On environmental policy and permitting

Philippe Bontems*  Jean-Marc Bourgeon†

April 2005

Abstract

The literature on environmental policy under adverse selection usually assumes that firms’ revenues vary monotonically with a private information parameter. However, it is easy to demonstrate using standard production setups that regularity is not the rule. We show that policy requirements are very sensitive to this assumption. In particular, the optimal instrument resembles more an “adaptable” pollution standard than the economic instrument of an environmental tax. We also show that permitting, which results in some firms over-investing in pollution-control equipment, does not serve the objective of improving the environment but rather allows the agency to increase the revenues of the policy and/or to diminish its enforcement cost.

Key-words: Environmental Taxation, Law Enforcement, Technological Standard, Adverse Selection.

JEL Classification: D62, D82, H21, H26, H32.
1 Introduction

Economists have long advocated the use of economic instruments such as taxation or marketable permits rather than the typical command-and-control policies to reduce pollution. However, the design of incentive schemes aimed at changing polluters’ behavior when the affected parties are privately informed about the costs and benefits they incur is often complex. Several studies have suggested regulatory strategies for dealing with pollution control when the necessary information is decentralized in the economy (see Lewis 1996 for a survey). A main finding is that a conventional emission fee is not sufficient for dealing with privately informed polluting firms. Firms are usually better informed than the regulator about their capabilities to reduce pollution. By overstating the burden associated with their abatement effort they are able to extract information rents. Such rents are generally costly for society, one reason being that raising public funds through taxation involves a deadweight efficiency loss. The regulator has thus to induce firms to reduce pollution at sufficiently high levels while limiting their information rents. This implies that the amount of pollution reduction may be distorted compared to its efficient level and that the taxation scheme is non-linear in general so polluters do not equate their marginal costs of abatement.

These features of the optimal pollution tax under private information are usually derived under the assumption that firms’ revenues vary monotonically with a unidimensional private information parameter (Baron 1985, Laffont 1994, Lewis 1996). However, we show using a standard production setup that in some realistic situations regularity of firms’ revenues is not the rule and that the stringency of the environmental regulation influences the extent of the asymmetric information problem. Indeed, consider a sector composed of firms with the same production technology involving a polluting input which substitutes for another, firm-specific one (like human capital, land fertility, or process knowledge). If firms in the sector are not constrained, they use the pollut-
ing input only and they all extract the same revenue from production. On the other extreme, if polluting emissions are completely forbidden, as firms are diversely affected depending on the cost of the firm-specific inputs they use to replace the polluting input, the sector presents the largest heterogeneity. Defining the firms’ private information parameter as the price differential between the polluting input and the firm-specific one, the firms’ profit decreases with the private information parameter, meaning that, when not allowed to pollute, firms with low (potential) marginal revenues from polluting generate more profits than firms with larger (potential) marginal revenues. An increase in the emission cap allows firms to substitute part of the costly input with the cheaper polluting one. The firms with the largest price differential are those that benefit the most from the relaxation of the policy. Hence, the firms’ marginal revenue from pollution increases with the private information parameter. As a result, the profit that a firm generates may increase or decrease with the private information parameter, depending on the regulation.

By reconsidering the optimal design of emissions regulation in such situations, we show that policy requirements are very sensitive to the assumption of regularity. Actually, compared to perfect information levels, firms extracting low pollution profits are tempted to increase their discharge (overstatement of their type), while firms extracting larger revenues may prefer to reduce their emissions (understatement). Hence, the non-regularity of their revenues introduces the possibility that a large subset of firms are pooled.

As a consequence, the optimal instrument resembles an “adaptable” pollution standard more than it does the economic instrument of an environmental tax. It is not quite the usual textbook command-and-control policy, but hopefully its real-life implementation. To implement the optimal allocation schedule, the policy must include exemption regimes which allow for some flexibility. For example, firms with private information parameter belonging to the lower tail of the distribution, which pollute
less than is allowed by the emission limit, should be exempted from part of the license fee. In contrast, firms in the upper tail of the distribution, which exceed the standard, should pay supplementary charges for emission level excess.

One way for the environmental agency to reduce the heterogeneity of a sector is to induce or require firms to adopt sounder production technologies, or at least more effective abatement technologies. Indeed, by relaxing the constraint on the polluting input, an abatement technology diminishes the need for firms to substitute for a firm-specific factor and thus ultimately reduces the heterogeneity among the firms in the sector. One method sometimes employed consists in offering subsidies dedicated to (partly) covering the private costs of improvement in abatement technologies. Alternatively, a more coercive and widespread approach consists in assigning firms a right or a “permit” to operate only if some minimum requirements with respect to pollution control technology are met. We examine how such a permitting policy interacts with the design of the emission regulation described just above, and in the extended setting where emissions are not readily observable by the regulator, who therefore has to audit firms in order to deter tax evasion. We show that permitting, which results in some firms over-investing in pollution-control equipment, does not serve the objective of improving the environment but rather allows the agency to increase the money raised by the policy and/or to diminish its monitoring cost.

A by-product of this analysis is that we show that costly monitoring of emissions generates an efficiency-rent extraction trade-off that is qualitatively different than the adverse selection one. Because auditing firms is costly, and because the probability of inspection is related to the firm’s revenues, rent extraction is a concern for the regulator even in the absence of tax-raising considerations. However, all firms enjoy strictly positive revenues in this latter case. Indeed, probability of inspection is determined by the sum of the firm’s revenues and its expected payments in case of evasion. As tax payments cancel out in the sum, they are defined only up to a constant. It is
thus always possible to specify an emission-payment schedule guaranteeing firms net positive revenues.

