Journal of Public Economics 37 (1988) 243–260. North-Holland

EXTERNAL MARKETS AND REGULATION

James J. ANTON

Department of Economics, State University of New York at Stony Brook, Stony Brook, NY 11794, USA

Paul J. GERTLER*

The RAND Corporation, 1700 Main Street, Santa Monica, CA 90406, USA

Received December 1986, revised version received June 1988

Regulated firms often sell to 'external' markets in addition to their regulated or 'internal' markets. The welfare of consumers in these 'external' markets is typically outside of the regulator's domain of concern. As a result, 'external' markets provide the regulator with additional policy options for the strategic influence of firm behavior in the presence of asymmetric information. Without an 'external' market, asymmetric information introduces an incentive cost and, consequently, the optimal policy is forced to distort 'internal' production below the first-best level. We show that profit opportunities presented by an 'external' market can be used to absorb some (all, when the firm is an 'external' price-taker) of the incentive cost, and thus insulate the 'internal' market from quantity distortions.

1. Introduction

Regulated firms often sell to 'external' markets in addition to their regulated or 'internal' markets. The welfare of consumers in these 'external' markets is typically outside of a regulator's domain of concern. As a result, 'external' markets provide regulators with additional policy options for the strategic influence of firm behavior in the presence of asymmetric information. Without an 'external' market, asymmetric information introduces an incentive cost and, consequently, the optimal policy is forced to distort 'internal' production below the first-best level [Baron and Myerson (1982), Sappington (1983), Baron and Besanko (1984) and Laffont and Tirole (1986)]. We show that profit opportunities presented by an 'external' market can be used to absorb some (all, when the firm is an 'external' price taker) of the incentive cost, and thus insulate the 'internal' market from quantity distortions.

*We are indebted to Robert Porter, Robert Rosenthal, Warren Sanderson, Yoram Weiss, the referees, and the workshop participants in the Economics Departments at the State University of New York at Stony Brook and at the University of Wisconsin at Madison for their helpful comments. Any errors are our responsibility.

0047-2727/88/$3.50

There are many examples of regulated firms selling in 'external' markets. For instance, some electric power and water utilities sell excess capacity to other more constrained utilities. In 1985, Quebec's electric power utility completed an agreement to sell excess capacity to New England. In these cases, regulators set the utility's 'internal' price and 'external' sales in accordance with the welfare of 'internal' consumers. In another context, national industries, such as petroleum companies in oil-exporting nations like the U.K., the Gulf States, Mexico, and Norway, sell in both domestic (internal) and international (external) markets. Clearly, these governments are more concerned with domestic than with international welfare. In these examples, the regulators control firms' sales to 'external' markets.

In other settings, regulators cannot control the actions of firms in 'external' markets. In the United States, for example, the Medicare and Medicaid programs create and regulate prices in public patient (internal) markets for hospitals, nursing homes, and physicians. The prices in private patient (external) markets are not subject to regulatory control. Also, foreign-based multinational corporations' 'external' operations are beyond direct regulatory control. Of course, firm control of 'external' production is not a problem if the regulator can monitor effectively and employ forcing contracts to implement its desired policy (e.g. quotas, embargos, and tariffs).

We consider optimal regulatory policy towards a firm that is a monopolist in an 'internal' market and can participate in an 'external' market. The firm is assumed to have private information about its production costs that the regulator cannot directly observe. The case of pure adverse selection, where the regulator is assumed to control the firm's behavior in both markets as well as impose a lump-sum tax (subsidy), is analyzed first. When the regulator is not able to control or even observe 'external' production, a moral hazard is introduced since the firm will strategically choose 'external' production. This case is also analyzed.

2. First-best regulatory policy

2.1. Assumptions and notation

Consider regulating a single-product firm that acts as a monopolist in an 'internal' market and can participate in an 'external' market. Total output Q is divided into 'internal' production q and 'external' production q^e. The firm faces an 'internal' demand function $P(q)$ which is downward sloping, and an 'external' demand function or, equivalently, an 'external' revenue function $R^e(q^e)$. The 'external' marginal revenue function $MR^e(q^e)$ can be flat (i.e. the firm is a price-taker) or downward sloping. Total costs are $\theta C(Q)$, where total production Q is the sum of q and q^e, and $\theta \in [\underline{\theta}, \bar{\theta}]$ is a firm-specific cost parameter. In the first-best regulatory problem, the government knows the value of θ ex ante.

The firm is assumed to be a profit-maximizer, with profits given by:

$$\Pi = P(q)q + R^e(q^e) - \theta C(Q) - T, \tag{1}$$

where

$P(\cdot)$ = 'internal' inverse demand function,
q = 'internal' quantity,
$R^e(\cdot)$ = 'external' revenue function,
q^e = 'external' quantity,
θ = firm specific cost parameter,
$\theta C(\cdot)$ = total costs,
$Q = q + q^e$ (total output),
T = lump-sum tax.

The government maximizes the welfare of consumers in the 'internal' market, which is defined as the sum of consumers' surplus and tax revenue. It chooses (q, q^e, T) to maximize

$$W = V(q) - P(q)q + T \tag{2}$$

subject to

$$\Pi = P(q)q + R^e(q^e) - \theta C(Q) - T \geqq 0, \tag{3}$$

where $V(q) = \int_0^q P(t)\,dt$.[1] The constraint in (3) guarantees the firm non-negative profits. Clearly, the solution depends on the cost parameter θ.

2.2. Solution

We assume that a unique solution to the first-best problem exists and that the 'internal' market is always active.[2] The necessary and sufficient Kuhn–Tucker first-order conditions for welfare maximization are:

[1]The welfare function in (2) is standard. Several extensions are possible. Some formulations place a positive weight on the firm's profits. Also, Laffont and Tirole (1986) consider introducing a parameter λ that allows the marginal welfare of revenues to differ from unity. Both extensions are straightforward and do not alter the substance of our results. The net effect is that the marginal cost of information is multiplied by a constant. With a weight on firm's profits, for instance, this constant is less than one and, consequently, quantity distortions are reduced.

