JOURNAL OF ECONOMIC THEORY 66, 238–263 (1995)

On Countervailing Incentives*

GIOVANNI MAGGI

Economics Department, Princeton University, Princeton, New Jersey 08544

AND

ANDRÉS RODRÍGUEZ-CLARE

Graduate School of Business, University of Chicago, Chicago, Illinois 60637

Received March 11, 1993; revised February 10, 1994

This paper extends and unifies previous work on optimal contracts under countervailing incentives, shedding light in particular on the relation between countervailing incentives and pooling ("inflexible rules"). Our main result is that the nature of the optimal contract depends crucially on whether the agent's utility is quasiconcave or quasiconvex in the private parameter: the optimal contract is separating in the former case and it may entail pooling in the latter case. *Journal of Economic Literature* Classification Number: D82. © 1995 Academic Press, Inc.

1. INTRODUCTION

Most of the existing principal–agent models with asymmetric information are structured so that the agent has a systematic incentive to always overstate or always understate his private information. Consider, for instance, the regulation of a firm that has private information about its constant marginal cost θ. In order to secure a higher compensation from the principal, the firm is tempted to overstate its costs, and this implies a systematic incentive to overstate the value of θ. In this case, the properties of the optimal contract (see, for example, Guesnerie and Laffont [6]) are the following: (i) The firm receives no rents when θ is at its upper extreme, and the firm's rents are higher for smaller θ. (ii) The firm produces less than the first-best level for all values of the cost except the lowest. (iii) Under a certain regularity condition, the equilibrium is fully separating.

* We are grateful to Mario Epelbaum, Giuseppe Lopomo, Eric Talley, and seminar participants at Stanford University for helpful comments and discussion. We also thank three anonymous referees for detailed comments. Giovanni Maggi and Andrés Rodríguez-Clare gratefully acknowledge support from CEPR and the Alfred P. Sloan Foundation, respectively.

0022-0531/95 $6.00

In many situations, however, the agent may have an incentive to either understate or overstate his private information, depending on its realization. Following the regulation example, this situation arises, for instance, when there is a fixed cost of production $K(\theta)$ which depends negatively on the marginal cost θ ($K'(\theta) < 0$), a case considered by Lewis and Sappington [9]. This is a plausible case, since a low marginal cost is likely to be associated with high overhead costs. A similar situation arises when the agent's reservation utility $\bar{v}(\theta)$ is decreasing in the marginal cost θ. This is likely to happen, as agents with low marginal costs are more prone to have attractive alternative options. A type-dependent reservation utility appears, for instance, in Brainard and Martimort [2], Feenstra and Lewis [4], Lewis and Sappington [10], and Moore [12].

In these situations, the agent is tempted to (i) overstate the fixed cost or reservation utility, which requires *understanding* θ, and (ii) overstate the marginal cost, which requires *overstating* θ. To use Lewis and Sappington's expression, the agent faces "countervailing incentives."[1] As Fudenberg and Tirole [5] point out, there are no general results for this class of problems. In this paper, we attempt a unified analysis of this issue. To focus the discussion, we state our results for the case in which countervailing incentives arise because of type-dependent reservation utility, although the results allow alternative interpretations.

Most of the principal-agent models with countervailing incentives share the result that the optimal contract involves pooling. Lewis and Sappington [9, 10] interpret this result as a possible explanation for the prevalence of "inflexible rules" in agency contracts. In this paper we will show that in fact pooling is not a general consequence of countervailing incentives; it depends critically on the assumption made in all those models that the reservation utility is a concave function of the marginal cost ($\bar{v}''(\theta) \leqslant 0$). We show that when $\bar{v}(\theta)$ is strictly convex ($\bar{v}''(\theta) > 0$), the optimal contract is fully separating. This result provides an important qualification to the view that countervailing incentives imply the optimality of inflexible rules.

We also show that the distortion of production introduced by asymmetric information and the distribution of the agent's information rents

[1] Countervailing incentives can arise also when the contract involves more than one action to be taken by the agent. For example, the problem of optimal bundling for a monopolist, first formulated by Adams and Yellen [1], involves countervailing incentives whenever the consumer's willingness to pay is inversely correlated across goods. Another example of countervailing incentives arising in situations where the agent performs multiple actions is the case of costly state falsification, as in Lacker and Weinberg [7] and Maggi and Rodríguez-Clare [11].

depend critically on the shape of $\bar{v}(\theta)$. Our main findings can be summarized as follows[2].

1. If $\bar{v}(\theta)$ is strictly concave, all types earn positive rents except for a single interior type. Pooling arises around the type that earns zero rents. Production is lower than the first-best level for low realizations of θ and higher than the first-best level for high realizations of θ. The first-best production level is implemented at both extremes of the support and at an interior type. This is the case studied by Lewis and Sappington [9].

2. If $\bar{v}(\theta)$ is linear, the structure of the contract is the same as in case 1, with the difference that a nondegenerate interval of types earn zero rents.

3. If $\bar{v}(\theta)$ is slightly convex, a nondegenerate interval of types earn zero rents, and the equilibrium is fully separating. One interesting feature in this case is that, over the range of types that earn no rents, the production level is independent of the distribution of θ. The distortion of production is qualitatively the same as in the two previous cases.

4. If $\bar{v}(\theta)$ is highly convex, we have the unusual result that the agent's rents are bell-shaped: rents are highest for an interior type, and the two extreme types earn zero rents. Only an interior type receives the efficient contract, and production is distorted at both extremes of the distribution. Furthermore, the sign of the distortion is opposite to that in cases 1, 2, and 3. The equilibrium is fully separating.

In the light of these results, the case in which the reservation utility is linear in θ, frequently assumed in the literature (e.g., Brainard and Martimort [2], Feenstra and Lewis [4], and Lewis and Sappington [10]), is revealed to be a "knife-edge" case: if $\bar{v}(\theta)$ becomes slightly convex, the pooling vanishes; if $\bar{v}(\theta)$ becomes slightly concave, the range of types that earn no rents collapses to a single point.

The shape of $\bar{v}(\theta)$, or, more concretely, whether $\bar{v}(\theta)$ is concave or convex, depends on the specific outside opportunities the agent faces. For example, suppose the agent owns some human capital, measured by H, which can be taken as an index of the agent's productivity, so that the marginal cost of production is given by $\theta = 1/H$. If the same human capital is employed in some alternative sector, it will yield returns $\pi(H) = \pi(1/\theta) \equiv \bar{v}(\theta)$. If the returns to H in the alternative sector are not too diminishing, $\bar{v}(\theta)$ will be convex,[3] but if these returns are highly diminishing, $\bar{v}(\theta)$ will be concave.

[2] The results that follow apply to the case in which the conflicting incentives to lie are balanced enough, so that the standard procedure (solving the problem imposing the participation constraint at one extreme only) does not work. If, for example, the incentive to overstate θ dominates the incentive to understate θ for all θ, the standard techniques and results apply.

[3] For example, if $\pi(H) = \log(H)$, then $\bar{v}(\theta) = -\log(H)$, which is decreasing and convex.

Although we focus most of our discussion on the case of constant marginal cost, we also extend our results to a more general cost function. In this more general case, the optimal contract is separating when the agent's "net" utility (i.e., utility in excess of the reservation level) is strictly quasiconcave in the private parameter, whereas it involves pooling only if the agent's net utility is quasiconvex in the private parameter.

The paper is structured as follows: in Section 2 we lay out the basic model. In Section 3 we give a heuristic discussion of the results. In Section 4 we derive the results formally. Section 5 concludes.

2. The Model

We consider a situation where a "principal" contracts with an "agent" to produce a certain amount of some good, q, and compensates the agent with a monetary transfer, y. The cost for the agent of producing q is given by θq, where θ denotes the agent's privately observed constant marginal cost. We also assume that the agent's utility is linear in income, so that it is of the form $y-\theta q$. The agent also has an outside opportunity, from which he can derive a utility level $\bar{v}(\theta)$: we assume that agents with higher marginal costs have less attractive outside opportunities, $\bar{v}'(\theta)<0$.

The agent knows his type before signing the contract but the principal does not. The principal's uncertainty about θ is represented by a (common knowledge) probability distribution $F(\theta)$ with associated density function $f(\theta)$ strictly positive on the support $\Theta=[\theta_0, \theta_1]$. It is assumed that both $F(\theta)/f(\theta)$ and $(F(\theta)-1)/f(\theta)$ are increasing functions of θ.[4]

The principal designs a contract (to which she can perfectly commit) to maximize her expected utility in the exchange with the agent. Such a contract must give all types at least their reservation utility $\bar{v}(\theta)$. We assume that the principal is risk neutral, with a utility function represented by $W(q)-y$, where $W(\ )$ is increasing and concave ($W'>0$, $W''<0$).[5]

[4] These monotonicity conditions are commonly assumed in the literature to ensure that the equilibrium is fully separating. They were first used together in Champsour and Rochet [3]. If these assumptions are violated, there is usually pooling in the optimal contract. Since here we are mostly concerned with the relation between pooling and countervailing incentives, to focus on this aspect we assume that these monotonicity conditions hold. Note that both monotonicity conditions are satisfied by the uniform distribution, for example.

