C} = 1 - \bar{e} \geq 0$). The optimal transfer function $t(C)$ in the absence of uncertainty is no longer concave and therefore is not the upper envelope of its tangents. The implementation through a menu of linear contracts is not feasible, and the noise in cost observation may matter.

¹⁶ For example, consider the indivisible project case ($q = 1$). Let $E(C, \beta)$ denote the effort level required to reach cost C when the firm has efficiency β . Make the reasonable assumptions that

$$\frac{\partial E}{\partial \beta} > 0; \frac{\partial E}{\partial C} < 0; \frac{\partial^2 E}{\partial C \partial \beta} \leq 0; \frac{\partial^2 E}{\partial C^2} \leq 0.$$

Then the firm's second-order condition requires that the firm's cost increase with β . If the regulator's Hamiltonian has a unique solution in the control variables (assumption 2), the equilibrium effort decreases with β (in particular, the firm's second-order condition is satisfied). It is then possible to show that the optimal transfer function $t(C)$ is convex in C , so that it can be represented as the envelope of its tangents. In other words, the optimal allocation can be implemented through a menu of linear contracts and therefore is immune to the introduction of noise in the cost observation.

Adding these two inequalities, we obtain

$$\psi[\bar{e}(\beta|\hat{\beta})] - \psi[\bar{e}(\beta|\beta)] \geq \psi[\bar{e}(\hat{\beta}|\hat{\beta})] - \psi[\bar{e}(\hat{\beta}|\beta)]. \quad (\text{A2})$$

Imagine that

$$\bar{e}(\hat{\beta}|\hat{\beta}) > \bar{e}(\beta|\hat{\beta}) \quad (\text{A3})$$

(which would contradict the lemma). We also know that, by definition of \bar{e} ,

$$\bar{e}(\beta|\hat{\beta}) - \bar{e}(\beta|\beta) = \hat{\beta} - \beta = \bar{e}(\hat{\beta}|\hat{\beta}) - \bar{e}(\hat{\beta}|\beta) > 0. \quad (\text{A4})$$

Last, (A3), (A4), and the strict convexity of ψ contradict (A2). Q.E.D.

LEMMA 2. $\bar{e}(\beta|\hat{\beta})$ is nonincreasing in β .

Proof. Fix $\beta > \beta'$ and define $\Delta(\hat{\beta}) \equiv \bar{e}(\beta'|\hat{\beta}) - \bar{e}(\beta|\hat{\beta})$. We have $\Delta(\hat{\beta}) = [e(\beta') - \beta'] - [e(\beta) - \beta]$. Thus $\Delta(\hat{\beta})$ does not depend on $\hat{\beta}$. But from the previous lemma $\Delta(\beta') \leq 0$. Thus, for all $\hat{\beta}$, $\Delta(\hat{\beta}) \leq 0$. Q.E.D.

Lemma 2 implies that $\bar{e}(\beta|\hat{\beta})$ is almost everywhere differentiable in β . So is $e(\beta) = \bar{e}(\beta|\hat{\beta}) + (\beta - \hat{\beta})$.

LEMMA 3. $U(\beta|\hat{\beta})$, as a function of β , is nondecreasing on $[\underline{\beta}, \hat{\beta}]$ and nonincreasing on $[\hat{\beta}, \bar{\beta}]$.

Proof. Let us first show monotonicity on $[\underline{\beta}, \hat{\beta}]$. Assume that $\beta < \beta' < \hat{\beta}$ and $U(\beta|\hat{\beta}) > U(\beta'|\hat{\beta})$. Thus

$$s(\beta) - \psi[\bar{e}(\beta|\hat{\beta})] > s(\beta') - \psi[\bar{e}(\beta'|\hat{\beta})]. \quad (\text{A5})$$

On the other hand, we know that a firm with cost β' prefers to announce β' rather than announce β . Thus

$$s(\beta') - \psi[\bar{e}(\beta'|\beta')] \geq s(\beta) - \psi[\bar{e}(\beta|\beta')]. \quad (\text{A6})$$

Adding (A5) and (A6), we get

$$\psi[\bar{e}(\beta|\beta')] - \psi[\bar{e}(\beta'|\beta')] > \psi[\bar{e}(\beta|\hat{\beta})] - \psi[\bar{e}(\beta'|\hat{\beta})]. \quad (\text{A7})$$

As in the proof of lemma 2 and using lemma 1, we get

$$\bar{e}(\beta|\beta') - \bar{e}(\beta'|\beta') = \bar{e}(\beta|\hat{\beta}) - \bar{e}(\beta'|\hat{\beta}) > 0, \quad (\text{A8})$$

and from the definition of \bar{e} ,

$$\bar{e}(\beta'|\beta') < \bar{e}(\beta'|\hat{\beta}). \quad (\text{A9})$$

Equations (A8) and (A9) and the convexity of ψ contradict (A7).