[2]Formal conditions for existence and uniqueness follow directly from geometric intuition. To ensure that the 'internal' market is always active we assume that $P(0)$ is greater than both $MR^e(0)$ and $\bar{\theta}C'(0)$, and that the fixed costs of the highest cost type, $\bar{\theta}C(0)$, are not large enough to justify shutting down the firm. Furthermore, all marginal cost curves (as θ varies) cross $P(q)$ from below and if they cross $MR^e(q^e)$ they do so from below. Note that the crossing properties are automatic if $V(q) + R^e(q^e) - \theta C(q + q^e)$ is concave in (q, q^e) for each θ. To avoid the trivial case where q^e is always zero or always positive, we assume that $P(q)$ eventually falls below $MR^e(0)$, say at q^f, and that $\theta^f \equiv MR^e(0)/C'(q^f) \in (\underline{\theta}, \bar{\theta})$.

$$P(q)=\theta C'(Q), \tag{4}$$

$$MR^{e}(q^{e})\leqq\theta C'(Q), \quad \text{with equality if } q^{e}>0, \tag{5}$$

and (3) with equality. From (4), 'internal' quantity $q^{f}(\theta)$ is chosen so that 'internal' price equals marginal cost and, from (5), 'external' quantity $q^{ef}(\theta)$ is chosen so that 'external' marginal revenue does not exceed marginal cost. It is straightforward to verify that these quantities are non-increasing in θ. Finally, the optimal tax $T^{f}(\theta)$ is determined by (3)–(5). It sets profits to zero and is non-increasing in θ.

Intuitively, the optimal policy maximizes 'internal' consumers' surplus by marginal cost pricing, and maximizes 'external' profits by setting marginal revenue equal to marginal cost whenever 'external' production is positive. Through the tax, the government uses 'external' profits to subsidize the 'internal' market. These profits may be used to cover fixed costs when necessary or may be taken in the form of tax revenue.

Formally, condition (5) requires that $q^{ef}(\theta)$ maximizes 'external' profits for any given values of q and θ. Let $r(q,\theta)$ denote this reaction function. Clearly, $r(q,\theta)=0$ when $MR^{e}(0)\leqq\theta C'(q)$. Otherwise, $r(q,\theta)$ solves (5) with equality.

Substituting the reaction function for q^{e} in (4), the first-best 'internal' quantity, $q^{f}(\theta)$, solves

$$P(q)=\theta C'(q+r(q,\theta)). \tag{6}$$

Let θ^{f} be the lowest cost type that is not directed to sell in the 'external' market. Formally, θ^{f} solves

$$MR^{e}(0)=\theta C'(q^{f}(\theta)). \tag{7}$$

Then, for 'high cost' types, $\theta\geqq\theta^{f}$, the optimal policy sets 'external' production to zero, and for 'low cost' types, $\theta<\theta^{f}$, 'external' production is positive.

There is a welfare gain from directing the firm to sell in the 'external' market only if the 'external' market is profitable. This occurs for 'low cost' types since the marginal revenue from the first 'external' unit, $MR^{e}(0)$, is greater than the 'internal only' marginal cost. For a 'high cost' type no welfare gains are available in the 'external' market.

The markets are linked via the cost function. For instance, when decreasing marginal costs are present, production in the 'external' market allows the firm to exploit scale economies, and thus increase both 'internal' production and the tax. In the case of constant marginal costs, conditions (4) and (5) imply that the 'internal' and 'external' quantity decisions are independent of one another.

3. Non-implementability of the first-best

In the first-best, the government sets policy based on the observed value of the firm's cost parameter θ. When the actual value of θ is private information of the firm, the government does not observe θ. Therefore, if it wishes to implement the first-best policy, it must obtain cost information from the firm.[3] Since the firm is aware of this policy rule, it can achieve higher profits by misrepresenting the value of its cost parameter, and as a result, circumvent the effectiveness of regulation.

Formally, the firm reports a cost parameter and the government uses this report to set its policy according to the first-best rules. Essentially, the firm is faced with a menu of policy options from which it chooses by specifying its cost report. The choice is made so as to maximize profits. In the case of *pure adverse selection*, the government cannot observe θ but can control both 'internal' and 'external' output. When the government loses control of 'external' output, a *moral hazard* is introduced since the firm will choose 'external' output strategically.

3.1. Pure adverse selection

Suppose the government does not observe θ. Under the first-best policy, if the firm reports $\hat{\theta}$ and its actual cost parameter is θ, profits are:

$$\pi(\hat{\theta}|\theta) = P(q^{f}(\hat{\theta}))q^{f}(\hat{\theta}) + R^{e}(q^{ef}(\hat{\theta})) - \theta C(Q^{f}(\hat{\theta})) - T^{f}(\hat{\theta}). \tag{8}$$

Notice that the revenue and tax terms depend only on the cost report, whereas the firm's costs also depend on its actual type. By adding zero to (8) in the form of $(\hat{\theta}-\hat{\theta})C(Q^{f}(\hat{\theta}))$, we can express the profit function as:

$$\pi(\hat{\theta}|\theta) = \Pi(\hat{\theta}) + (\hat{\theta}-\theta)C(Q^{f}(\hat{\theta})). \tag{9}$$

The first term in (9) is profit under the first-best policy for a type $\hat{\theta}$ and is zero. The second term is the gain from reporting $\hat{\theta}$ when the actual cost parameter is θ and is the difference between reported and actual costs.

Since the first term in (9) is zero, the profit-maximization problem reduces to choosing $\hat{\theta}$ to maximize the second term. Moreover, since the cost of production is positive, $C(\cdot)$ is increasing in Q, and $Q^{f}(\cdot)$ is non-increasing in θ, the firm will always report a cost parameter higher than its actual type. Thus, the firm always pretends it has higher than actual production costs.