The mechanism design approach of environmental policy that we adopt in this paper is closely related to the analysis of countervailing incentives of incentive theory (Lewis and Sappington 1989a and 1989b, Maggi and Rodriguez-Clare 1995 and Jullien 2000). Some of these works, developed in the context of production regulation, suggest that countervailing incentives are favored by the principal to diminish information rents left to the agent. For instance, Lewis and Sappington (1989b) show that the principal benefits from a somewhat artificial introduction of countervailing incentives by giving the agent an input necessary for production. We show that countervailing incentives may emerge naturally in the environmental regulation context, with the similar consequence that optimal contracts consist of a standard (a rule), with allowances given to exceptional cases only.

The recent extension of the regulation framework to the realistic case of a costly monitoring of pollution is surveyed by Cohen (1999). In the asymmetric information context, Swierzbinski (1994) and Bontems and Bourgeon (2005) consider the optimal environmental tax when monitoring is costly and enforcement penalties are limited. We extend this framework by considering the permitting policy as part of the regulatory arsenal.

Finally, Stranlund (1997) also examines firms’ adoption of new abatement technologies assuming that the regulator enforces an abatement standard by auditing abatement efforts at random. He shows that public funding helps the regulator to save on monitoring costs. Indeed, by allowing firms to adopt more efficient abatement technologies, public subsidies lower the marginal cost of abatement and thus diminish the incentives to commit fraud. Our approach differs from Stranlund’s on several grounds. While he considers that the agency knows the profits that firms derive from polluting, the policy we analyze is constrained by asymmetric information. Moreover, we determine
the optimal emission payment scheme and audit mechanism simultaneously, while the audit policy he considers is exogenously given. Finally, we allow the regulator to intervene directly in the choice of technology through permitting, while Stranlund considers public subsidies.

The rest of the paper is organized as follows: Section 2 presents the model and discusses the effects of an environmental regulation on the heterogeneity of a sector. Section 3 analyzes the benchmark case of the optimal regulation under perfect information. Section 4 presents the regulatory mechanism when information rents are costly due to the social cost of public funds. We first consider the optimal design of emission regulations without permitting and then analyze the effects of permitting policies. Section 5 is devoted to the situation where monitoring emissions is costly. The last section concludes.

2 Heterogeneity induced by environmental regulations

It is usually assumed in the literature on environmental policy under asymmetric information that firms differ in a (one-dimensional) private information parameter and that their revenues vary monotonically with it (Lewis 1996). The optimal regulation allows the regulator to screen firms, and generally consists in a non-linear allocation-payment schedule, provided that the firm’s marginal revenue also varies monotonically with the private information parameter, the so-called Spence-Mirrlees “single-crossing” condition. Whereas this latter property is a plausible assumption which may be considered as the mere characteristic of the private information parameter, it does not imply, and thus hardly justifies, the other common assumption of monotonic revenues. In fact, it is easy to show using standard production setups that it is not generally the case, and particularly when we consider the pollution induced by inputs that may substitute for other components of production which are firm-specific, like investment in process
knowledge, human capital, or labor.

To illustrate, consider a sector composed of firms that may use 2 inputs in quantities $z_i$, $i = 0, 1$, where input 0 stands for a polluting input. To simplify the algebra, assume that one unit of the polluting input generates one unit of emission, that inputs 0 and 1 are perfectly substitutable, and that the firms’ production function is given by

$$Q = (z_0 + z_1)^a,$$

where $a < 1$. Denote by $w_i$ the price of input $i$ and assume that $w_0$ is a market price (known to the agency) while $w_1$ is the firm-specific cost of input 1, with $w_1 > w_0$. Private information parameter $\theta$ can be defined as the firm-specific price differential between inputs 0 and 1, i.e., $\theta = w_1 - w_0$.

Significant examples of comparable production settings can be found in the agricultural sector, where pesticides and fertilizers are substitutes for labor and land quality. Asymmetric information naturally appears with pesticides when farms are self-managed since the opportunity cost of labor is often specific to farmers. Use of pesticides allows farmers to spend less time on the farm, but even if time savings are approximately the same for all farmers, their earnings off the farm may widely differ. Similarly, heterogeneity of individual situations also appears with fertilizer. Fertilizer allows farmers to compensate for the productivity of their plots (the quality of the land), which is not observable by the agency. As the same amount of fertilizer spread on different plots has diverse impacts on production depending on the quality of the soil, the corresponding earnings may widely differ from one farm to another.

Absent any environmental regulation, firms would use input 0 only, extracting the same revenue from polluting. When firms are restricted to (gross) emission level $y$, their restricted cost function is given by

$$c(Q; y, \theta) = \min_{z_0, z_1} \{w_0 y + (\theta + w_0)z_1 : Q = (z_0 + z_1)^a; z_0 \leq y\}$$

1Generalizations to more than 2 inputs and other pollution functions are straightforward but cumbersome.
and straightforward computations give

\[
c(Q; y, \theta) = \begin{cases} 
w_0/Q^{1/a} & y \geq Q^{1/a}/a \\
(w_0 + \theta)Q^{1/a} - \theta y & y < Q^{1/a}/a 
\end{cases}
\]

which is not generally monotonic in \(\theta\). The same is true for the associated restricted revenue function, defined as

\[
B(\theta, y) = \max_Q pQ - c(Q; y, \theta)
\]

and given by

\[
B(\theta, y) = \begin{cases} 
\left(\frac{ap}{w_0}\right)^{(1-a)/(1-a)}(1 - a)/a & p \leq y^{1-a}w_0/a \\
\left[\frac{ap}{(\theta + w_0)\theta}\right]^{1/(1-a)}(1 - a)/a + \theta y & p > y^{1-a}w_0/a
\end{cases}
\]