[5] We assume that the revenue from production is high enough relative to $q\theta+\bar{v}(\theta)$ for all θ, so that the principal will always induce the agent to produce a strictly positive amount. This assumption is made to avoid problems of optimal cutoff.

The Revelation Principle guarantees that there is no loss of generality in restricting to truthful revelation mechanisms. Therefore the principal's problem can be written as

$$(\text{P}) \quad \max_{q(\theta),\, y(\theta)} E[W(q) - y]$$

$$\text{s.t. IC: } \theta \in \underset{\theta_r}{\operatorname{argmax}} \{y(\theta_r) - \theta q(\theta_r)\}, \qquad \text{for all } \theta$$

$$\text{PC: } U(y(\theta), q(\theta), \theta) \equiv y(\theta) - \theta q(\theta) - \bar{v}(\theta) \geqslant 0, \qquad \text{for all } \theta,$$

where θ_r is the agent's report, and IC and PC stand for incentive-compatibility constraint and participation constraint, respectively. $U(y, q, \theta)$ is the agent's net utility, that is, the utility he gets in excess of the reservation level. For future reference, with a slight abuse of notation, we let $U(\theta) \equiv U(y(\theta), q(\theta), \theta)$.

To solve problem P we replace the IC condition with necessary and sufficient conditions: a local (first-order) condition and a monotonicity condition, given by the following lemma.

LEMMA 1. *Necessary and sufficient conditions for IC are*:

(i) $dU/d\theta = U_\theta = -\bar{v}'(\theta) - q(\theta)$ (*local optimality*)

(ii) $q(\theta)$ *nonincreasing* (*monotonicity*).

The proof of Lemma 1 is standard and is omitted.[6]

In the next section we provide an intuitive discussion of the problem. The results will be derived formally in Section 4.

3. Informal Discussion of Results

Condition (i) of Lemma 1, to which we will simply refer to as condition (i), is crucial to understanding most of our results. To induce truthtelling, the principal has to provide the agent with "information rents," $U(\theta)$, to compensate for the incentives to misreport θ. Condition (i) specifies how information rents must change with θ for the contract to be incentive compatible, or, in other words, it specifies the required slope of $U(\theta)$ to induce truthtelling.

To overstate his marginal cost, the agent has an incentive to report a θ *higher* than his true θ. The gains from overstating the marginal cost increase with the level of production q. This explains the second term on

[6] See Laffont [8], chapter 10.

the RHS of condition (i). On the other hand, to overstate his reservation utility, the agent has an incentive to report a θ *lower* than his true θ, since $\bar{v}' < 0$. This explains the first term on the RHS of condition (i). Information rents will be increasing or decreasing in θ depending on which of these incentives dominates.

The existence of information rents leads the principal to distort the level of production away from the first-best level, which we denote by $q^F(\theta)$ and is given implicitly by $W'(q^F) = \theta$.[7] In the standard case, in which $\bar{v}'(\theta) = 0$ for all θ, $U(\theta)$ is decreasing regardless of the level of q. Therefore, the PC is satisfied for all types if we set $U(\theta_1) = 0$. The concern of the principal will then be to *increase* the slope of $U(\theta)$ (make it less negative) by decreasing q below the full-information level, trading off the savings of information rents with the costs of the induced distortion. It is a standard result that the optimal production schedule, which we denote by $q^L(\theta)$, is given implicitly by

$$W'(q^L) - \theta = F(\theta)/f(\theta).$$

This result carries over to the case where $-\bar{v}'(\theta)$ is sufficiently "small": if $-\bar{v}'(\theta) < q^L(\theta)$ for all θ, then $q^L(\theta)$ is the optimal production schedule, just as when $-\bar{v}'(\theta) = 0$ for all θ. Also note that since $W'(q^L(\theta)) \geqslant \theta$, then $q^L(\theta) \leqslant q^F(\theta)$, with equality only for θ_0: to reduce information rents, the principal designs a contract that induces underproduction for all types except the one with the lowest unit cost.

When $-\bar{v}'$ is sufficiently high, the dominating incentive is to understate θ and the slope of $U(\theta)$, as given by condition (i), will be positive for all types. Therefore, setting $U(\theta_0) = 0$ satisfies the PC for all types. In this case, the concern of the principal will be to *decrease* the slope of $U(\theta)$. To do this she will increase the level of production above the full-information level, making it more costly for the agent to understate his marginal costs and decreasing the net incentives to understate θ. When the dominating incentive for all types is to understate θ, the optimal production schedule, which we denote by $q^H(\theta)$, is given implicitly by

$$W'(q^H) - \theta = (F(\theta) - 1)/f(\theta).$$

More formally, $q^H(\theta)$ is optimal when $-\bar{v}'(\theta) > q^H(\theta)$ for all θ. In this case, setting production at $q^H(\theta)$ gives information rents that are increasing in θ ($U_\theta = -\bar{v}'(\theta) - q^H(\theta) > 0$). Note that $q^H(\theta) \geqslant q^F(\theta)$ with equality only for θ_1: there is overproduction in the optimal contract.

The previous discussion shows that when $-\bar{v}'$ is sufficiently high or sufficiently low for all θ, the incentive to misreport θ has the same sign for all

[7] To avoid corner solutions, we asume that $W'(q) \to 0$ as $q \to \infty$ and $W'(q) \to \infty$ as $q \to 0$.

types. A more interesting situation arises when $-\bar{v}'$ assumes intermediate values. Consider first the simplest case, in which the agent's reservation utility is linear in θ (so that $-\bar{v}'$ is constant), and suppose that $-\bar{v}'$ intersects $q^{\mathrm{F}}(\theta)$ for an interior θ, as in Fig. 1. Let $\hat{\theta}$ and $\hat{\hat{\theta}}$ be defined implicitly by $-\bar{v}' = q^{\mathrm{L}}(\hat{\theta})$ and $-\bar{v}' = q^{\mathrm{H}}(\hat{\hat{\theta}})$, respectively. In this case, for $\theta < \hat{\theta}$ we have $-\bar{v}' < q^{\mathrm{L}}(\theta)$, and hence, as above, the optimal contract involves $q(\theta) = q^{\mathrm{L}}(\theta)$. On the other hand, for $\theta > \hat{\hat{\theta}}$ we have $-\bar{v}' > q^{\mathrm{H}}(\theta)$, and hence, as above, the optimal contract involves $q(\theta) = q^{\mathrm{H}}(\theta)$. For $\hat{\theta} \leqslant \theta \leqslant \hat{\hat{\theta}}$, the optimal contract involves $q(\theta) = -\bar{v}'$, so that $U'(\theta) = U_\theta = 0$. Since $U(\theta)$ is decreasing for $\theta < \hat{\theta}$, constant for $\hat{\theta} \leqslant \theta \leqslant \hat{\hat{\theta}}$, and increasing for $\theta > \hat{\hat{\theta}}$, then it is optimal to impose $U(\theta) = 0$ for θ in $[\hat{\theta}, \hat{\hat{\theta}}]$, as shown in Fig. 1. All types in $[\hat{\theta}, \hat{\hat{\theta}}]$ are assigned the same production level (pooling) and earn no information rents. This conclusion matches the results of Feenstra and Lewis [4], Brainard and Martimort [2], and Lewis and Sappington [10]: all these models correspond to the case where $\bar{v}(\theta)$ is linear.