The source of the gain from over-reporting depends on the firm's actual θ. Specifically, a 'high cost' firm pretends that it is a higher 'high cost' firm and, consequently, increases profits through both a reduction in taxes, and an

[3] Additionally, it may be possible to conduct a cost audit or otherwise monitor the firm as in Baron and Besanko (1984).

increase in 'internal' price. Similarly, some 'low cost' types increase their profits by pretending that they are 'high cost' types. By reporting a 'high cost' $\hat{\theta}$, these 'low cost' firms give up the 'external' market and derive all profit from the 'internal' market. These are the types closer to θ^{f}.

A type with a very low value of θ will over-report but not beyond θ^{f}.[4] By submitting a 'low cost' report, these types still participate in both markets. Thus, they trade-off further increases in the 'internal' price for continued access to the 'external' market. For these types, the opportunity cost of 'external' market profits is greater than potential profits from 'internal only' profit maximization.

In summary, the firm subverts some of the welfare gain from regulation by overstating its cost type. This distortion is qualitatively different for the very 'low cost' firms due to their greater ability to earn profits in the 'external' market. As we show later, this differential profit incentive can be used to mitigate the welfare loss from adverse selection.

3.2. *Adverse selection plus moral hazard*

Suppose, in addition to the unobservability of θ, the firm and not the government controls 'external' production. Then, if a type θ firm reports $\hat{\theta}$ and optimally sets 'external' production at $r(q^{\mathrm{f}}(\hat{\theta}),\theta)$, profits reduce to:

$$\pi(\hat{\theta}|\theta)=\Pi(\hat{\theta})+\Pi^{\mathrm{e}}(q^{\mathrm{f}}(\hat{\theta}),\theta)-\Pi^{\mathrm{e}}(q^{\mathrm{f}}(\hat{\theta}),\hat{\theta}), \tag{10}$$

where

$$\Pi^{\mathrm{e}}(q,\theta)\equiv\max_{q^{\mathrm{e}}}\,[R^{\mathrm{e}}(q^{\mathrm{e}})-\theta C(q+q^{\mathrm{e}})]. \tag{11}$$

The profit differential in (10) reflects the added incentive of the firm to set 'external' production optimally in response to $q^{\mathrm{f}}(\hat{\theta})$. When the firm controls 'external' production, 'external' revenues are determined by the reaction function. In contrast to the pure adverse-selection case, government policy does not directly determine 'external' revenues but only indirectly influences them through the reaction function $r(q^{\mathrm{f}}(\hat{\theta}),\theta)$. Thus, the incentive to misreport in (10) depends on the revenue differential in the 'external' market as well as the cost differential. Since $\Pi^{\mathrm{e}}(q^{\mathrm{f}}(\hat{\theta}),\theta)$ is decreasing in θ, the firm will always over-report.

The net impact of moral hazard on the incentive to overstate costs [comparing (9) to (10)] depends on returns to scale in the cost function.

[4]Consider the following example. Let $P(q)=1-q$, $C(Q)=Q^2/2$, $\theta=1/4$, and suppose that the firm is a price-taker in the 'external' market, where $p^{\mathrm{e}}=MR^{\mathrm{e}}=2/3$. Then $\theta^{\mathrm{f}}=2$ and direct calculation reveals that profits are maximized by reporting $\theta=1/2$. Hence, this 'low cost' type reports a $\hat{\theta}$ greater than θ but less then θ^{f}.

When marginal cost rises with output ($C''>0$), it is straightforward to verify that the derivative in $\hat{\theta}$ of $\pi(\hat{\theta}|\theta)$ is always larger in (10) than it is in (9). This means that when $\pi(\hat{\theta}|\theta)$ reaches a maximum in (9) it is still increasing in (10). Therefore, moral hazard leads to a greater overstatement of costs. Since 'internal' production falls with $\hat{\theta}$ when $C''>0$, the firm can operate in the 'external' market with lower marginal costs.

When marginal costs fall with output, this incentive is reversed as larger 'internal' production reduces marginal costs. The firm still over-reports, but the net impact of moral hazard is ambiguous.

4. Second-best regulatory policy: Pure adverse selection

4.1. The framework

The government is placed in a second-best setting when the firm has private cost information. In order to implement a regulatory policy that responds to the actual cost parameter, the government must obtain information from the firm. As was shown in section 3, the first-best policy cannot be implemented because the firm has an incentive to misreport the value of its cost parameter. Any attempt to implement the first-best policy results in the adoption of prices and quantities that are appropriate for the reported cost parameter instead of the actual one. Hence, an optimal second-best policy must take into account the limitations created by adverse selection. In this section, we characterize the optimal policy for this second-best problem.

Following Baron and Myerson (1982), we employ a direct regulatory mechanism that specifies how the policy instruments are set in response to the cost report. The instruments are the 'internal' quantity, the 'external' quantity and a lump-sum tax. A policy is defined by three functions of the cost report $\Omega(\theta)=\{q(\theta), q^{e}(\theta), T(\theta)\}$. Regulation occurs in Stackelberg fashion. The government specifies a policy $\Omega(\cdot)$ and the firm responds with a cost report $\hat{\theta}$, thus determining specific values for prices, quantities, and the tax. The government has a prior probability distribution on the cost parameter and we assume that the density $f(\theta)$ is positive and differentiable on $\theta\in[\underline{\theta},\bar{\theta}]$.

The policy is chosen to maximize expected consumers' surplus in the 'internal' market plus tax revenue. Without loss in generality, we restrict the choice set to the class of incentive-compatible policies. A policy $\Omega(\cdot)$ is incentive-compatible if it gives the firm no incentive to misreport information. The Revelation Principle, as discussed by Dasgupta, Hammond and Maskin (1979) and Myerson (1979), guarantees that there is no loss in generality when the optimal policy is found by examining just the class of direct incentive-compatible policies.