Hence, without regulation, or a very lax one (where \(y\) is very large), firms generate the same profit whatever the price of the firm-specific input since they do not use it. Hence, with lax environmental regulations, the private information parameter can be considered merely as a “potential” marginal gain from polluting, in case environmental constraints become stringent. At the other extreme, when a regulation is in place and no allowances are given to firms \((y = 0)\), the firms’ revenues are decreasing with \(\theta\), meaning that the lower the firm-specific cost of input 1, the less the impact of the prohibition. An increase in the emission quota allows firms to substitute part of the costly input 1 with the same amount of the cheaper input 0. Firms with the larger cost of the input 1 are those that benefit the most from the relaxation of the policy.

We can thus draw the following consequences from revenue function (1):

**Claim 1** Absent an environmental regulation, the sector appears homogeneous. The private information parameter is not relevant since firms can pollute as much as they want.

**Claim 2** Under a stringent emission cap, firms are heterogeneous both in their total and marginal revenues from polluting. The firm’s marginal revenue from polluting
increases with the private information parameter (we have $B_y(\theta, y) = \theta$),\(^2\) while the firms’ total revenue may be decreasing or increasing with it, depending on the stringency of the environmental regulation.

It is easy to find other examples that exhibit a revenue function which is not monotonically with respect to the private information parameter while the marginal revenue is. As we show in the following section, the regulator may use this feature to diminish the social costs of information rents.\(^3\) Observe that Claim 2 needs a caveat. Indeed, to attenuate the stringency of the environmental regulation, firms may choose to reduce their net emissions into the environment, denoted by $q$, by abating their gross pollution level according to $q = y - a$, where $a$ is the abatement effort. By relaxing the constraint on the polluting input, this would ultimately reduce the heterogeneity among the firms in the sector. However, abating pollution is costly, with an abatement cost that depends on the firms’ investment in pollution-control technologies. To reduce the heterogeneity of the firms in the sector, the agency may thus decide to impose a technological standard corresponding to required pollution-control equipment $I_0$. This is the case when firms have to receive allowances from the agency to operate before starting production in a specific area. This permitting policy consists of requiring firms to adopt at least several minimum pollution-control apparatus (e.g., a minimum size of a water purification facility or of waste containment equipment, a minimum scrubber’s capacity, etc.). We shall take into account this possibility in the following, and to strengthen our results we assume that only the revenue from polluting is firm-specific while the abatement technology is not (e.g., end-of-pipe abatement technology).\(^4\) For a given abatement effort and investment level, the firm’s abatement cost is thus per-

\(^2\)We denote by $f_x$ the partial derivative of function $f(\cdot)$ with respect to variable $x$.

\(^3\)Countervailing incentives were first analyzed by Lewis and Sappington (1989a) in the context of regulation of natural monopolies. Also see Lewis and Sappington (1989b), Maggi and Rodriguez-Clare (1995), and Jullien (2000).

\(^4\)Were the abatement cost firm-specific, it would be obvious that regulating the firms’ technology choices would allow the agency to diminish the detrimental consequences of the firms’ informational advantage.
fectly known to the agency. More precisely, for an investment \( I \) in pollution-control technologies and an abatement effort \( a \), the profit of a \( \theta \)-type firm is given by

\[
\tilde{\pi}(y, a, I, \theta) = B(\theta, y) - A(a, I) - I, \tag{2}
\]

where \( A(a, I) \) is convex, satisfying \( A(0, I) = 0 \) and \( A_{Ia} < 0 \) (investment in pollution-control technology allows firms to reduce the marginal cost of abatement).\(^5\) To pursue the analysis, we shall consider a simple form of (1) given by

\[
B(\theta, y) = K/\theta + \theta y, \tag{3}
\]

where it is assumed that \( y \leq \overline{y} \), i.e., that there is a maximum gross pollution level, which, for simplicity purposes, is assumed identical for all firms and corresponds to their selfish pollution level.\(^6\) Firms’ revenues are thus composed of a fixed \((K/\theta)\) and a variable \((\theta y)\) component. \(K/\theta\) corresponds to the minimal revenue that firms can obtain without polluting, and \(\theta y\) the additional amount procured by emitting gross pollution level \(y\).

### 3 The Optimal Environmental Regulation

#### 3.1 The agency’s task

Since pollution abatement is costly, without environmental regulation firms neither invest in a pollution-control technology nor abate emissions. They extract as much revenue as possible by polluting. To induce producers to internalize environmental damage, the regulator’s task is to design an emission-payment schedule, and eventually to complement it with permitting. An emission-payment schedule consists of individualized pairs of a monetary transfer and an emission level, one for each firm’s possible value of the private information parameter \( \theta \). Emission-payment schedules can

---

\(^5\)Convexity implies \( A_{aa} > 0, A_{II} > 0 \) and \( A_{aa}A_{II} - A_{Ia}^2 \geq 0 \). We also assume that \( A_I(a^*, 0) < 1 \) and \( A_a(0, I) = 0 \) to allow for interior solutions.