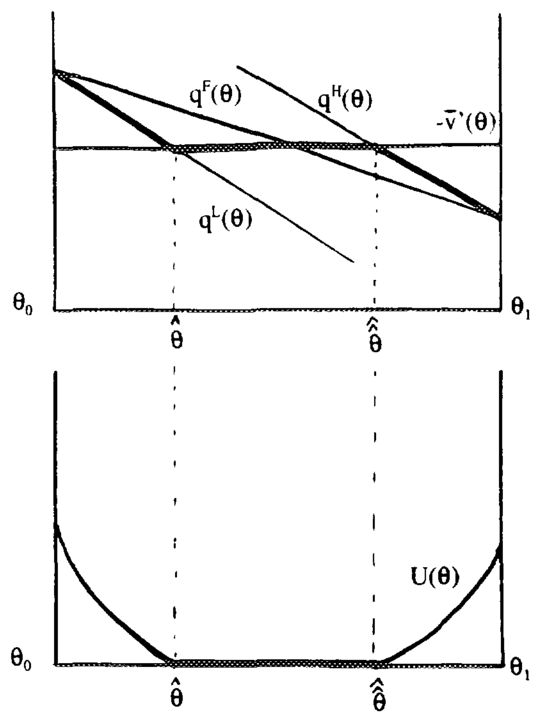


FIGURE 1

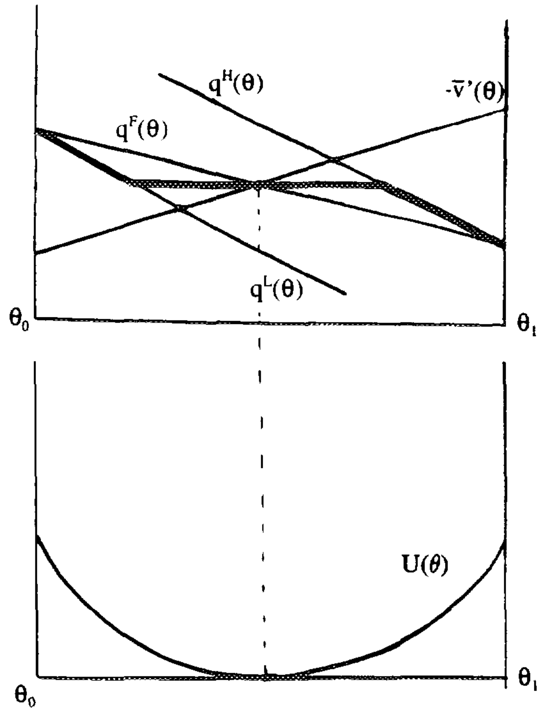


FIGURE 2

This discussion should convince the reader that the form of the contract depends crucially on the shape of the curve $\bar{v}'(\theta)$, which determines how the incentives to misreport θ change over the support Θ. Suppose $-\bar{v}'(\theta)$ is increasing, as in Fig. 2. In this case, the principal cannot keep a whole interval of types on their reservation utility, because to do so, $q(\theta)$ would have to equal $-\bar{v}'(\theta)$ for a whole interval and this would violate the monotonicity condition for IC (condition (ii) of Lemma 1). Therefore the solution has to be "patched" (that is, it has to be smoothed into a monotonic function in the least costly way), and this leads to pooling, as in Fig. 2. In this case, all types earn information rents, except for an interior type. This is the logic behind the result of Lewis and Sapington [9].

However, this pooling result does not carry over to the case where $-\bar{v}'(\theta)$ is decreasing, because in this case the principal *can* keep a whole interval of types on their reservation level without violating the monotonicity condition. Here two different solutions are possible. If $-\bar{v}'(\theta)$ is decreasing at a slow rate (see Fig. 3), the optimal contract will be

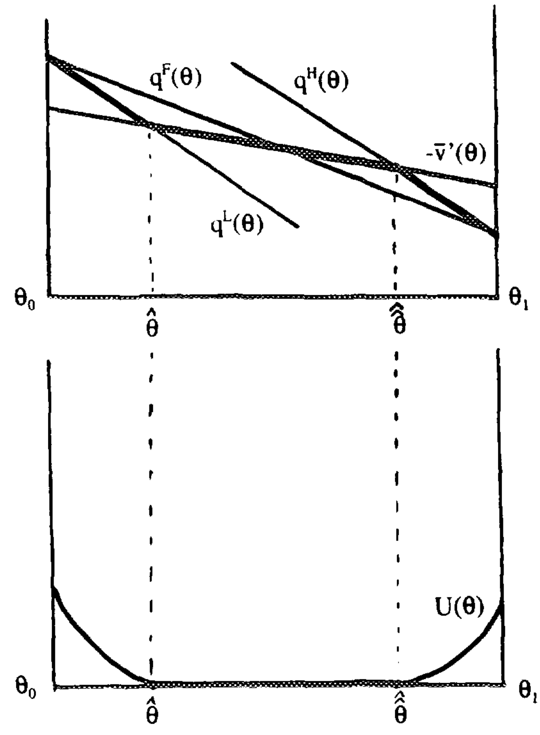


FIGURE 3

separating and will assign no information rents for an interval of types, as in the case of $-\bar{v}'(\theta)$ constant. However, if $-\bar{v}'(\theta)$ decreases at a fast rate (namely, if $-\bar{v}'(\theta)$ decreases faster than $q^F(\theta)$), as in Fig. 4, the structure of the contract changes dramatically. The incentives to lie reverse sign relative to the previous case: types close to θ_0 now have an incentive to understate θ and types close to θ_1 have an incentive to overstate θ. The condition $U'(\theta) = U_\theta$ implies that rents have to be increasing over the interval of θ where the dominating incentive is to understate θ, and vice versa. The strength of countervailing incentives is highest where $\bar{v}'(\theta)$ crosses $q^F(\theta)$, at which point incentives to overstate and to understate are perfectly balanced and the agent has no incentive to lie. As a consequence, rents will be *bell-shaped*, reaching a maximum for some interior type. The production allocation turns out to be separating in this case as well.

The next section offers a formal derivation of these results and of the other features of the optimal contract.

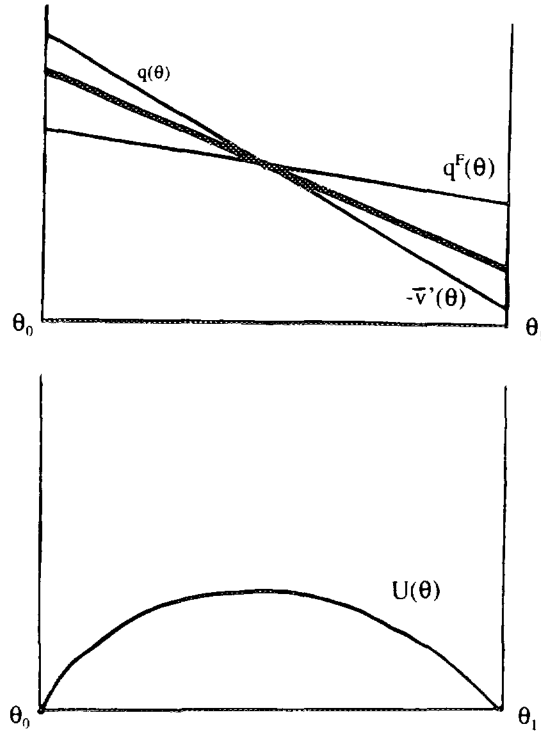


FIGURE 4

4. Formal Derivation of Results

The usual procedure to find the solution for (P) involves replacing the IC constraint with condition (i) in Lemma 1, assuming that the PC binds only at one extreme, and then checking that the solution satisfies the PC for all θ and the monotonicity condition given by condition (ii). However, in our case the sign of U_θ may change over $[\theta_0, \theta_1]$. Therefore, it is possible that the solution given by the standard procedure violates the PC for some θ. For this reason, we have to introduce the PC explicitly in the optimization problem.

First we solve the problem ignoring condition (ii), and then we check whether the solution satisfies it. It turns out that condition (ii) is satisfied as long as $\bar{v}(\theta)$ is convex ($\bar{v}'' \geqslant 0$). This is the case we focus on in the next subsection. In Section 4.2 we consider the case where $\bar{v}(\theta)$ is strictly concave ($\bar{v}'' < 0$), where we have to introduce condition (ii) explicitly in the optimization problem. For expositional reasons, we assume that $\bar{v}''(\theta)$ does not change sign over Θ. Relaxing this assumption would complicate the results without adding new insights.

4.1. $\bar{v}(\theta)$ *Convex*

In this section we solve the principal's problem, imposing the first-order condition for incentive compatibility (condition (i) of Lemma 1) and then checking that the solution we find satisfies also condition (ii) of Lemma 1 to guarantee global incentive compatibility. Letting $V(q, \theta) \equiv W(q) - \theta q - \bar{v}(\theta)$ and substituting the transfer $y(\theta)$ away in the principal's objective, the problem we solve in this section is

$$\text{(P1)} \quad \max_{q(\theta),\, U(\theta)} \int_{\theta_0}^{\theta_1} [V(q(\theta), \theta) - U(\theta)]\, f(\theta)\, d\theta$$

$$\text{s.t. } U'(\theta) = -\bar{v}'(\theta) - q(\theta)$$

$$U(\theta) \geqslant 0 \qquad \text{for all} \quad \theta \in \Theta.$$

In this optimization problem, the control variable is q, the state variable is U, the costate variable is μ, the Hamiltonian is

$$H(U, q, \mu, \theta) = [V(q, \theta) - U]\, f(\theta) - \mu[\bar{v}'(\theta) + q],$$

and the Lagrangian is

$$L = H + \tau U,$$

where τ is the multiplier of the constraint $U(\theta) \geqslant 0$.

To solve this problem, we make use of a set of sufficient conditions due to Seierstad and Sydsaeter (S-S) [13] for an optimal control problem with pure state constraints. These sufficient conditions allow jumps in the costate variable. Our strategy will be first to look for a solution with a continuous costate variable. If none exists, we will allow jumps in the costate variable and use the restrictions given in S-S for the jump points.