The solution procedure is as follows. First, we define the set of feasible policies. Then, we find a characterization that simplifies the incentive-compatibility constraints so that they can be directly substituted into the government's objective function. A Kuhn–Tucker analysis then provides the optimal second-best policy.

4.2. Feasibility

A policy is feasible if and only if it is incentive-compatible and the firm is a willing participant in the regulated market. A policy is incentive-compatible if the firm has no incentive to misreport. Using (8) to define $\pi(\hat{\theta}|\theta)$ for a policy $\Omega(\cdot)$, a policy is incentive-compatible when

$$\text{(IC)} \qquad \Pi(\theta) \geqq \pi(\hat{\theta}|\theta)$$

for all reports $\hat{\theta}$ and types θ, where $\Pi(\theta) \equiv \pi(\theta|\theta)$. Under (IC), the firm finds it optimal to report the actual value of its privately observed cost parameter.

If the firm is to be a willing participant, then the policy must ensure that the firm earns non-negative profits. Since the government does not observe the firm's type ex ante, every potential type must be able to earn non-negative profits. Thus, we have the individual rationality condition:

$$\text{(IR)} \qquad \Pi(\theta) \geqq 0, \quad \text{for all } \theta.$$

Note that a firm is guaranteed non-negative profits only if it reports its actual type.

A policy $\Omega(\cdot)$ is feasible if it satisfies (IC) and (IR). By a standard argument (see the appendix), feasibility is equivalent to the conditions:

$$Q(\theta) = q(\theta) + q^{e}(\theta) \text{ is a non-increasing function in } \theta, \tag{12}$$

$$\Pi(\theta) = \Pi(\bar{\theta}) + \int_{\theta}^{\bar{\theta}} C(q(t) + q^{e}(t))\,\mathrm{d}t, \tag{13}$$

$$\Pi(\bar{\theta}) \geqq 0. \tag{14}$$

Condition (12) states that total output cannot be larger for less-efficient types. Condition (13) implies that the profit of a type θ firm is determined by backwards integration of the cost function, and is equal to the profit of the least efficient firm plus an incentive term which depends on the production of all higher types. This implies that more efficient firms are guaranteed higher profits, as the integral is decreasing in θ. Finally, by imposing the requirement that the highest cost type, $\bar{\theta}$, receives non-negative profits, condition (14) guarantees that all types earn non-negative profits.

The characterization of feasibility provided by (12)–(14) allows us to simplify the government's objective function:

$$W\langle\Omega\rangle = \int_{\theta}^{\bar{\theta}} [V(q(\theta)) - P(q(\theta))q(\theta) + T(\theta)] f(\theta)\,\mathrm{d}\theta,$$

which is expected 'internal' consumers' surplus plus tax revenue. As described in the appendix, the terms $-P(q(\theta))q(\theta)+T(\theta)$ can be substituted out of the objective function to obtain:

$$W\langle\Omega\rangle = \int_{\theta}^{\bar{\theta}} [V(q(\theta)) + R^{e}(q^{e}(\theta)) - \rho(\theta)C(q(\theta)+q^{e}(\theta))] \times f(\theta)\,\mathrm{d}\theta - \Pi(\bar{\theta}), \tag{15}$$

where $\rho(\theta) \equiv \theta + F(\theta)/f(\theta)$.

Notice that $\rho(\theta)$ replaces θ as the multiplicative factor on the cost function in (15). The weight $\rho(\theta)$ is always larger than θ since it includes the incentive term $F(\theta)/f(\theta)$, the ratio of the cumulative distribution to the density function. This ratio reflects the loss in consumers' surplus that arises from the cost of providing incentives for the firm to reveal private information. In fact, when multiplied by $C'(\cdot)$, the ratio becomes the marginal cost of information in the sense of a shadow price [see Baron and Besanko (1984)].

Thus, the problem reduces to choosing the 'internal' and 'external' production schedules, $q(\theta)$ and $q^{e}(\theta)$, to maximize (15) subject to the restriction in (12) that total quantity be non-increasing in θ. The integrand in (15) is strictly concave in the variables q and q^{e}. Therefore, the Kuhn–Tucker conditions are necessary and sufficient for q and q^{e} to maximize the integrand at a given θ. Under the simplifying assumption that $\rho(\theta)$ is strictly increasing in θ, this pointwise solution produces a total quantity that is non-increasing in θ, implying that the Kuhn–Tucker solution is global.[5]

4.3. Solution

Consider the pointwise program in θ and associated Kuhn–Tucker conditions:

[5]This is a standard regularity condition for models with private information. It is valid for several important distributions including the normal, beta, and uniform families. When it fails, the optimal schedule involves 'bunching' and quantity is held constant over one or more ranges of θ types. Theorem 4 in Guesnerie and Laffont (1984) provides conditions for determining the bunching intervals. The appendix discusses this issue in more detail.

$$\max_{q,q^e} V(q)+R^e(q^e)-\rho(\theta)C(q+q^e)$$

$$P(q)-\rho(\theta)C'(q+q^e)=0. \tag{16}$$

$$MR^e(q^e)-\rho(\theta)C'(q+q^e)\leqq 0, \quad \text{with equality if } q^e>0. \tag{17}$$

These Kuhn–Tucker conditions are identical to those of the first-best except that $\rho(\theta)$ has replaced θ. Therefore, the second-best 'internal' and 'external' quantities are given by $q^s(\theta)=q^f(\rho(\theta))$ and $q^{es}(\theta)=q^{ef}(\rho(\theta))$, respectively.

Marginal costs are higher in the second-best, since $\rho(\theta)>\theta$. Therefore, a comparison of (16) and (17) to (4) and (5) shows that total production is smaller in the second-best than in the first-best.