\(^6\)Revenue function (3) corresponds to the extreme case of a free polluting input assuming \( a = 1/2 \).
be straightforwardly interpreted as an environmental tax scheme, but the mechanism design approach we follow allows us to encompass any policy resulting in an emission-payment schedule. Indeed, we shall see that the real-life implementation of the optimal schedule resembles an environmental emission standard policy. The difficulty of the agency’s task comes from the fact that it only knows that \( \theta \) is distributed among firms over a non-negative interval \( \Theta = [\theta, \bar{\theta}] \) according to a cumulative distribution function \( G(\theta) \) and continuous density \( g(\theta) \).\(^7\)

While it is possible that the environmental regulation also individualizes the level of investment in an abatement technology, we consider in the following only “permitting” regulations, which are commonly observed in large sectors. Under such regulations, accreditations are given by specialized inspectors to firms complying with abatement equipment rules before production begins. Once they are allowed to produce, firms periodically (annually) declare their emission level to the administration and are charged (or given) the corresponding payment, whatever the size of their abatement technology. The administration may decide to verify the firm’s declaration, but these inspections are not systematic and do not usually require highly specialized inspectors.

We assume that, because of the so-called “second dividend,” the agency is not indifferent to the funds raised using an environmental policy. Indeed, using taxation to correct externalities also allows the government to diminish the tax burden that weighs on the rest of society. This reduction decreases the deadweight losses associated with other existing tax systems, such as income taxation, that cause distortions in the economy. Levying \( t \) on a polluting firm with an environmental tax allows the government to diminish other taxes by the same amount. This induces an indirect social gain of \( \lambda t \), where \( \lambda \) is a per monetary unit measure of these deadweight losses \((1 + \lambda \) is commonly called the shadow cost of public funds).\(^8\)

\(^7\)We also make the usual monotonicity assumptions \( d((1 - G)/g)/d\theta < 0 \) and \( d(G/g)/d\theta > 0 \).

\(^8\)For a discussion of the second dividend effect, see Bovenberg and Goulder (1996).
Before analyzing the optimal regulation under asymmetric information, it is useful to characterize the benchmark case of a perfect information policy designed by a well-informed and plenipotent regulator who may require of firms any investment in pollution control equipment and any abatement levels, as well as payment and emission levels.

3.2 The perfect information benchmark

The regulator’s task is thus to design an emission-payment schedule \( \{q(\cdot), t(\cdot)\} \), where \( t(\theta) \) is the amount paid by \( \theta \)-type firms for \( q(\theta) \) emission allowances, and to impose an investment-abatement schedule \( I(\cdot), a(\cdot) \) to maximize the expected social welfare given by

\[
W = \int_\theta^\theta [\pi(y(\theta), a(\theta), I(\theta), \theta) - D(q(\theta)) + \lambda t(\theta)]g(\theta)d\theta, \tag{4}
\]

where \( D(q) \) is the environmental damage, an increasing and convex function of individual emissions. To illustrate, we shall assume that this damage is given by \( D(q) = q^2/2 \). The optimal policy is characterized in the following proposition whose proof is left to the reader.

**Proposition 1** The perfect information policy \( \{t^{PI}(\theta), q^{PI}(\theta), I^{PI}(\theta), a^{PI}(\theta) : \theta \in \Theta\} \) corresponds to payment and emission pattern \( t^{PI}(\cdot) \) and \( q^{PI}(\cdot) \) satisfying

\[
t^{PI}(\theta) = \pi(q^{PI}(\theta), \theta) \tag{5}
\]

\[
q^{PI}(\theta) = (1 + \lambda)\theta \tag{6}
\]

while investment and abatement schemes \( I^{PI}(\cdot) \) and \( a^{PI}(\cdot) \) correspond to first-best levels \( I^*(\cdot) \) and \( a^*(\cdot) \) implicitly defined by

\[
-A_I(a^*(\theta), I^*(\theta)) = 1, \tag{7}
\]

and

\[
\theta - A_a(a^*(\theta), I^*(\theta)) = 0. \tag{8}
\]

12
Since the public objective increases with firms’ payments, the regulator captures their entire profits as stated in (5). This is the obvious consequence of the second-dividend effect. Indeed, when $\lambda = 0$, there is no reason to charge firms per se since revenues left to firms are not socially costly. Another consequence of the second-dividend effect is an emission schedule (6) larger than the first-best levels. Indeed, first-best levels are defined by the Pigovian rule of equating marginal private profits to marginal social damage, which would be the case with (6) were $\lambda = 0$. On the contrary, with distortions in the rest of the economy ($\lambda > 0$), the regulator allows firms to over-pollute when compared to first-best levels. Indeed, since the agency captures all pollution revenues, increased emission levels means that more money is collected. This tax-revenue objective does not affect the optimal investment and abatement levels which correspond to first-best. This comes from the fact that the marginal benefit from polluting of the $\theta$-type firm is the same whatever $y$ chosen by the firm. First-best levels satisfy (8) and (7), which state that the abatement level must allow the firm to equate the marginal abatement cost with the marginal benefit of polluting, and that the optimal investment level equalizes the marginal benefit of an increased abatement capacity to the marginal financial cost of the investment. This is exactly the private trade-off faced by a firm. Indeed, without permitting and for a given level of net emission $q < \bar{y}$, a $\theta$-type firm freely chooses its optimal abatement effort $a^*(\theta)$ and technology investment $I^*(\theta)$ to reach

$$\pi(q, \theta) \equiv \max_{a, I} \tilde{\pi}(q + a, a, I, \theta),$$

which is the revenue level, gross of emission payments, that it reaches under the public policy for given $q$ and $t$ whatever $\lambda$. Hence, it is important to stress that

**Corollary 1** It is not necessary to intervene in the firms’ investment and abatement choices under perfect information.