The first-order condition for the maximization of the Hamiltonian with respect to q is

$$V_q(q, \theta) = \mu(\theta)/f(\theta). \tag{1}$$

Since $V_{qq} < 0$, condition (1) is also sufficient for the maximization of the Hamiltonian. The other sufficient conditions are

$$d\mu/d\theta = -\partial L/\partial U = f(\theta) - \tau(\theta) \qquad \text{(costate equation)} \qquad (2)$$

$$dU(\theta)/d\theta = -\bar{v}'(\theta) - q(\theta) \qquad \text{(state equation)} \qquad (3)$$

$$\tau(\theta)U(\theta) = 0, \quad \tau(\theta) \geqslant 0, \quad U(\theta) \geqslant 0 \qquad \text{(complementary slackness)} \qquad (4)$$

$$\mu(\theta_0)U(\theta_0) = 0, \quad \mu(\theta_0) \leqslant 0, \quad \mu(\theta_1)U(\theta_1) = 0, \; \mu(\theta_1) \geqslant 0$$
$$\text{(transversality conditions).}^{8} \qquad (5)$$

Since $H(U, q, \mu, \theta)$ is concave in U, if the configuration $(U(\theta), q(\theta), q(\theta), \tau(\theta))$ satisfies (1), (2), (3), (4), and (5), it also satisfies the sufficient conditions for an optimum.

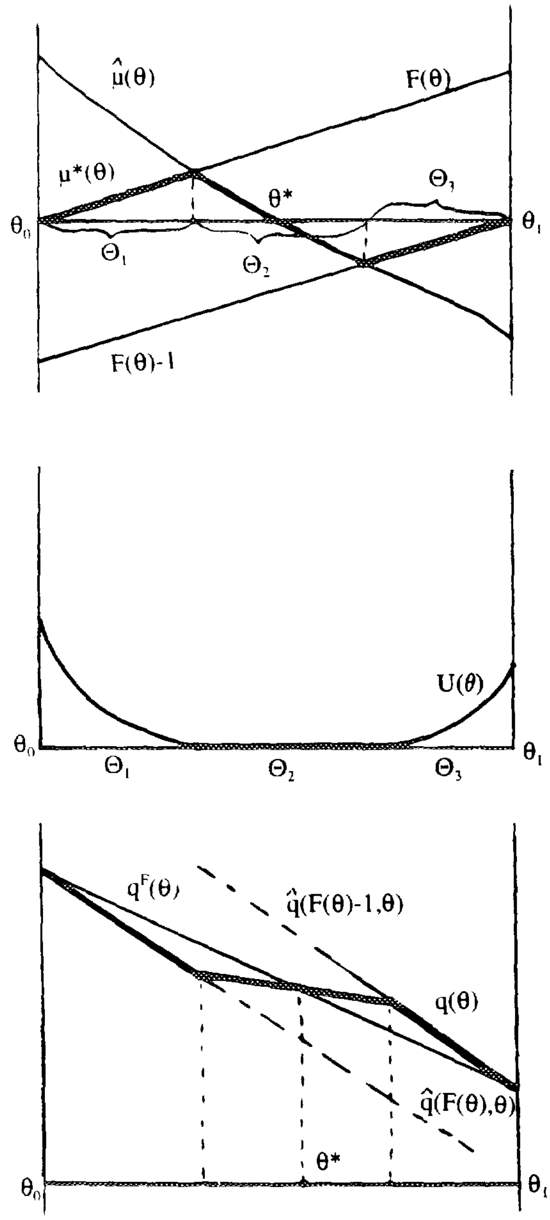
Let $\hat{q}(\mu, \theta)$ denote the value of q that maximizes the Hamiltonian given μ and θ, defined implicitly by (1), and let $\hat{\mu}(\theta)$ be the solution in μ to the following equation: $\bar{v}'(\theta) + \hat{q}(\mu, \theta) = 0$. $\hat{\mu}(\theta)$ is the value of the costate variable such that the agent's utility is constant ($U'(\theta) = U_\theta = 0$). The slope of $\hat{\mu}(\theta)$ is crucial in determining the optimal contract. If the PC is to be binding on a nondegenerate interval, then $\mu(\theta)$ must be equal to $\hat{\mu}(\theta)$ on that interval. From (2) and (4), we have $\tau(\theta) = f(\theta) - \hat{\mu}'(\theta) \geqslant 0$; thus at an optimum it must be $\mu'(\theta) \leqslant f(\theta)$. The PC can therefore be binding on a nondegenerate interval only if $\hat{\mu}'(\theta) \leqslant f(\theta)$.

Our strategy to solve the problem is to conjecture a solution and verify that it satisfies the sufficient conditions. The critical part is to construct the right solution for $\mu(\theta)$. Consider the following schedule, illustrated in Fig. 5:

$$\mu^*(\theta) = \begin{cases} F(\theta) & \text{if } \; F(\theta) < \hat{\mu}(\theta) \\ \hat{\mu}(\theta) & \text{if } \; F(\theta) - 1 < \hat{\mu}(\theta) \leqslant F(\theta) \\ F(\theta) - 1 & \text{if } \; \hat{\mu}(\theta) \leqslant F(\theta) - 1. \end{cases}$$

For future reference, also define Θ_1, Θ_2, and Θ_3 as the subintervals of Θ where $\mu^*(\theta)$ is equal to $F(\theta)$, $\hat{\mu}(\theta)$, and $F(\theta) - 1$, respectively.

[8] In the optimal control problem with pure state constraints studied by S-S, the state variable is fixed at the initial point and "free" at the final point. In this case, the transversality condition is $\mu(\theta_1)U(\theta_1) = 0$, $\mu(\theta_1) \geqslant 0$. However, in our case, the state variable is free at both extremes. We can get the sufficient conditions for this case by applying the theory of optimal control for the case where there can be jumps in the state variables (see Seierstad and Sydsaeter [13, p.194]). This gives us the transversality conditions in (5).



$\hat{\mu}(\theta)$
$F(\theta)$
$\mu^*(\theta)$
Θ_3
θ^*
θ_0
θ_1
Θ_1
Θ_2
$F(\theta)-1$
$U(\theta)$
θ_0
θ_1
Θ_1
Θ_2
Θ_3
$q^F(\theta)$
$\hat{q}(F(\theta)-1,\theta)$
$q(\theta)$
$\hat{q}(F(\theta),\theta)$
θ^*
θ_0
θ_1

FIGURE 5

In order for $\mu^*(\theta)$ to satisfy the sufficient conditions, we need to check that $\hat{\mu}'(\theta) \leqslant f(\theta)$ for $\theta \in \Theta_2$. Differentiating the equality $\bar{v}'(\theta) + \hat{q}(\hat{\mu}(\theta), \theta) = 0$ with respect to θ and using (1) gives an expression for $\hat{\mu}'(\theta)$ over the interval Θ_2:

$$\hat{\mu}'(\theta) = -\frac{\bar{v}''(\theta) + \partial\hat{q}/\partial\theta}{\partial\hat{q}/\partial\mu}$$
$$= -\bar{v}''(\theta)\, f(\theta)\, V_{qq}(\hat{q}, \theta) + f'(\theta)\, V_q(\hat{q}, \theta) - f(\theta). \tag{6}$$

$(\hat{\mu}'(\theta) - f(\theta))$ may be either positive or negative, depending on the magnitude of $\bar{v}''$. It can be shown that if $\bar{v}''(\theta) = 0$ for all θ, then necessarily $\hat{\mu}'(\theta) < f(\theta)$ for all $\theta \in \Theta_2$.[9] Therefore, by continuity, the condition is always satisfied as long as $\bar{v}''(\theta)$ is small. When $\bar{v}''(\theta)$ is high enough, on the other hand, the condition $\hat{\mu}'(\theta) \leqslant f(\theta)$ for all $\theta \in \Theta_2$ is not satisfied and $\mu^*(\theta)$ cannot be part of an optimum, in which case a different construction is required. We will consider these two cases separately in the remainder of this section. For expositional convenience, we will assume that $(\hat{\mu}'(\theta) - f(\theta))$ does not change sign (the generalization just involves partioning the interval so that this assumption holds in each subinterval). First we focus on the case where $(\hat{\mu}'(\theta) - f(\theta)) < 0$, which can be seen as the case where $\bar{v}''$ is nonnegative and "small."

4.1.1. *$\bar{v}(\theta)$ slightly convex.* In this subsection we derive the solution to P1 for the case where $\hat{\mu}'(\theta) < f(\theta)$ for all $\theta \in \Theta_2$, that is, where $\bar{v}(\theta)$ is slightly convex. The following lemma characterizes the costate variable in this solution. The proof of Lemma 2, together with those of all the lemmas that follow, is relegated to the Appendix.

LEMMA 2. *If $\hat{\mu}'(\theta) < f(\theta)$ for all $\theta \in \Theta_2$, a solution to P1 entails $\mu(\theta) = \mu^*(\theta)$, $q(\theta) = \hat{q}(\mu^*(\theta), \theta)$, and $U'(\theta) = -\bar{v}'(\theta) - \hat{q}(\mu^*(\theta), \theta)$, with $U(\theta) = 0$ for all $\theta \in \Theta_2$.*

We now verify that $\hat{q}(\mu^*(\theta), \theta)$ satisfies the monotonicity condition (ii) of Lemma 1. For the intervals Θ_1 and Θ_3, this follows directly from (1) and the assumptions on F/f and $(F-1)/f$. The following lemma states that if $\bar{v}'' \geqslant 0$, this is true also for the interval Θ_2, where $q = \hat{q}(\hat{\mu}(\theta), \theta)$.