Condition (17) requires that 'external' production in the second-best, $q^{es}(\theta)$, maximizes 'external' profits for any given values of q and $\rho(\theta)$. Recall that this quantity is just $r(q,\rho(\theta))$, the firm's reaction function. Substituting into (16), we see that second-best 'internal' quantity, $q^s(\theta)$, solves

$$P(q)=\rho(\theta)C'(q+r(q,\rho(\theta))). \tag{18}$$

Let θ^s be defined by:

$$MR^e(0)=\rho(\theta)C'(q^s(\theta)). \tag{19}$$

Then, for 'high cost' types, $\theta\geqq\theta^s$, the optimal policy sets 'external' production to zero, and for 'low cost' types, $\theta<\theta^s$, 'external' production is positive.

Condition (18) is the first-best condition (6) for the 'internal' production choice with $\rho(\theta)$ replacing θ. Since $\rho(\theta)>\theta$ and $\theta C'(q+r(q,\theta))$ is an increasing function of θ, the second-best 'internal' quantity is no larger than the first-best 'internal' quantity.

The second-best critical value, θ^s, is less than the first-best critical value θ^f. This is easily confirmed by comparing (7) with (19) and recalling that $\rho(\theta)>\theta$. Hence, the 'external' market is used less often in the second-best policy than in the first-best. Intuitively, since the marginal cost of information is positive, some 'low cost' types in the first-best are treated as 'high cost' types in the second-best and kept out of the 'external' market.

4.4. Discussion

We have shown that, relative to the first-best, the effects of adverse selection are (i) to reduce total quantity for each cost type, (ii) to restrict 'external' market access to fewer cost types, and (iii) to reduce the availability of 'external' profits for subsidization. The distortion in total quantity is distributed between the two markets. When 'internal' quantity is distorted, the regulator experiences both a loss in consumers' surplus and a loss in lump-sum tax revenues. When 'external' quantity is distorted only tax revenues are affected. Therefore, for a given loss in tax revenues, the regulator prefers to impose the quantity distortions on 'external' production and avoid a loss in consumers' surplus.

We are interested in the extent to which the regulator can concentrate the distortion in total quantity on the 'external' market, and thus insulate the 'internal' market from the effects of adverse selection. Full insulation occurs when there are no quantity distortions in the 'internal' market.

Insulation is possible only with 'low cost' types, as 'high cost' types do not improve welfare by selling in the 'external' market. The degree of insulation depends on the linkage between the markets, which is summarized by the first-order conditions. Using (4) and (5) we rewrite the first-best conditions (suppressing θ as an argument) as:

$$P(q)=\theta C'(Q)=(1-(1/\eta))p^{e}(q^{e}), \tag{20}$$

where η is the price elasticity of 'external' demand. Recall that the second-best conditions are identical to (20) with $\rho(\theta)$ replacing θ, and that $\rho(\theta)$ is greater than θ. Hence, we can assess the distributional impact of private information in the context of a simple comparative static experiment in θ.

Consider the role of cost structure in the market linkage. As can be seen from (20), constant marginal costs imply that 'internal' and 'external' quantities are set independently of one another. In this case, the presence of an 'external' market does nothing to insulate the 'internal' market from private information distortions. In all other cost structures, marginal cost varies with production and consequently quantity choices are interdependent. This linkage allows the 'external' market to mitigate 'internal' quantity distortions.

The extent of the insulation then depends upon the size of the 'external' demand elasticity η. The 'internal' market is fully insulated when the firm is an 'external' price-taker. In this case, the value of η is $-\infty$ and the right-hand side of (20) reduces to a constant p^{e}. Now consider an increase in θ to $\rho(\theta)$. The 'internal' quantity remains fixed at the level that equates 'internal' price to p^{e}, which is the first-best 'internal' quantity, and all of the decline in total quantity occurs in 'external' production. This is because it is cheaper to take the quantity distortion in the 'external' market, as $P(q^{f})$ rises above p^{e} as soon as 'internal' quantity is reduced. Therefore, the 'external' market absorbs the full burden of the total quantity distortion necessary to provide incentives. In general, the more elastic 'external' demand is relative to 'internal' demand, the better the 'external' market insulates the 'internal' market.

5. Second-best regulatory policy: Adverse selection plus moral hazard

5.1. Framework and feasibility

Now consider the case where the decision to sell in the 'external' market is controlled by the firm. It is clear that the firm will always set q^e to maximize 'external' profits subject to the policy choice of the regulator. Thus, for any given level q of 'internal' production, a firm of type θ sets 'external' production according to $r(q,\theta)$. From the perspective of the regulator, this introduces a moral hazard since 'external' production can only be influenced indirectly through the reaction function.

The instruments available to the regulator now consist of 'internal' production and the tax, so that a policy is defined by $\Omega(\theta)=\{q(\theta),T(\theta)\}$. The solution procedure follows the same logic as before with the important exception that $r(q(\hat{\theta}),\theta)$ everywhere replaces $q^e(\hat{\theta})$, in order to account for the firm's profit-maximizing choice of 'external' production.

As described in the appendix, a policy is feasible if and only if

$$q(\theta) \text{ is a non-increasing function in } \theta, \tag{21}$$

$$\Pi(\theta)=\Pi(\bar{\theta})+\int_{\theta}^{\bar{\theta}} C(q(t)+r(q(t),t))\,\mathrm{d}t, \tag{22}$$

$$\Pi(\bar{\theta})\geqq 0. \tag{23}$$

Note that (21) requires 'internal' production to be non-increasing; when only adverse selection is present, total production is non-increasing.