---

9Hence, when $\lambda = 0$, the omnipotent regulator may decide to subsidize firms.
Given convexity of $A(\cdot)$, optimal investment level and abatement effort are increasing functions of the firm’s type. Indeed, by differentiating (8) and (7), we come to

$$I^*(\theta) = \frac{-A_{Ia}}{A_{aa}A_{II} - A_{II}^2} > 0$$

(10)

and

$$a^*(\theta) = \frac{A_{II}}{A_{aa}A_{II} - A_{II}^2} > 0$$

(11)

which mean that the higher the firm’s profitability of polluting emission, the larger its pollution-control equipment and cleaning effort. Observe also that we have

$$\frac{dt^{Pi}}{dq^{Pi}} = 2\theta - [K/\theta^2 - a^*(\theta)]/(1 + \lambda).$$

Hence, for $K$ sufficiently high, the perfect information payment-emission schedule is decreasing then increasing, which means that compared to other firms, firms in the lower tail of the distribution are submitted to larger payments for fewer allowances. This comes from the fact that the regulator wants to capture the firms’ profits which are decreasing with $\theta$ in the lower tail of the distribution. Such a tax-scheme is obviously not implementable under asymmetric information, since these firms would choose to pollute more in order to decrease their tax-payment.

4 Adverse selection and environmental policy

Under asymmetric information, the agency encounters two types of adverse selection problems when designing an environmental policy for firms with non-monotonic revenues. As firms’ revenues are declining in the lower tail of the distribution, it is impossible for the agency to capture the entire profits of firms with low marginal pollution benefits because they may diminish their payment by pretending to have higher marginal pollution benefits. On the upper tail of the firms’ distribution, revenues are increasing but the difficulty is the same: it is impossible to capture the entire profit of firms with high marginal pollution benefits because they may pretend to have lower
marginal pollution benefits. By mimicking a lower-type firm, they would have fewer emission allowances, but since the revenue they extract by polluting is higher than for these lower-type firms, they end up having positive revenues.

We analyze how the agency deals with these adverse selection problems by first investigating the situation where firms are not required to get an operating permit. The agency regulates the emission level using only an emission-payment scheme. We then introduce the permitting policy as part of the regulatory arsenal and analyze the interaction with the emission control policy. This two-step procedure will allow us to separate the effects of the permitting policy from the rest of the environmental regulation.

4.1 Environmental policy without permitting

The regulation process under asymmetric information is modeled as a two-stage game. First, the regulator chooses an emission-payment schedule. Second, firms choose their abatement technology and effort and report their type (i.e., they choose a net emission and payment pair). More formally, the environmental agency’s mechanism consists of \( q(\tilde{\theta}), t(\tilde{\theta}) \), where \( \tilde{\theta} \) is the reported type. From the revelation principle, we can restrict the search for the optimal mechanism to the set of direct and incentive-compatible mechanisms. The expected profit of a \( \theta \)-type firm that announces \( \tilde{\theta} \) is given by

\[
U(\theta, \tilde{\theta}) = \pi(q(\tilde{\theta}), \theta) - t(\tilde{\theta})
\]

for all \( \theta \) and \( \tilde{\theta} \), which states that firms are better off announcing their true parameter than lying about it. This prevents \( \theta \)-type firms from finding it worthwhile to mimic other types of firms by choosing an emission-payment pair which is not intended for
them. Let us denote by $R(\theta)$ the profit of a $\theta$-type firm that truthfully announces its type, i.e.,

$$R(\theta) = U(\theta, \theta).$$  \hspace{1cm} (13)

Using (12), the incentive-compatibility constraints can be equivalently written as

$$R(\theta) \geq R(\tilde{\theta}) + \pi(q(\tilde{\theta}), \theta) - \pi(q(\tilde{\theta}), \tilde{\theta})$$  \hspace{1cm} (IC)

for all $\theta$ and $\tilde{\theta}$, meaning that the difference of net revenues between the $\theta$ and $\tilde{\theta}$-type firms must be at least equal to the difference between their gross profits. The agency must also take into account the firm’s participation constraint, which is given by

$$R(\theta) \geq 0$$  \hspace{1cm} (IR)

for all $\theta \in \Theta$. The agency’s program may thus be written as

$$\max_{q,R} \left\{ \int_{\theta}^{\tilde{\theta}} [(1 + \lambda)\pi(q(\theta), \theta) - D(q(\theta)) - \lambda R(\theta)]g(\theta)d\theta : (IC), (IR) \right\}$$  \hspace{1cm} I

Compared to the perfect information policy where (IR) is binding for all firms, a strictly positive revenue corresponds to an “information rent” for the firms that enjoy it. The term $E[\lambda R]$ thus corresponds to the information cost of the policy, due to constraints (IC).

Problem I presents two sets of inequality constraints and cannot be solved directly. However, incentive constraints can be simplified thanks to the following lemma, whose proof is standard and thus omitted:

**Lemma 1** Incentive constraints (IC) reduce to

$$\dot{R}(\theta) = -K/\theta^2 + a^*(\theta) + q(\theta)$$  \hspace{1cm} (IC1)

and

$$q(\theta) \geq q(\tilde{\theta}) \text{ for all } (\theta, \tilde{\theta}) \in \Theta^2 : \theta > \tilde{\theta}$$  \hspace{1cm} (MN)
Proof: standard.