LEMMA 3. *If $\bar{v}''(\theta) \geqslant \theta$, $(d/d\theta)\, \hat{q}(\hat{\mu}(\theta), \theta) \leqslant 0$, with equality if $\bar{v}''(\theta) = 0$.*

Lemma 3 guarantees that if $\bar{v}''(\theta)$ is small the costate variable $\mu^*(\theta)$ and the production schedule $\hat{q}(\mu^*(\theta), \theta)$ are optimal for problem (P).

[9] To see this, impose $\bar{v}''(\theta) = 0$ in (6), use (1) and the hazard rate assumptions, and use the fact that $\mu \in [F(\theta) - 1, F(\theta)]$.

We now proceed to characterize the utility profile $U(\theta)$ and the allocation $q(\theta)$ in the optimal contract. In particular, it is interesting to compare the optimal allocation $q(\theta)$ with the full-information allocation, $q^F(\theta)$.

The following proposition, which follows directly from the previous lemmas, summarizes the main features of the optimal contract when $\hat{\mu}'(\theta) \leqslant f(\theta)$ for all $\theta \in \Theta_2$, that is, when $\bar{v}(\theta)$ is slightly convex:

PROPOSITION 1. *Suppose $\bar{v}''(\theta) \geqslant 0$ for all θ and $\hat{\mu}'(\theta) \leqslant f(\theta)$ for all $\theta \in \Theta_2$ ($\bar{v}''$ nonnegative and small). Then*:

(i) *If there exists a $\theta^* \in \Theta$ such that $\hat{\mu}(\theta^*) = 0$, then*

$$\begin{cases} q(\theta) = q^F(\theta) & \text{for} \quad \theta \in \{\theta_0, \theta^*, \theta_1\} \\ q(\theta) < q^F(\theta) & \text{for} \quad \theta_0 < \theta < \theta^* \\ q(\theta) > q^F(\theta) & \text{for} \quad \theta^* < \theta < \theta_1. \end{cases}$$

If $\hat{\mu}(\theta) > 0$ for all θ then $q(\theta) < q^F(\theta)$ for all θ. If $\hat{\mu}(\theta) < 0$ for all θ then $q(\theta) > q^F(\theta)$ for all θ.

(ii) *The PC is binding for $\theta \in \Theta_2$.*

(iii) *If $\bar{v}''(\theta) > 0$ for all $\theta \in \Theta_2$, the optimal contract is fully separating (i.e., there is no pooling). If $\bar{v}''(\theta) = 0$ for all $\theta \in \Theta_2$, the optimal contract involves pooling for $\theta \in \Theta_2$.*

Figure 5 shows the qualitative behavior of $\mu(\theta)$ and $U(\theta)$ in the case where $\mu(\theta)$ crosses the zero line and compares the optimal allocation $q(\theta)$ with $q^F(\theta)$. An interesting case occurs when $V_q(q, \theta)$ and $\bar{v}'(\theta)$ are such that $F(\theta) \geqslant \hat{\mu}(\theta) \geqslant F(\theta) - 1$ for all θ, with $\hat{\mu}'(\theta) \leqslant f(\theta)$ for all $\theta \in \Theta_2$. In this case, the optimal contract involves no information rents for *all* types: $\Theta_2 = \Theta$. Yet in this case the contract specifies the first-best allocation for only one type (in the interior of Θ), and the equilibrium is fully separating.

A number of interesting conclusions emerge from this analysis. First, a whole interval of types receive the same net utility but not the same production assignment, whereas in Lewis and Sappington [9] an interval of types produce the same amount but obtain different information rents. Second, on the interval Θ_2 the contract has the interesting feature of being independent of the distribution $f(\theta)$: this is because on Θ_2 we have $\mu(\theta) = \hat{\mu}(\theta)$, which depends solely on the cost function of the agent, and hence $\hat{q}(\hat{\mu}(\theta), \theta)$ does not depend on $f(\theta)$. Third, in principal–agent models of this kind generally the sign of the production distortion $(q - q^F)$ coincides with the sign of the incentive to lie in equilibrium (the sign of U_θ). However, this is not true here: on the interval Θ_2 we have $U_\theta = 0$, yet production is distorted (in either direction) over this interval. This result will carry over to the case in which $\bar{v}(\theta)$ is highly convex.

4.1.2. $\bar{v}(\theta)$ *highly convex*. In this subsection we derive the solution to P1 for the case where $\hat{\mu}'(\theta) > f(\theta)$ for all $\theta \in \Theta_2$. In this case, the solution constructed in the previous section cannot be an optimum, and therefore we must look for a different solution. We conjecture, and later verify, that in this case the solution involves $\tau(\theta) = 0$ for all θ, so that $\mu'(\theta) = f(\theta)$ for all θ. This implies that the schedule $\mu(\theta)$ is of the form $\mu(\theta) = F(\theta) + \mu_0$, where μ_0 is a constant.

Given this conjecture, all that remains is to choose μ_0. The optimal value of μ_0 is such that $\hat{q}(F(\theta) + \mu_0, \theta)$ minimizes information rents while satisfying the PC. First consider the simplest case, in which $\hat{\mu}(\theta)$ does not intersect $F(\theta)$ or $F(\theta) - 1$. In this case, the balance of incentives is unambiguous, and it is easy to show that the optimal choice of μ_0 is either -1 or 0. For instance, if $\hat{\mu}(\theta) > F(\theta)$ for all θ, the optimal costate variable is $\mu(\theta) = F(\theta)$ ($\mu_0 = 0$), and the optimal contract has the standard structure, with the PC binding only at θ_1, information rents decreasing in θ, and underproduction. If instead $\hat{\mu}(\theta)$ intersects either or both $F(\theta)$ and $F(\theta) - 1$, countervailing incentives are balanced and the optimal value of μ_0 will lie between -1 and 0.

To determine the optimal choice for μ_0 in this case, note that since $U'(\theta) = -\bar{v}'(\theta) - \hat{q}(\mu, \theta)$, we have $U(\theta_1) = U(\theta_0) + R(\mu_0)$, where

$$R(\mu_0) \equiv \int_{\theta_0}^{\theta_1} U'(\theta)\, d\theta = -\int_{\theta_0}^{\theta_1} [\bar{v}'(\theta) + \hat{q}(\mu_0 + F(\theta), \theta)]\, d\theta. \tag{7}$$

$R(\mu_0)$ is the utility differential between the high type and the low type that results when the costate variable follows $\mu_0 + F(\theta)$. Note that $R(\mu_0)$ is an increasing function. If $R(0) < 0$, it is optimal to set $\mu_0 = 0$ and $U(\theta_1) = 0$ (as in the standard solution): this will satisfy the transversality condition and the PC for all θ. If $R(-1) > 0$, on the other hand, the optimum involves setting $\mu_0 = -1$ and $U(\theta_0) = 0$. The most interesting case is the one where countervailing incentives are balanced enough, in which case the schedule $R(\mu_0)$ equals zero for some μ_0 in $(-1, 0)$. In this case, the only way to satisfy both the transversality conditions and the PC is to set μ_0 such that $R(\mu_0) = 0$ and $U(\theta_0) = U(\theta_1) = 0$, so that the PC is binding at both extremes. This is stated formally in Lemma 4, which uses the following definition:

$$\mu_0^* \equiv \begin{cases} 0 & \text{if} \quad R(0) \leqslant 0 \\ R^{-1}(0) & \text{if} \quad R(-1) < 0 < R(0) \\ -1 & \text{if} \quad R(-1) \geqslant 0. \end{cases}$$

LEMMA 4. *If* $\hat{\mu}'(\theta) > f(\theta)$ *for* $\theta \in \Theta_2$ *the solution to* P1 *entails* $\mu(\theta) = \mu_0^* + F(\theta)$, $q(\theta) = \hat{q}(\mu_0^* + F(\theta), \theta)$, *and* $U'(\theta) = -\bar{v}'(\theta) - \hat{q}(\mu_0^* + F(\theta), \theta)$.

Lemma 4 refers to the solution to P1, but it is easy to show that given our assumptions on the distribution of θ, the production schedule in Lemma 4 is nonincreasing, and therefore we have identified the optimum for problem P.

The next proposition, which follows immediately from Lemma 4, summarizes the most interesting features of the optimal solution to P when $\hat{\mu}'(\theta) > f(\theta)$ for $\theta \in \Theta_2$. Let θ^* be defined implicitly by $\mu_0^* + F(\theta^*) = 0$ (it is easy to check that θ^* exists and is unique).