Monotonicity on $[\hat{\beta}, \bar{\beta}]$ is proved in the same way (using the incentive compatibility constraint for β this time). Q.E.D.

LEMMA 4. $s(\beta)$ is nonincreasing.

Proof. By definition

$$s(\beta) = U(\beta|\beta) + \psi[\bar{e}(\beta|\beta)]. \quad (\text{A10})$$

From lemmas 2 and 3, the two functions on the right-hand side are nonincreasing. So is s . Q.E.D.

Lemmas 2 and 4 imply that the functions \bar{e} (and therefore e) and s are almost everywhere differentiable. Hence $U(\beta) = s(\beta) - \psi[e(\beta)]$ is also almost everywhere differentiable. This completes the proof of the first part of proposition 1.

Appendix B

The Local Second-Order Condition Implies the Global One

LEMMA 5. If $\partial U/\partial \beta$ is (strictly) monotonic in $\hat{\beta}$, then the local second-order condition implies the global one.

Proof of Lemma 5. The local second-order condition implies that announcing the truth $\hat{\beta}$ gives a local maximum for the firm. Is there another $\beta \neq \hat{\beta}$ that satisfies the first-order condition? That is, does there exist $\beta \neq \hat{\beta}$ such that

$$\frac{\partial U}{\partial \beta}(\beta, \hat{\beta}) = \frac{\partial U}{\partial \beta}(\hat{\beta}, \hat{\beta}) = 0?$$

This would imply that

$$\frac{\partial U}{\partial \beta}(\beta, \hat{\beta}) = \frac{\partial U}{\partial \beta}(\beta, \beta) = 0.$$

But this is inconsistent with the (strict) monotonicity of $\partial U/\partial \beta$ with respect to its second argument. Q.E.D.

LEMMA 6. $\partial^2 U/\partial \beta \partial \hat{\beta}$ is (strictly) positive if the local second-order condition is (strictly) satisfied.

Proof of Lemma 6. Differentiating equation (7) with respect to $\hat{\beta}$ gives

$$\frac{\partial^2 U}{\partial \beta \partial \hat{\beta}} = \psi''[\bar{e}(\beta|\hat{\beta})][1 - \bar{e}(\beta)]. \quad (\text{B1})$$

Using (10) and the convexity of ψ , we obtain our conclusion. Q.E.D.

Appendix C

The Planner's Optimization Problem

1. Necessary Conditions

Consider the subconstrained program P' . The Hamiltonian is

$$H = \{S(q) - (1 + \lambda)[\psi(e) + (\beta - e)q] - \lambda U\} + \mu[-\psi'(e)], \quad (\text{C1})$$

where μ is the multiplier associated with (11). The Pontryagin principle yields

$$\frac{\partial H}{\partial q} = 0 = S'(q) - (1 + \lambda)(\beta - e), \quad (\text{C2})$$

$$\frac{\partial H}{\partial e} = 0 = -(1 + \lambda)[\psi'(e) - q] - \mu\psi'(e), \quad (\text{C3})$$

$$\dot{\mu} = -\frac{\partial H}{\partial U} = \lambda. \quad (\text{C4})$$

Furthermore, $\underline{\beta}$ is a free boundary so that

$$\mu(\underline{\beta}) = 0. \quad (\text{C5})$$

Integrating (C4) and using (C5), we obtain

$$\mu(\beta) = \lambda(\beta - \underline{\beta}). \quad (\text{C6})$$

The necessary conditions given in proposition 2 follow.

2. Sufficiency Conditions and the Existence of a Solution for the Planner's Problem

Let us consider how to prove the existence of a solution and to characterize it. Two difficulties may exist. First, the program may be nonconcave. Second, incentive compatibility imposes that the state variable U be almost everywhere differentiable, while Pontryagin's principle assumes that the state variables are piecewise differentiable (with a finite number of pieces).