By (IC1), to deter imitation, information rents must evolve according to the marginal profitability advantage at the assigned emission level. As mentioned above, this marginal profitability advantage is composed of two opposite terms: As $\theta$ increases, the revenue without polluting (equal to $K/\theta$) decreases, while the revenue extracted from gross pollution level (given by $\theta[a^*(\theta) + q(\theta)]$) increases. Given that marginal pollution profits increase with the type of firm, the “single crossing” condition holds in this context. However, since $\dot{R}(\theta)$ given by (IC1) may be either positive or negative, marginal information rents are not always monotonic. In particular, since $a^*$ and $q$ are increasing and $K/\theta^2$ is decreasing, $R$ may be first decreasing and then increasing. Hence, we would have at least one firm’s type for which the (marginal) incentives to understate and to overstate their type exactly compensate. For firms with lower or larger types, information rents would be higher. This is particularly interesting for the agency since in order to increase the proceeds of the policy, the agency must reduce information rents as much as possible, i.e., $R(\theta) = 0$ for a large set of firms. The marginal increase of information rents is null for these firms, i.e., we have $\dot{R}(\theta) = 0$ for a subset of type, which implies gross pollution levels $y(\theta) = K/\theta^2$ for these firms. But the corresponding emission schedule $q(\theta) = K/\theta^2 - a^*(\theta)$ would be decreasing. Consequently, constraints (MN) are not satisfied, i.e., the policy is not incentive-compatible. Incentive-compatible policies thus entail binding emission constraints (MN), and the corresponding emission schedule presents a “flat” segment, i.e., $q(\theta) = \bar{q}$ whenever (MN) binds. The resulting schedule is presented in the following proposition.

Proposition 2 The optimal emission schedule solution of program I is given by

$$\hat{q}(\theta) = \begin{cases} 
q_1(\theta) & \text{for all } \theta < \theta_0 \\
\bar{q} & \text{for all } \theta_0 \leq \theta \leq \theta_1 \\
q_2(\theta) & \text{for all } \theta > \theta_1 
\end{cases}$$

(14)

\footnote{This problem has been investigated in detail by Guesnerie and Laffont (1984).}
where

\[ q_1(\theta) \equiv q^{PI}(\theta) + \lambda G(\theta)/g(\theta) \]

\[ q_2(\theta) \equiv q^{PI}(\theta) - \lambda[1 - G(\theta)]/g(\theta) \]

and \( \bar{q} \) satisfies

\[ \int_{\theta_0}^{\theta^*} [\bar{q} - q_1(\theta)]g(\theta)d\theta = \int_{\theta_1}^{\theta^*} [q_2(\theta) - \bar{q}]g(\theta)d\theta \quad (15) \]

with

\[ \bar{q} = q_1(\theta_0) = q_2(\theta_1) = K/\theta^* - a^*(\theta^*). \quad (16) \]

**Proof:** See appendix.

The emission schedule (14) entails three different regimes, as depicted in Fig. 1. For the first regime, \( \theta < \theta_0 \), the emission schedule is strictly higher than \( q^{PI}(\theta) \), except for \( \bar{\theta} \), the lowest type of firms. In this interval, information rents decrease. Low \( \theta \) firms may pretend to extract pollution revenues, thus to have a low revenue without polluting, so as to reduce their payment. However, the larger the allowance for higher-type firms, the less beneficial the mimicking. Indeed, by increasing the allowances for higher-type firms, the agency also increases the corresponding payments. With a \( dq \) increase in emissions over the perfect information level \( q^{PI}(\theta) \) for the \( \theta \)-type firms, the agency can increase their payment by \( \theta dq \). But this increase diminishes by \( dq \) the marginal benefits that lower-type firms extract from mimicking \( \theta \)-type firms. Hence, the agency can raise by \( dq \) the tax collected on the \( G(\theta) \) firms with type lower than \( \theta \), which yields a social benefit \( \lambda G(\theta)dq \) that the agency has to balance with the social welfare loss of the \( dq \) increase in emissions for the \( g(\theta) \) \( \theta \)-type firms. This loss is negligible since we have \( (1 + \lambda)\theta - D'(q^{PI}(\theta)) = 0 \). Trade-offs for larger emission increases result in optimal schedule \( q_1(\cdot) \).

For the third regime, \( \theta > \theta_1 \), we have the opposite result, as the optimal emission schedule is strictly lower than with perfect information, except for \( \bar{\theta} \), the highest type.
In this interval, information rents increase. As firms wish to reduce their payment, they are tempted to pretend that they extract fewer pollution revenues. However, the lower the allowances for low-type firms, the less beneficial the mimicking. Indeed, a marginal decrease $dq$ from $q^{PI}(\theta)$ has a negligible effect on the welfare generated by the $g(\theta)$ $\theta$-type firms, but allows for social gains $\lambda(1 - G(\theta))dq > 0$ on all firms of type higher than $\theta$ due to decreased information rents. It is thus optimal to reduce emission assignments on this interval, with the optimal schedule given by $q_2(\cdot)$.

Finally, the second regime of the emission schedule presents a “flat” (pooling) segment for intermediary types ($\theta_0 \leq \theta \leq \theta_1$). These firms are therefore induced to pollute according to constant level $\bar{q}$. To explain the pooling interval, recall that the agency would like to extract the profits from firms of intermediate types. But setting $q(\theta) = K/\theta^2 - a^*(\theta)$ over a subset results in a decreasing emission schedule, and leads to the violation of incentive constraints. The agency can do no better than to induce the same level of net pollution $\bar{q}$, in $[\theta_0, \theta_1]$. The extent of this pooling interval is given by the intersection between $\bar{q}$ and the first and third regime of scheme $\hat{q}$. It is optimally determined by (15), which states that in $[\theta_0, \theta_1]$, decreased marginal damage for low-types firms, due to a $\bar{q}$ emission level lower than $q_1(\theta)$ for $\theta < \theta^*$, must compensate for increased marginal damage for high-type ($\theta > \theta^*$), due to $\bar{q}$ being greater than $q_2(\theta)$. $\theta^*$ corresponds to firms for which the countervailing incentives to overstate and understate their type have the same magnitude and annihilate each other. These firms cannot extract information rent and thus their tax payment corresponds to their entire profit, as in the perfect information case. All other firms enjoy positive rents.