PROPOSITION 2. *Suppose that $\hat{\mu}'(\theta) > f(\theta)$ for all $\theta \in \Theta_2$ ($\bar{v}''$ positive and high). Then*:

$$\text{(i)} \quad \begin{cases} q(\theta) = q^{\mathrm{F}}(\theta) & \text{for} \quad \theta = \theta^* \\ q(\theta) > q^{\mathrm{F}}(\theta) & \text{for} \quad \theta \in [\theta_0, \theta^*[\\ q(\theta) < q^{\mathrm{F}}(\theta) & \text{for} \quad \theta \in]\theta^*, \theta_1]. \end{cases}$$

(ii) *If $\theta^* \in (\theta_0, \theta_1)$, the PC is binding at both extremes: $U(\theta_0) = U(\theta_1) = 0$. If $\theta^* = \theta_0$, the PC binds only at $\theta = \theta_1$. If $\theta^* = \theta_1$, the PC binds only at $\theta = \theta_0$.*

(iii) *The optimal contract is fully separating.*

The most interesting case is the one where $\theta^* \in (\theta_0, \theta_1)$, shown in Fig. 6. In this case we have the unusual result that information rents are highest for an *interior* θ and zero for both extremes ($U(\theta_0) = U(\theta_1) = 0$), and that the efficient allocation is distorted at both the top and the bottom of the distribution. As we see in Fig. 6, $U_\theta = U'(\theta)$ is decreasing in θ and is zero for θ^*. Hence, the agent has an incentive to understate θ for $\theta < \theta^*$ and to overstate θ for $\theta > \theta^*$, in sharp contrast with the case of Proposition 1. This explains why it is the middle types who obtain the information rents and why the production distortion is opposite to that of Proposition 1.

4.2. $\bar{v}(\theta)$ *Concave*

This case is the one analyzed in Lewis and Sappington [9]. Although our results coincide with theirs, we report our analysis in this paper because this case stands in an interesting relation with the cases examined above, and because we use a different methodology.

In this case we have $\hat{\mu}'(\theta) < 0 < f(\theta)$ for all θ (this can be shown using the hazard rate assumptions). However, Lemma 3 does not apply and the production schedule $\hat{q}(\hat{\mu}(\theta), \theta)$ is increasing in θ for $\theta \in \Theta_2$. This implies that if $\hat{\mu}(\theta)$ crosses either $F(\theta)$ or $F(\theta) - 1$, the procedure used in the previous section does not yield the optimum and we have to impose the mononicity condition explicitly in the problem.

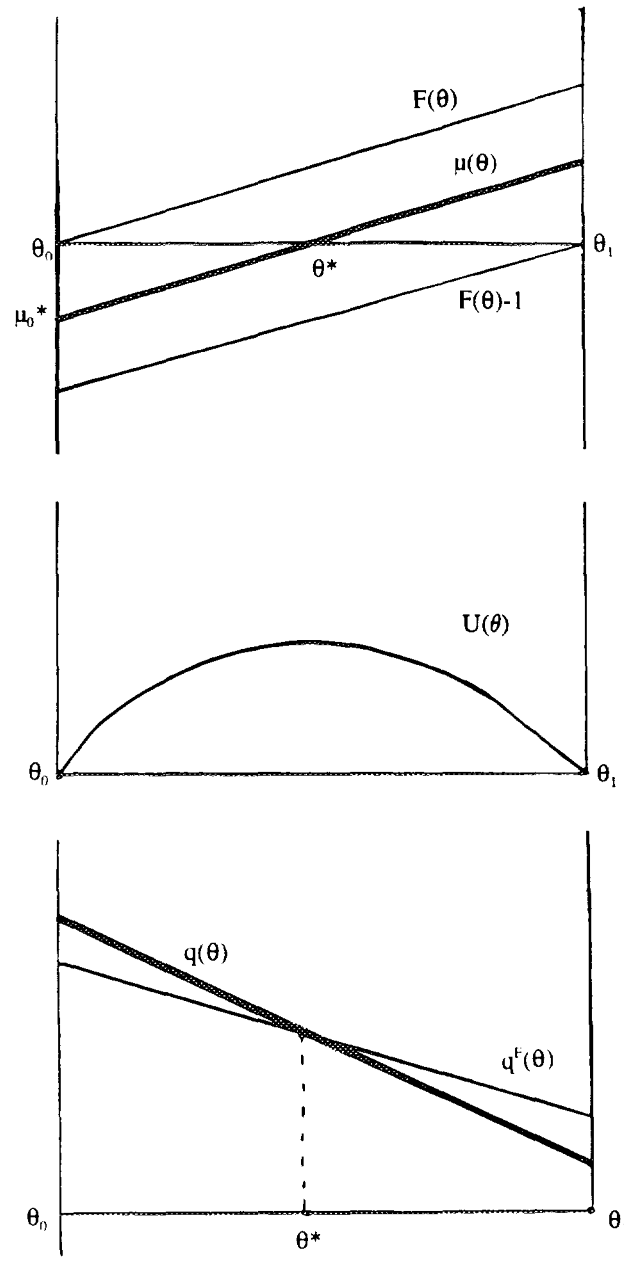


FIGURE 6

We will treat q as a state variable together with U, and $x \equiv dq/d\theta$ as the control variable. The optimization problem now becomes

$$(\text{P2}) \qquad \max_{x(\theta),\, q(\theta),\, U(\theta)} \int_{\theta_0}^{\theta_1} [V(q(\theta), \theta) - U(\theta)]\, f(\theta)\, d\theta$$

$$\text{s.t. } dq/d\theta = x \tag{8}$$

$$U'(\theta) = U_\theta \tag{9}$$

$$x(\theta) \leqslant 0, \quad U(\theta) \geqslant 0 \qquad \text{for all} \quad \theta \in \Theta.$$

The Hamiltonian for this optimal control problem is

$$H = [V(q, \theta) - U]\, f(\theta) - \mu[\bar{v}'(\theta) + q] + \lambda x,$$

and the Lagrangian is

$$L = H + \tau U,$$

where λ is the costate variable associated with the state q.
The maximization of the Hamiltonian implies

$$\lambda x = 0, \quad \lambda \geqslant 0, \quad x \leqslant 0. \tag{10}$$

The costate equations (which apply only where $\mu(\theta)$ and $\lambda(\theta)$ are differentiable) are

$$d\mu/d\theta = -\partial L/\partial U = f(\theta) - \tau(\theta) \tag{11}$$

$$d\lambda/d\theta = -\partial L/\partial q = -(V_q(q, \theta)\, f(\theta) - \mu(\theta)), \tag{12}$$

with

$$\tau(\theta) U(\theta) = 0, \quad \tau(\theta) \geqslant 0, \quad \text{and} \quad U(\theta) \geqslant 0 \qquad \text{for all } \theta \tag{13}$$

and the transversality conditions

$$\lambda(\theta_0) = \lambda(\theta_1) = 0 \tag{14}$$

$$\mu(\theta_0) U(\theta_0) = 0, \quad \mu(\theta_0) \geqslant 0 \tag{15a}$$

$$\mu(\theta_1) U(\theta_1) = 0, \quad \mu(\theta_1) \leqslant 0. \tag{15b}$$

Again, from the sufficient conditions in Seierstad and Sydsaeter [13], we know that if the configuration $(x(\theta), q(\theta), U(\theta), \mu(\theta), \lambda(\theta), \tau(\theta))$ satisfies (8)–(15), then $(x(\theta), q(\theta), U(\theta))$ solves P2. However, there may be no solution to this problem with a continuous $\mu(\theta)$. In this case, we have to allow for jumps in $\mu(\theta)$. Jumps may occur in $\mu(\theta)$ where the state constraint is binding ($U=0$) and $dU/d\theta=0$, and such jumps must be downward. That is, if there is a jump at θ_m, then

$$\mu(\theta_m^-)-\mu(\theta_m^+)>0, \tag{16}$$

where $\mu(\theta_m^-)$ and $\mu(\theta_m^+)$ are the limits of $\mu(\theta)$ as θ approaches θ_m from the left and the right, respectively.

If the schedule $\hat{\mu}(\theta)$ intersects $F(\theta)$ or $F(\theta)-1$, the solution cannot involve $\mu(\theta)=F(\theta)$ or $F(\theta)-1$, because either the transversality conditions or the PC would be violated. Nor can $\mu(\theta)$ coincide with $\hat{\mu}(\theta)$ for a whole interval because the allocation $\hat{q}(\hat{\mu}(\theta), \theta)$ violates condition (ii) of Lemma 1. Therefore there cannot be an interval where the PC is binding. It turns out that the only way to meet the transversality condition and the PC is for $\mu(\theta)$ to follow $F(\theta)$ for low θ's, to jump down to $F(\theta)-1$ at some interior point, and then to follow $F(\theta)-1$ until $\theta=\theta_1$.

The complete characterization of the optimal solution is given in the Appendix, in Lemma 5. The following proposition summarizes the structure of the optimal contract:

PROPOSITION 3. *Suppose* $\bar{v}''<0$. *Then there exist* θ_p, θ_q, *and* θ^* *with* $\theta_p\leqslant\theta_q$ *and* $\theta^*\in[\theta_p, \theta_q]$ *such that*:

(i) *If* $\theta_p<\theta_q$, *then*

$$\begin{cases} q(\theta)=q^F(\theta) & \text{for} \quad \theta\in\{\theta_0, \theta^*, \theta_1\} \\ q(\theta)<q^F(\theta) & \text{for} \quad \theta_0<\theta<\theta^* \\ q(\theta)>q^F(\theta) & \text{for} \quad \theta^*<\theta<\theta_1. \end{cases}$$

If $\theta_p=\theta_q$, *then either* $\theta_p=\theta_q=\theta_0$, *in which case* $q(\theta)>q^F(\theta)$ *for* $\theta<\theta_1$ *and* $q(\theta)=q^F(\theta)$ *for* $\theta=\theta_1$, *or* $\theta_p=\theta_q=\theta_1$, *in which case* $q(\theta)<q^F(\theta)$ *for* $\theta>\theta_0$ *and* $q(\theta)=q^F(\theta)$ *for* $\theta=\theta_0$.