It is then straightforward to substitute for $T-P(q)q$ in the objective of the government and obtain:

$$W\langle\Omega\rangle=\int_{\theta}^{\bar{\theta}}[V(q(\theta))+R^e(r(q(\theta),\theta))$$

$$-\rho(\theta)C(q(\theta)+r(q(\theta),\theta))]f(\theta)\,\mathrm{d}\theta-\Pi(\bar{\theta}). \tag{24}$$

Consider how moral hazard has altered the objective of the government. Like the first-best, 'external' production enters into 'external' revenue and total cost as $r(q,\theta)$, not at $r(q,\rho(\theta))$ as in the pure adverse-selection case. Since $r(q,\theta)$ declines in θ total costs are higher for any q than in the pure adverse-selection case. This reflects profit-maximization by a firm of type θ as opposed to an uninformed government. In contrast to the first-best and like the pure adverse-selection case, however, asymmetric information forces the government to weight total cost with $\rho(\theta)$ in order to account for the incentive costs of inducing revelation.

The regulator chooses $q(\cdot)$ to maximize (24) subject to (21). Characterizing the solution is more difficult in this case because the reaction of the firm appears in the objective function. In particular, the objective may not be concave since concavity is determined by second-order properties of $r(q,\theta)$ which, in turn, are determined by third-order derivatives of $C(\cdot)$ and $R^e(\cdot)$. This is a familiar problem in models with moral hazard and a formal treatment is presented in the appendix. In the discussion below we focus on the pointwise maximum of the integrand in (24); as demonstrated in the appendix, the optimal $q(\cdot)$ schedule either coincides with the pointwise maximum or can be solved for with the pointwise maximum. For several important special cases, including constant MR^e, the pointwise maximum is the optimal schedule.

5.2. Solution and discussion

The first-order condition for the pointwise solution of 'internal' quantity q is given by:

$$P(q)=\rho^*(\theta,q)C'(q+r(q,\theta)), \tag{25}$$

where

$$\rho^*(\theta,q)=\begin{cases}\rho(\theta), & \text{for } MR^e(0)<\theta C'(q),\\ \theta+\dfrac{F(\theta)}{f(\theta)}\left[1+\dfrac{\partial r(q,\theta)}{\partial q}\right], & \text{for } MR^e(0)>\theta C'(q).\end{cases} \tag{26}$$

Thus, when moral hazard is present, $\rho^*(\theta,q)$ replaces $\rho(\theta)$ as the weight in the marginal cost of information. When $r(q,\theta)=0$, $\rho^*(\theta,q)$ collapses to $\rho(\theta)$. When $r(q,\theta)>0$, the firm's marginal reaction to the regulator's choice of q distorts $\rho(\theta)$.

It is useful to consider how $\rho^*(\theta,q)$ relates to θ and $\rho(\theta)$. Since $\partial r/\partial q \geqq -1$, we always have $\rho^*(\theta,q)\geqq\theta$. Furthermore, it is easy to verify that the relative ordering of $\rho^*(\theta,q)$ and $\rho(\theta)$ is determined by returns to scale (i.e. the sign of C''). Specifically:

$$\rho^*(\theta,q) \lesseqqgtr \rho(\theta), \quad \text{as } C'' \gtreqqless 0.$$

This is because C'' determines the sign of $\partial r/\partial q$. For instance, with increasing returns marginal cost falls with output and it is profitable for the firm to increase 'external' production as 'internal' production rises. Thus, the effect of moral hazard is to make $\rho^*(\theta,q)$ larger than $\rho(\theta)$.

Consider the critical value, θ^*, that differentiates 'high cost' from 'low cost' types. The following results are established in the appendix. Let $q^0(\theta)$ be defined as the solution to $MR^e(0)=\theta C'(q)$. Note that the reaction of a type θ firm ceases to be positive when 'internal' production is $q^0(\theta)$. Then θ^* is the solution to $P(q^0(\theta))=\rho(\theta)C'(q^0(\theta))$. For a 'high cost' type, $\theta \geqq \theta^*$, the optimal 'internal' quantity is $q^*(\theta)=q^s(\theta)$. When $C''>0$, we have $\theta^* \geqq \theta^f$; when $C''=0$, we have $\theta^*=\theta^f$; and when $C''<0$, we have $\theta^* \in (\theta^s,\theta^f)$. The net effect of moral hazard relative to pure adverse selection is that fewer cost types are denied access to the 'external' market, and in some cases no types are denied access.

The optimal 'internal' quantity, $q^*(\theta)$, is determined by (25). The only difference between (25) and the analogous first-best condition (6) is that θ is replaced with $\rho^*(\theta,q)$. Since $\rho^*(\theta,q) \geqq \theta$, we must have $q^*(\theta) \leqq q^f(\theta)$, so that 'internal' quantity is distorted.

To investigate the insulation properties, it is useful to consider the case of constant 'external' marginal revenue. As long as there is some linkage between the two markets (i.e. $C'' \neq 0$), we have $\partial r/\partial q=-1$ and therefore $\rho^*(\theta,q)=\theta$. Again, $q^*(\theta)=q^f(\theta)$, which is a constant. As in the pure adverse-selection case, the 'internal' market is fully insulated. Moreover, we have $\theta^*=\theta^f$, since 'external' production is controlled by the firm instead of the uninformed government, and 'external' production is set at $q^{ef}(\theta)$, the first-best 'external' quantity. Thus, the effects of moral hazard and adverse selection exactly offset one another so that we are left with the first-best quantities over the first-best range of cost types.[6] As in the pure adverse-selection case, the 'internal' market is better insulated the more elastic 'external' demand is relative to 'internal' demand.[7]

6. Conclusion

We have demonstrated that regulators can use the profit opportunities created by an 'external' market to absorb incentive costs and thus insulate the 'internal' market from the quantity distortions that arise in environments with asymmetric information. Insulation is accomplished by allowing 'low cost' types to operate in the 'external' market, thus placing some of the quantity distortion on 'external' production. The degree of insulation is determined by the structure of production costs and 'external' demand. A necessary condition for insulation to occur is that marginal costs vary with

[6]As noted in the appendix, there is a kink in the objective function at the 'internal' quantity $q^0(\theta)$ where the reaction becomes positive. This means that $q^*(\theta)$ may equal $q^0(\theta)$ near θ^f before it equals the first-best quantity.