Any emission-payment schedule can be interpreted as a non-linear taxation scheme, and therefore environmental taxation generally proves to be the best policy instrument. However, when a large subset of firms are pooled, the optimal schedule resembles more an emission standard-cum-license fee.\footnote{Using a uniform distribution, it is straightforward that the proportion of firms that pollute at
More precisely, schedule (14) can be implemented through an “adaptable” emission standard policy whose interpretation is the following. To be allowed to pollute up to the $\bar{q}$ emission quota, firms have to pay a license fee equal to $\pi(\bar{q}, \theta^*)$. This is the choice for firms of intermediate-type, which belong to the set $[\theta_0, \theta_1]$. The policy also includes two sorts of exemption regimes. Firms in the lower tail of the distribution, i.e., of type lower than $\theta_0$, must be compensated for emission levels lower than the standard and thus are exempted from part of the license fee according to their emission deficits. Firms in the upper tail of the distribution, i.e., of type larger than $\theta_1$, must be allowed to exceed the standard by paying supplementary charges according to their excess in emission level.

4.2 Permitting and the environmental policy

We now derive the effect of a permitting policy on the optimal regulation. This policy consists in requiring firms to adopt a minimum size of pollution-control equipment $I_0$ in order to obtain their operating permit. Since private equipment choices $I^*(\theta)$ increase with the firm’s type, i.e., $dI^*/d\theta \geq 0$, only low-profitability firms are constrained by the technological standard $I_0$. Indeed, firms with a profitability parameter below the threshold type $\theta_s \equiv I^*-1(I_0)$ would have chosen a technology level lower than $I_0$ (i.e., $I^*(\theta) < I_0$). For all the firms that would have purchased less than the required level $I_0$, we have

$$-A_f(\hat{a}(\theta), I_0) - 1 < 0,$$

i.e., for all $\theta \leq \theta_s$, the marginal net private value of pollution-control technology is negative at the emission level $\hat{a}(\theta)$ chosen by the firm, which satisfies the following first-order condition:

$$A_a(\hat{a}(\theta), I_0) = \theta. \quad (17)$$

the quota is given by $\lambda/(1 + 2\lambda)$. This proportion increases with $\lambda$ but cannot exceed 50%, which is relatively low. However, we can expect a larger proportion of firms under the emission quota with distributions more concentrated around their mean values.
What can motivate the agency to impose excessive pollution-control equipment on firms that pollute the least? The answer resides in the impact of this overcapacity on the rents extraction potential of the environmental policy. With over-investment in abatement technology, firms are induced to increase their gross pollution level, which is less beneficial the lower the firm’s marginal benefit from polluting. As a result, firms’ revenues in the lower tail of the distribution (the types lower than \( \theta_s \)) are less heterogeneous. Indeed, the supplementary charge in investment is larger the lower the type, while the increase in pollution revenues is larger the larger the type. This “flattens” the revenue profile of firms in the lower tail of the distribution, which reduces the benefits that low-type firms extract from mimicking higher types.

This can be seen from (IC1) and (17), which indicate that information rents evolve positively with the firms’ abatement levels and that abatement levels increase with the capacity of the pollution-control equipment. Indeed, for low \( \theta \), standard \( I_0 \) modifies (IC1), which reads as

\[
\dot{R}(\theta) = -\frac{K}{\theta^2} + \hat{a}(\theta) + q(\theta). \tag{18}
\]

Differentiating (17), we obtain that for these firms, an increase in the technological standard affects marginal information rents by

\[
d\dot{R}/dI_0 = d\hat{a}/dI_0 = -A_{Ia}/A_{aa} > 0.
\]

Since information rents of low-type firms decrease (we have \( \dot{R} < 0 \) for \( \theta < \theta^* \)), a standard allows the agency to attenuate this decrease and thus to reduce the rents in the lower tail of the distribution of firms. Consequently, for given net emission levels the agency is able to increase the corresponding payments by imposing a technological standard.

We now investigate more closely the policy mix. The regulation process is modeled as a three-stage game. First, the regulator chooses an emission-payment schedule \( \{t(\cdot), q(\cdot)\} \) and a pollution-control technology standard \( I_0 \). Second, firms choose their abatement technology. Firms are inspected and those that do not comply with requirement \( I_0 \) are not allowed to produce (i.e., the plant is closed down). Third, remaining

21
firms report their type (i.e., they choose their net emission and payment pairs) and simultaneously decide on their abatement effort.

The objective of the regulator is to maximize the expected social welfare given by

\[ \int_{\theta}^{\bar{\theta}} [(1 + \lambda)\tilde{\pi}(q(\theta)) + \hat{a}(\theta), \hat{I}(\theta), \theta) - D(q(\theta)) - \lambda R(\theta)]g(\theta)d\theta, \]

where \( \hat{I}(\theta) \geq I_0 \) is the investment choice induced on the \( \theta \)-type firms, which verifies

\[ A_I(\hat{a}(\theta), \hat{I}(\theta)) + 1 \geq 0 \quad (19) \]

with strict inequality whenever \( \hat{I}(\theta) = I_0 \), and \( \hat{a}(\theta) \) is the corresponding abatement effort, given by

\[ A_a(\hat{a}(\theta), \hat{I}(\theta)) = \theta. \quad (20) \]

As in the previous section, \( R(\theta) \) and \( q(\theta) \) are subject to (MN) and (18) respectively. The optimal policy is characterized in the following proposition.