(ii) *The PC is binding only at* θ_m.

(iii) $q(\theta)=q^*$ *for all* θ *in* $[\theta_p, \theta_q]$.

Proposition 3 reproduces the results of Lewis and Sappington [9]; therefore we refer the reader to their paper for an extensive interpretation of the results. Figure 7 illustrates the optimal contract in the case $\bar{v}''<0$ and $\theta_0<\theta_p<\theta_q<\theta_1$.

4.3. *Extension to More General Utility Functions*

In this section we extend the analysis in two ways. First, we derive an index-free condition (i.e., one that is independent of the specific

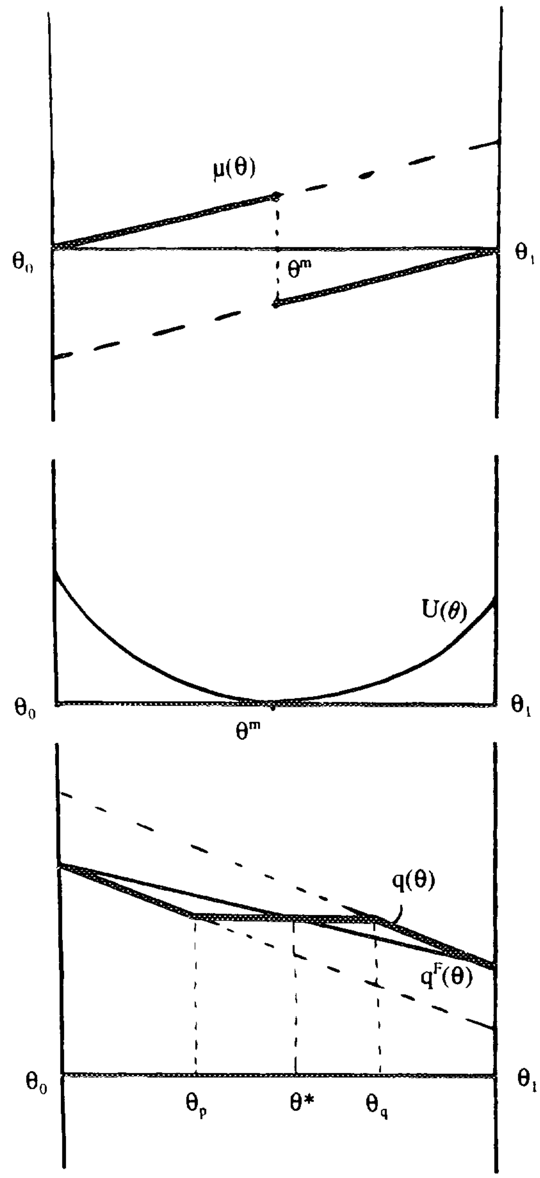


FIGURE 7

parameterization of the technology) for the contract to be fully separating. Second, we generalize the results to the case where the agent's utility is of the more general form $y - C(q, \theta)$, with $C_q > 0$, $C_{qq} > 0$, $C_\theta > 0$, and $C_{q\theta} > 0$ (this last assumption is the standard single-crossing property).

If we assume that the adjusted welfare is concave in q, that is $(\partial^2/\partial q^2)(W(q) f(\theta) - C(q, \theta) f(\theta) - \mu C_\theta(q, \theta)) < 0$ for all $\mu \in [F(\theta) - 1, F(\theta)]$ and for all $\theta \in \Theta$, then the analysis extends in an obvious way. The only amendment that deserves emphasis concerns the occurrence of pooling versus separation. As before, whether the optimal contract is pooling or separating depends on whether the monotonicity condition $q'(\theta) < 0$ is satisfied when $\mu = \hat{\mu}(\theta)$. We have

$$(d/d\theta)\, \hat{q}(\hat{\mu}(\theta), \theta) = (\partial\hat{q}/\partial\mu)\, \hat{\mu}'(\theta) + \partial\hat{q}/\partial\theta, \tag{17}$$

where $\hat{q}(\mu, \theta)$ is still given implicitly by (1) but now $V(q, \theta) \equiv W(q) - C(q, \theta) - \bar{v}(\theta)$. From the implicit definition of $\hat{\mu}(\theta)$, that is $-C_\theta[\hat{q}(\hat{\mu}, \theta), \theta] - \bar{v}'(\theta) = 0$, we can obtain

$$\hat{\mu}'(\theta) = -\frac{(C_{\theta\theta} + \bar{v}''(\theta))/C_{q\theta} + \partial\hat{q}/\partial\theta}{\partial\hat{q}/\partial\mu}. \tag{18}$$

Plugging (18) into (17), we get

$$(d/d\theta)\, \hat{q}(\hat{\mu}(\theta), \theta) = -(C_{\theta\theta} + \bar{v}''(\theta))/C_{q\theta}. \tag{19}$$

Now recall that when $q = \hat{q}(\hat{\mu}(\theta), \theta)$ we have $U'(\theta) = -(C_\theta + \bar{v}') = 0$, by the definition of $\hat{\mu}(\theta)$. Therefore, the necessary condition to have separation in the optimal contract is

$$C_{\theta\theta} + \bar{v}''(\theta)\,|_{C_\theta + \bar{v}' = 0} > 0,$$

or equivalently

$$U_{\theta\theta}\,|_{U_\theta = 0} < 0. \tag{*}$$

But condition (*) is equivalent to the condition that the agent's net utility is strictly quasiconcave in θ. If we suppose that U is either strictly quasiconcave or quasiconvex in θ for all θ, we can say that the presence of countervailing incentives yields pooling only if U is quasiconvex in θ, and implies that the equilibrium is separating if U is strictly quasiconcave in θ.

5. Conclusion

In this paper we have extended previous work on optimal contracts under countervailing incentives in the direction of more general utility functions for the agent. We have shown that the structure of the optimal contract, and in particular the occurrence of pooling versus separation, depends crucially on whether the agent's net utility is quasiconcave or quasiconvex in the private parameter. Pooling may obtain only if U is quasiconvex; separation obtains if U is strictly quasiconcave. Therefore pooling cannot be considered an intrinsic consequence of the presence of countervailing incentives.

We also found that if U is quasiconcave, the optimal contract can have one of two different structures. In one class of cases, information rents are decreasing for low types, equal to zero for an intermediate interval of types, and increasing for high types, production being distorted in the same qualitative way as in Lewis and Sappington [10]. In another class of cases, information rents are bell-shaped, with both extreme types earning no rents, and production is distorted in the opposite way relative to the Lewis and Sapington [10] solution. Moreover, only an interior type receives the efficient allocation in this case: the contract entails distortions at both extremes of the distribution.

APPENDIX

Proof of Lemma 2. The utility profile given in the lemma satisfies the PC for all θ since PC is binding for $\theta \in \Theta_2$ and $dU/d\theta < 0$ for $\theta \in \Theta_1$ and $dU/d\theta > 0$ for $\theta \in \Theta_3$. The sufficient conditions for the optimization problem are all satisfied by construction. Q.E.D.

Proof of Lemma 3. We have

$$d\hat{q}/d\theta = (\partial\hat{q}/\partial\theta) + (\partial\hat{q}/\partial\mu)(d\mu/d\theta)$$

and

$$\mu'(\theta)\,|_{\theta\in\Theta_2} = \hat{\mu}'(\theta) = -\frac{\bar{v}''(\theta) + \partial\hat{q}/\partial\theta}{\partial\hat{q}/\partial u}.$$

Therefore

$$d\hat{q}/d\theta\,|_{\theta\in\Theta_2} = -\bar{v}''(\theta) \leqslant \theta. \qquad \text{Q.E.D.}$$

Proof of Lemma 4. First note that $\hat{\mu}'(\theta) > f(\theta)$ implies that $\mu_0 + F(\theta)$ intersects $\hat{\mu}(\theta)$ at most once, say at θ'' (if there is no intersection, that is, if $\hat{\mu}(\theta) > F(\theta)$ or $\hat{\mu}(\theta) < F(\theta) - 1$ for all θ, then it is easy to see that $\mu_0 = 0$ or -1 as in the traditional case, were information rents are positive except at one extreme of the support), and lies above $\hat{\mu}(\theta)$ for $\theta < \theta''$ and below it for $\theta > \theta''$. Let

$$J(\theta;\, \mu_0) \equiv -\int_{\theta_0}^{\theta} [\bar{v}'(s) + \hat{q}(\mu_0 + F(s),\, s)]\, ds.$$

To prove the lemma we first need to show that $J(\theta;\, \mu_0)$ is quasiconcave in θ or equivalently that $J''(\theta;\, \mu_0)\,|_{J'(\theta;\, \mu_0)=0} \leqslant 0$. Note first that $J'(\theta;\, \mu_0) = 0$ only when $\bar{v}'(\theta) + \hat{q}(\mu_0 + F(\theta),\, \theta) = 0$, which in turn happens only when $\mu_0 + F(\theta) = \hat{\mu}(\theta)$. Therefore, $J'(\theta;\, \mu_0) = 0$ only for $\theta = \theta''$. Quasiconcavity follows because $J'(\theta;\, \mu_0)$ is decreasing in θ in a neighborhood of θ''. Necessarily $U(\theta) = J(\theta;\, \mu_0) + U(\theta_0)$, and hence, by quasiconcavity, $U(\theta) \geqslant \min(U(\theta_0),\, U(\theta_1))$ for all θ. Therefore, if both $U(\theta_0)$ and $U(\theta_1)$ are nonnegative the PC is satisfied for all θ.