[7]In general, moral hazard introduces two offsetting effects relative to the case of pure adverse selection so that $q^*(\theta)$ may be larger or smaller than $q^s(\theta)$: in (25) the ρ^* and C' terms always move in opposite directions.

output levels, thus providing a cost link between the two markets. Given this linkage, the degree of insulation rises when 'external' demand is more elastic. If the firm is an 'external' price-taker, complete insulation occurs as 'internal' quantity can be set at the first-best level for 'low cost' types.

Appendix

Several claims in the text follow from standard arguments. We sketch the proof in these cases.

A. $\Omega(\cdot)$ is feasible iff (12)–(14) hold.

This is standard. Consider necessity first. Clearly, (IR) implies (14). To see (12), apply (IC) at $(\hat{\theta},\theta)$ and $(\theta,\hat{\theta})$. Combining and simplifying yield an upper and lower bound on $\Pi(\theta)-\Pi(\hat{\theta})$. Then (12) follows because $C(\cdot)$ is an increasing function. Dividing by $(\hat{\theta}-\theta)$ and taking limits as $\hat{\theta}\to\theta$ yield $\Pi'(\theta)=-C(Q(\theta))$ a.e., and (13) follows by integration.

For sufficiency, note that (IR) is a trivial consequence of (13) and (14) because $C(\cdot)$ is non-negative. To verify (IC), begin with the identity $\pi(\hat{\theta}|\theta)=(\hat{\theta}-\theta)C(Q(\hat{\theta}))+\Pi(\hat{\theta})$ and the substitute for $\Pi(\hat{\theta})$ with (13). (IC) then follows by (12).

B. If $\Omega(\cdot)$ is feasible then $W\langle\Omega\rangle$ is given by (15).

Solve for $T(\theta)-P(q(\theta))q(\theta)$ with the definition of profits (1) and then use (13) to substitute for $\Pi(\theta)$. The resulting expression contains a double integral that can be simplified by using integration by parts. (15) then follows from the definition of $\rho(\theta)$. At an optimum we have $\Pi(\bar{\theta})=0$, and $T(\theta)$ can be found with (13) by using the quantity schedules.

C. $\Omega(\cdot)$ is feasible iff (21)–(23) hold.

Consider necessity. Apply (IC) at $(\hat{\theta},\theta)$ and $(\theta,\hat{\theta})$ and use (10) and (11) with $\Omega(\theta)$ to combine and simplify. This yields:

$$\Pi^{e}(q,\theta)-\Pi^{e}(q,\hat{\theta})\geqq\Pi(\theta)-\Pi(\hat{\theta})\geqq\Pi^{e}(\hat{q},\theta)-\Pi^{e}(\hat{q},\hat{\theta}), \tag{A.1}$$

where $q=q(\theta)$ and $\hat{q}=q(\hat{\theta})$. From the definition of Π^{e} in (11), we find that the cross-partial derivative $\Pi^{e}_{q\theta}\leqq 0$ [the right- and left-hand derivatives are both non-positive along $MR^{e}(0)=\theta C'(q)$]. Thus, $q(\cdot)$ is non-increasing.

To verify (22), let θ be a continuity point of $q(\cdot)$. Since Π^{e} is convex in θ, we have $\Pi^{e}(\hat{q},\theta)-\Pi^{e}(\hat{q},\hat{\theta})\geqq(\hat{\theta}-\theta)\Pi^{e}_{\theta}(\hat{q},\hat{\theta})$. As Π^{e}_{θ} is continuous in θ, dividing (A.1) by $(\hat{\theta}-\theta)$ and letting $\hat{\theta}\to\theta$ yields:

$$\Pi'(\theta)=-\Pi^{e}_{\theta}(q(\theta),\theta)=-C(q(\theta)+r(q(\theta)\,\theta))\quad\text{a.e.}$$

(22) then follows by integration. (23) is obvious.

Consider sufficiency. (IR) is trivial. To verify (IC), use (22) and (10) with $\Omega(\theta)$ to obtain:

$$\Pi(\theta)-\pi(\hat{\theta}|\theta)=\int_{\theta}^{\hat{\theta}} C(q(t)+r(q(t),t))\,\mathrm{d}t+\Pi^{e}(q(\hat{\theta}),\hat{\theta})-\Pi^{e}(q(\hat{\theta}),\theta).$$

As $\Pi^{e}_{\theta}(x,t)=-C(x+r(x,t))$, we have $\Pi^{e}(x,\hat{\theta})-\Pi^{e}(x,\theta)=-\int_{\theta}^{\hat{\theta}} C(x+r(x,t))\,\mathrm{d}t$. Thus,

$$\Pi(\theta)-\pi(\hat{\theta}|\theta)=\int_{\theta}^{\hat{\theta}} [C(q(t)+r(q(t),t))-C(q(\hat{\theta})+r(q(\hat{\theta}),t))]\,\mathrm{d}t.$$

This is non-negative since $x+r(x,t)$ is increasing in x [as $1+r_q(x,t)\geqq 0$] and $q(t)\geqq q(\hat{\theta})$ for $t\leqq\hat{\theta}$. Thus (IC) holds.

D. The program from (25) and (21) is

$$\max_{q(\cdot)} \int_{\theta}^{\bar{\theta}} w(q(\theta),\theta)f(\theta)\,\mathrm{d}\theta$$

$$\text{s.t.}\quad q(\cdot)\ \text{non-increasing,} \tag{A.2}$$

where $w(q,\theta)\equiv V(q)+R^{e}(r(q,\theta))-\rho(\theta)C(q+r(q,\theta))$.
Before solving (A.2) we need to develop a few of the properties of $w(q,\theta)$.