In step 1 we show that there exists a solution by restricting the analysis to the Pontryagin framework. In step 2 we show that the solution to step 1 is indeed the solution.

Step 1. It is easy to show using assumption 1 that the Pontryagin necessary conditions (C2) and (C3) have a unique interior solution if λ or $(\bar{\beta} - \underline{\beta})$ is small enough. Next, if Kamien and Schwartz's (1971) sufficient condition holds, the solution to the first-order condition is optimal. The sufficient condition is satisfied if the maximized Hamiltonian is concave in the state variable U . This is always the case here because the Hamiltonian is linear in U , and the equations defining the control variables, (C2) and (C3), are independent of U .

Step 2. The space of almost everywhere differentiable increasing functions on $[\underline{\beta}, \bar{\beta}]$ is a closed convex subset of the Banach space $L^\infty([\underline{\beta}, \bar{\beta}], R)$. Let A be the subspace of piecewise-continuous functions. The objective function is continuous in e and s . Since any increasing function of L^∞ can be approximated as closely as desired in the supnorm topology by functions in A and since we have a solution to the maximization in A , it is a solution to the maximization.

Appendix D

Nonlinearity and Cost Disturbances

Let us show that a scheme that is not linear in cost cannot implement the optimal solution for all distributions of the disturbance. We know that $t(\beta, C)$ must satisfy

$$s^*(\beta) = Et\{\beta, [\beta - e^*(\beta)]q^*(\beta) + \epsilon\}. \quad (D1)$$

If t is not linear in cost, there exist β , C_1 , C_2 , and C_3 such that

$$\frac{t(\beta, C_1) - t(\beta, C_2)}{C_1 - C_2} \neq \frac{t(\beta, C_1) - t(\beta, C_3)}{C_1 - C_3}. \quad (D2)$$

Define $\epsilon_i \equiv C_i - [\beta - e^*(\beta)]q^*(\beta)$, and consider the family of discrete distributions with three atoms at ϵ_1 , ϵ_2 , and ϵ_3 and no weight elsewhere (since these distributions can be approximated by continuous distributions, we could actually restrict ourselves to continuous distributions). It is clear that by varying the weights on the three disturbance levels and given (D2), (D1) cannot always be satisfied.

References

- Baron, David P. "A Model of the Demand for Investment Banking Advising and Distribution Services for New Issues." *J. Finance* 37 (September 1982): 955-76.