Now consider separately the three cases:

(a) $R(0) \leqslant 0$: if we set $\mu_0 = 0$ and $U(\theta_0) = -R(0) = -J(\theta_1;\, 0)$, then $U(\theta_0) \geqslant 0$ and $U(\theta_1) = 0$, so the PC is satisfied for all θ. The transversality condition (TC) is also satisfied because $\mu(\theta_0) = 0$ and $U(\theta_1) = 0$.

(b) $R(-1) \geqslant 0$: setting $\mu_0 = -1$ and $U(\theta_0) = 0$, then $U(\theta_1) = J(\theta_1;\, -1) = R(-1) \geqslant 0$, so the PC is satisfied for all θ. The TC is also satisfied because $U(\theta_0) = 0$ and $\mu(\theta_1) = 0$.

(c) $R(-1) < 0 < R(0)$: if we choose μ_0 such that $R(\mu_0) = 0$ and we set $U(\theta_0) = 0$, then $U(\theta_1) = J(\theta_1;\, \mu_0) = R(\mu_0) = 0$, and hence the PC is satisfied for all θ. The TC is also satisfied since $U(\theta_0) = U(\theta_1) = 0$. Q.E.D.

LEMMA 5. *Assume $\bar{v}''(\theta) < 0$ for all θ. Then the optimal contract involves*

$$\mu(\theta) = \begin{cases} F(\theta) & \text{for } \theta < \theta_m \\ F(\theta) - 1 & \text{for } \theta > \theta_m \end{cases}$$

$$\lambda(\theta) = \begin{cases} 0 & \text{for } \theta \notin [\theta_p, \theta_q] \\ -\int_{\theta_p}^{\theta} [V_q(q^*, s)\, f(s)]\, ds & \text{for } \theta \in [\theta_p, \theta_m] \\ -\int_{\theta_p}^{\theta_m} [V_q(q^*, s)\, f(s) - F(s)]\, ds - \int_{\theta_m}^{\theta} [V_q(q^*, s)\, f(s) - (F(s) - 1)]\, ds & \text{for } \theta \in [\theta_m, \theta_q] \end{cases}$$

$$q(\theta) = \begin{cases} \hat{q}(F(\theta),\, \theta) & \text{for } \theta < \theta_p \\ q^* & \text{for } \theta \in [\theta_p, \theta_q] \\ \hat{q}(F(\theta) - 1,\, \theta) & \text{for } \theta > \theta_q \end{cases}$$

$$U'(\theta) = -\bar{v}'(\theta) - q(\theta), \quad \text{with} \quad U(\theta_m) = 0,$$

where q^*, θ_p, θ_q, *and* $\theta_{\rm m}$ *satisfy the following conditions*:

$$\bar{v}'(\theta_{\rm m})+q^*\begin{cases}\geqslant 0 & \textit{if}\ \ \theta_{\rm m}=\theta_{\rm p}\\ =0 & \textit{if}\ \ \theta_{\rm m}\in(\theta_p,\theta_q)\\ \leqslant 0 & \textit{if}\ \ \theta_{\rm m}=\theta_q\end{cases} \tag{A1}$$

$$\int_{\theta_p}^{\theta_{\rm m}}[V_q(q^*,\theta)\,f(\theta)-F(\theta)]\,d\theta+\int_{\theta_{\rm m}}^{\theta_q}[V_q(q^*,\theta)\,f(\theta)-(F(\theta)-1)]\,d\theta=0 \tag{A2}$$

$$V_q(q^*,\theta_p)\,f(\theta_p)-F(\theta_p)\begin{cases}=0 & \textit{if}\ \ \theta_p>\theta_0\\ \geqslant 0 & \textit{if}\ \ \theta_p=\theta_0\end{cases} \tag{A3a}$$

$$V_q(q^*,\theta_q)\,f(\theta_q)-F(\theta_q)+1\begin{cases}=0 & \textit{if}\ \ \theta_q<\theta_1\\ \geqslant 0 & \textit{if}\ \ \theta_q=\theta_1\end{cases} \tag{A3b}$$

Proof of Lemma 5. We first show that a solution $(q^*, \theta_p, \theta_q, \theta_{\rm m})$ to conditions (A1)–(A3) exists. We do this by construction. Assume first that either $\theta_p>\theta_0$ or $\theta_q<\theta_1$. Without loss of generality, assume $\theta_p>\theta_0$. Then from (A3) we get $q^*(\theta_p)$ (which exists by the assumption in footnote 7) and $\theta_q(\theta_p)$ (which can be shown to exist). Using this in (A2) we get $\theta_{\rm m}(\theta_p)$, which can be shown to exist. Let $\xi(x)\equiv\bar{v}'(\theta_{\rm m}(x))+q^*(x)$. It is easy to verify that $\xi(x)$ is decreasing in x. Then let

$$\theta_p^*=\begin{cases}\theta_0 & \text{if}\ \ \xi(\theta_0)\leqslant 0\\ \xi^{-1}(0) & \text{if}\ \ \xi(\theta_0)>0\ \text{and}\ \xi(\theta_1)<0\\ \theta_1 & \text{if}\ \ \xi(\theta_1)\geqslant 0.\end{cases}$$

If $\theta_p^*>\theta_0$ then $\theta_p=\theta_p^*$, $q^*=q^*(\theta_p^*)$, $\theta_q=\theta_q(\theta_p^*)$, and $\theta_{\rm m}=\theta_{\rm m}(\theta_p^*)$. If $\theta_p^*=\theta_0$, then in the solution we have $\theta_p=\theta_0$. Next assume that $\theta_q<\theta_1$. With the same procedure we find a value θ_q^*. If $\theta_q^*<\theta_1$, then θ_q^* gives the complete solution. However, if $\theta_q^*=\theta_1$ we now have $\theta_p=\theta_0$ or $\theta_q=\theta_1$. We still have to determine q^* and $\theta_{\rm m}$. First fix $\theta_{\rm m}$. From (A2) we get $q^*(\theta_{\rm m})$ (with necessarily exists by the assumption in footnote 7) and plugging into (A1) we get an equation in $\theta_{\rm m}$ which gives the solution for $\theta_{\rm m}$. q^* is obtained from the function $q^*(\theta_{\rm m})$.

Now it is easy to verify that the solution proposed in Lemma 5 satisfies the sufficient conditions (9)–(15). Q.E.D.

References

1. W. J. Adams and G. L. Yellen, Commodity in bundling in the burden of monopoly, *Quart. J. Econ.* **90** (1976), 475–498.
2. S. L. Brainard and D. Martimort, Strategic trade policy with ignorant policymakers, mimeo.

3. P. Champsour and J. Rochet, Multiproduct duopolists, *Econometrica* **57** (1989), 533–557.
4. R. Feenstra and T. Lewis, Negotiated trade restrictions with private political pressure, *Quart. J. Econ.* **106** (1991), 1287–1308.
5. D. Fudenberg and J. Tirole, "Game Theory," MIT Press, Cambridge, MA, 1991.
6. R. Guesnerie and J. J. Laffont, 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** (1984), 329–369.
7. J. M. Lacker and J. A. Weinberg, Optimal contracts under costly state falsification, *J. Polit. Econ.* **97** (1989), 1345–1363.
8. J. J. Laffont, "The economics of uncertainty and information," MIT Press, Cambridge, MA, 1989.
9. T. Lewis and D. Sappington, Countervailing incentives in agency problems, *J. Econ. Theory* **49** (1989), 294–313.
10. T. Lewis and D. Sappington, Inflexible rules in incentive problems, *Amer. Econ. Rev.* **79** (1989), 69–84.
11. G. Maggi and A. Rodríguez-Clare, Costly distortion of information in agency problems, *Rand*, in press.
12. J. Moore, Optimal labor contracts when workers have a variety of privately observed reservation wages, *Rev. Econ. Stud.* **52** (1985), 37–68.
13. A. Seierstad and K. Sydsaeter, Sufficient conditions in optimal control theory, *Int. Econ. Rev.* **18** (1977), 367–391.