From the definition in (11) we have:

$$w(q,\theta)\leqq V(q)+\Pi^{e}(q,\rho(\theta)), \tag{A.3}$$

so that $w(q,\theta)$ is bounded above by a concave function in q. Whenever $MR^{e}(0)<\rho(\theta)C'(q)$, so that $r(q,\theta)=0$, (A.3) becomes an equality. Now define $q^{0}(\theta)$ by $MR^{e}(0)=\theta C'(q^{0}(\theta))$. For $\theta\geqq\theta^{s}$ we know that the RHS in (A.3) has a global maximum at $q^{s}(\theta)$. Defining θ^{*} by $q^{0}(\theta^{*})=q^{s}(\theta^{*})$, we have $\theta^{*}>\theta^{s}$ (see E. below). Thus, $q^{s}(\theta)$ is the global maximum of $w(q,\theta)$ when $\theta\geqq\theta^{*}$.

At $q=q^{0}(\theta)$ the function $w(q,\theta)$ has a kink, as $r_q(q,\theta)$ jumps from 0 to $\theta C''(q)/[MR^{e'}(0)-\theta C''(q)]$. The sign of C'' determines whether $r_q=0$ for q above or below $q^{0}(\theta)$. In each case, however, we see from (A.3) that $w_q(q,\theta)$ jumps downward at $q=q^{0}(\theta)$.

When $\theta<\theta^{*}$, the behavior of $w(q,\theta)$ depends on the reaction $r(q,\theta)$. Letting $m(\cdot)=MR^{e}(\cdot)$ and $\Delta=m'-\theta C''$, we have (suppressing arguments):

$$w_{qq}=\rho'-\left[\theta+\frac{F(\theta)}{f(\theta)}\frac{m'}{\Delta}\right]\frac{m'C''}{\Delta}-\frac{F(\theta)}{f(\theta)}C''r_{qq}.$$

It is tedious but straightforward to verify that the existence conditions for the first-best solution imply that the first two terms are negative. Also, if MR^e is concave and C' is convex, then $r_{qq}>0$ and, hence, $w_{qq}<0$ for $\theta<\theta^*$.

These sufficient conditions for $\theta<\theta^*$, the above discussion for $\theta>\theta^*$, and the kink property imply that $w(q,\theta)$ is strictly concave in q for all θ. Thus, the pointwise maximum of $w(q,\theta)$ in q is unique for each $\theta\in[\underline{\theta},\bar{\theta}]$ and, by the Maximum Theorem, the pointwise maximum is a continuous function of θ. Under the mild regularity condition that the pointwise maximum changes from increasing to decreasing a finite number of times, the existence and uniqueness of an optimal $q^*(\cdot)$ for (A.2) in the class of piecewise continuously differentiable functions follows from Theorem 3 in Guesnerie and Laffont (1984). Further, from Appendix C of Laffont and Tirole (1986), this $q^*(\cdot)$ is optimal in the set of non-decreasing functions as well.

The optimal $q^*(\theta)$ equals the pointwise maximum of $w(q,\theta)$ or it is constant. It is constant on a finite number of subintervals, say I_k, where $\int_{I_k} w_q(q_k^*,\theta)f(\theta)\,\mathrm{d}\theta=0$ and q_k^* is the pointwise maximum at one of the endpoints of I_k.

E. There is a unique θ^* such that $q^0(\theta^*)=q^s(\theta^*)$, and

$$\text{as } C''\begin{cases} > \\ =0, \\ < \end{cases} \text{we have } \theta^* \begin{cases} \geqq\theta^f, \\ =\theta^f, \\ \in(\theta^s,\theta^f). \end{cases}$$

When $C''=0$ we have C' constant. Then $q^0(\theta)$ reduces to a vertical line at the value of θ such that $MR^e(0)=\theta C'$. This is the definition of θ^f.

When $C''<0$, $q^0(\theta)$ is an increasing function. From the definition of θ^s in (19) and the definition of $q^0(\theta)$, we have:

$$\theta^s C'(q^0(\theta^s))=MR^e(0)>\theta^s C'(q^s(\theta^s)),$$

since $\rho(\theta^s)>\theta^s$. This implies $q^0(\theta^s)<q^s(\theta^s)$ when $C''<0$. Similarly, using (7), we find $q^0(\theta^f)=q^f(\theta^f)>q^s(\theta^f)$. Since $q^0(\theta)$ is increasing and $q^s(\theta)$ is decreasing, continuity implies a unique $\theta^*\in(\theta^s,\theta^f)$ where $q^0(\theta^*)=q^s(\theta^*)$.

When $C''>0$, $q^0(\theta)$ is a decreasing function. If $MR^e(\cdot)$ is constant, then $\theta^*=\theta^f$. Otherwise, (6) implies that $q^f(\theta)<q^0(\theta)$ when $\theta<\theta^f$. Since $q^s(\theta)<q^f(\theta)$, if $q^0(\theta)$ and $q^s(\theta)$ cross before $\bar{\theta}$, they do so at $\theta^*\geqq\theta^f$.

References

Baron, David P. and David Besanko, 1984, Regulation, asymmetric information, and auditing, Rand Journal of Economics 15, 447–470.

Baron, David P. and Roger B. Myerson, 1982, Regulating a monopolist with unknown costs, Econometrica 50, 911–930.

Dasgupta, Partha S., Peter J. Hammond and Eric S. Maskin, 1979, The implementation of social choice rules; Some results on incentive compatibility, Review of Economic Studies 46, 185–216.

Guesnerie, Roger and Jean-Jacques Laffont, 1984, A complete solution to a class of principal-agent problems with an application to the control of a self-managed firm, Journal of Public Economics 25, 329–369.

Laffont, Jean-Jacques and Jean Tirole, 1986, Using cost observations to regulate firms, Journal of Political Economy 94, 614–641.

Myerson, Roger B., 1979, Incentive compatibility and the bargaining problem, Econometrica 47, 61–74.

Sappington, David E.M., 1983, Optimal regulation of a multiproduct monopoly with unknown technological capabilities, Bell Journal of Economics 14, 453–463.