- Baron, David P., and Besanko, David. "Regulation, Asymmetric Information, and Auditing." *Rand J. Econ.* 15 (Winter 1984): 447-70.
- Baron, David P., and Myerson, Roger B. "Regulating a Monopolist with Unknown Costs." *Econometrica* 50 (July 1982): 911-30.
- Bergson, Abram. "Managerial Risks and Rewards in Public Enterprises." *J. Comparative Econ.* 2 (September 1978): 211-25.
- Bonin, John P. "On the Design of Managerial Incentive Structures in a Decentralized Planning Environment." *A.E.R.* 66 (September 1976): 682-87.
- Caillaud, Bernard; Guesnerie, Roger; Rey, Patrick; and Tirole, Jean. "The Normative Economics of Government Intervention in Production in the Light of Incentive Theory: A Review of Recent Contributions." Technical Report no. 473. Stanford, Calif.: Stanford Univ., IMSSS, 1985.
- Crampes, Claude. "Theory of Incentives and Rate of Return Regulation." Manuscript. Toulouse: Univ. Toulouse, GREMAQ, 1982.
- Demsetz, Harold. "Why Regulate Utilities?" *J. Law and Econ.* 11 (April 1968): 55-65.
- Domar, Evsey D. "On the Optimal Compensation of a Socialist Manager." *Q.J.E.* 88 (February 1974): 1-18.
- Ellman, Michael. "Bonus Formulae and Soviet Managerial Performance: A Further Comment." *Southern Econ. J.* 39 (April 1973): 652-53.
- Fan, Liang-Shing. "On the Reward System." *A.E.R.* 65 (March 1975): 226-29.
- Finsinger, Jorg, and Vogelsang, Ingo. "Alternative Institutional Frameworks for Price Incentive Mechanisms." *Kyklos* 34, no. 3 (1981): 388-404.
- . "Performance Indices for Public Enterprises." In *Public Enterprise in Less-developed Countries*, edited by Leroy P. Jones. New York: Cambridge Univ. Press, 1982.
- Freixas, Xavier, and Laffont, Jean-Jacques. "Marginal Cost Pricing versus Average Cost Pricing under Moral Hazard." *J. Public Econ.* 26 (March 1985): 135-46.
- Grossman, Sanford J., and Hart, Oliver D. "An Analysis of the Principal-Agent Problem." *Econometrica* 51 (January 1983): 7-45.
- Guesnerie, Roger, and Laffont, Jean-Jacques. "A Complete Solution to a Class of Principal-Agent Problems with an Application to the Control of a Self-managed Firm." *J. Public Econ.* 25 (December 1984): 329-69.
- Harris, Milton, and Raviv, Artur. "Optimal Incentive Contracts with Imperfect Information." *J. Econ. Theory* 20 (April 1979): 231-59.
- Holmström, Bengt. "Moral Hazard and Observability." *Bell J. Econ.* 10 (Spring 1979): 74-91.
- Holmström, Bengt, and Milgrom, Paul. "Aggregation and Linearity in the Provision of Intertemporal Incentives." Mimeographed. Evanston, Ill.: Northwestern Univ., 1984.
- Kamien, Morton I., and Schwartz, Nancy L. "Sufficient Conditions in Optimal Control Theory." *J. Econ. Theory* 3 (June 1971): 207-14.
- Laffont, Jean-Jacques, and Tirole, Jean. "Using Cost Observation to Regulate Firms." Discussion Paper no. 27. Paris: CERAS, 1984.
- . "The Dynamics of Incentive Contracts." Mimeographed. Toulouse: Univ. Toulouse, GREMAQ, 1985.
- Landsberger, M., and Chazan, D. "Admissibility of Net Income Distributions, Optimal Income Taxes and the Detection Mechanism under Tax Evasion." Mimeographed. Haifa: Haifa Univ., 1983.
- Loeb, Martin, and Magat, Wesley A. "A Decentralized Method for Utility Regulation." *J. Law and Econ.* 22 (October 1979): 399-404.

- Mirrlees, James A. "Notes on Welfare Economics, Information and Uncertainty." In *Essays on Economic Behavior under Uncertainty*, edited by Michael Balch, Daniel McFadden, and Shih-yen Wu. Amsterdam: North-Holland, 1974.
- . "The Optimal Structure of Incentives and Authority within an Organization." *Bell J. Econ.* 7 (Spring 1976): 105–31.
- Peck, Merton S., and Scherer, Frederic M. *The Weapons Acquisition Process: An Economic Analysis*. Boston: Harvard Univ., Grad. School Bus. Admin., 1962.
- Polinsky, A. Mitchell, and Shavell, Steven. "The Optimal Tradeoff between the Probability and Magnitude of Fines." *A.E.R.* 69 (December 1979): 880–91.
- Ponsard, Jean Pierre, and de Pouvourville, G. *Marches publiques et politique industrielle*. Paris: Economica, 1982.
- Sappington, David. "Optimal Regulation of Research and Development under Imperfect Information." *Bell J. Econ.* 13 (Autumn 1982): 354–68.
- Scherer, Frederic M. "The Theory of Contractual Incentives for Cost Reduction." *Q.J.E.* 78 (May 1964): 257–80.
- Shavell, Steven. "Risk Sharing and Incentives in the Principal and Agent Relationship." *Bell J. Econ.* 10 (Spring 1979): 55–73.
- Tam, Mo-Yin S. "On Incentive Structures of a Socialist Economy." *J. Comparative Econ.* 3 (September 1979): 277–84.
- . "Reward Structures in a Planned Economy: The Problem of Incentives and Efficient Allocation of Resources." *Q.J.E.* 96 (February 1981): 111–28.
- Tirole, Jean. "Procurement and Renegotiation." *J.P.E.* 94 (April 1986): 235–59.
- Vogelsang, Ingo. "Effort Rewarding Incentive Mechanism for Public Enterprise Managers." *Internat. J. Indus. Organization* 1 (1983): 253–73.
- Weitzman, Martin L. "The New Soviet Incentive Model." *Bell J. Econ.* 7 (Spring 1976): 251–57.
- Williamson, Oliver E. "Franchise Bidding for Natural Monopolies—in General and with Respect to CATV." *Bell J. Econ.* 7 (Spring 1976): 73–104.