**Proposition 3** Under the optimal policy,

(i) firms with a profitability parameter below threshold type \( \theta_s \equiv I^{*-1}(I_0) \leq \theta^* \) are constrained by the optimal technological standard \( I_0 \), satisfying

\[ \int_{\theta_s}^{\theta} \left\{ A_I(\hat{a}(\theta), I_0) + 1 + \frac{\lambda A_Ia(\hat{a}(\theta), I_0) G(\theta)}{1 + \lambda A_a(\hat{a}(\theta), I_0) g(\theta)} \right\} g(\theta)d\theta = 0 \quad (21) \]

where \( \hat{a}(\cdot) \) is given by (17), and

(ii) the optimal emission schedule is still given by (14)-(16).

**Proof:** See appendix.

As indicated in point (i), \( \theta_s \leq \theta^* \), i.e., only firms in the lower tail of the distribution are constrained by the emission standard. The motive for public intervention clearly appears in (21) whose integrand is the opposite of the “virtual surplus” derived from a marginal increase in the technological standard for the \( \theta \)-type firm. This increase
diminishes $\theta$-type firms’ profit, which corresponds to a social cost equal to $(1+\lambda)(A_I+1)$ for these $g(\theta)$ firms, but allows the agency to reduce the information rents of the $G(\theta)$ less productive firms by a marginal amount equal to $A_{Ia}/A_{aa}$, and thus save $\lambda G(\theta)A_{Ia}/A_{aa}$ on rents. This virtual surplus is negative for low-profitability firms (marginal gains for low types are small) but positive for more profitable firms. The optimal standard balances these marginal gains and losses.

Point (ii) of proposition 3 states that the optimal emission schedule is the same as the one imposed without permitting. The requirement of oversize pollution-control equipment is thus for tax-raising purposes only, and not to diminish the emissions of these firms. In particular, as $\theta_s \leq \theta^*$, the “licence fee” $\pi(\bar{q}, \theta^*)$ discussed above is not affected by the technological standard since $\theta^*$-type firms are not forced to over-invest in abatement technology. Only the rebates for emission levels lower than the standard are changed. Firms with type lower than $\theta_s$ suffer for two reasons from the permitting policy: Rebates are decreased and the increases in gross pollution revenues do not reimburse the cost of the over-investments (the difference between $I^*(\theta)$ and $I_0$ for the $\theta$-type firm).

5 Evasion and permitting

Before concluding, it is interesting to discuss another motive to impose permitting, namely the regulation evasion problem. This is particularly relevant in the environmental regulation context since, in some instances, the regulation relies on firms’ declarations, and the agency has to perform costly inspections to verify the firms’ statements. Stranlund (1997) shows in a symmetric information context about firms’ type that permitting does indeed allow reduction in the cost of inspections. We show in this section that the asymmetric information problem is another motive for requiring firms to adopt minimum pollution-control equipment. To isolate this effect from the tax-raising consideration analyzed in the previous section, we conduct this analysis
assuming $\lambda = 0$.\footnote{See Bontems and Bourgeon (2005) for an analysis of the intertwined problems of increasing tax proceeds and avoiding tax evasion.}

The regulation and inspection process is modeled as a four-stage game. First, the regulator chooses an emission-payment monitoring schedule and a minimum abatement-technology standard. Second, firms choose their abatement technology. Firms that purchased less than $I_0$ are not allowed to produce (i.e., the plant is closed down). Third, remaining firms report their type (i.e., choose their net emission and payment pair) and simultaneously decide on their abatement levels. Fourth, firms are monitored depending on their stated types. If no audit occurs, the firm pays only the amount corresponding to its report. In the case of auditing, it pays a transfer corresponding to its announcement and the result of the audit.

More formally, the environmental agency’s mechanism consists of $I_0$, $q(\tilde{\theta})$, $t(\tilde{\theta})$, $\mu(\tilde{\theta})$, and $f(\tilde{\theta}, q)$, where $\tilde{\theta}$ is the reported type and $q$ stands for the emissions as revealed by the audit. Once the firm has invested at least $I_0$ in abatement equipment, it reports a $\tilde{\theta}$ profitability parameter to the agency. After this report, it is allowed to emit $q(\tilde{\theta})$. It must pay $t(\tilde{\theta})$ and is audited with probability $\mu(\tilde{\theta})$. In the case of auditing, it pays fine $f(\tilde{\theta}, q) \geq 0$ if $q \neq q(\tilde{\theta})$.\footnote{Since fines are constrained to be non-negative, we exclude tax rebates. Consequently, in accordance with the general practice, the agency cannot reward firms in case of compliance. See Świerzbinski (1994) for an analysis of the incentive properties of tax rebates.}

A firm may misbehave in two ways. First, it may misreport its type and pollute according to the allowance for the reported type. This is the typical adverse-selection problem we have just investigated where the firm perfectly mimics another type but does not evade taxation. Auditing the firm fails to disclose such mimicking because inspection reveals only pollution levels and neither the type nor the profit of the firm. Second, a firm may decide to evade taxation by choosing an emission level different from its allowance, i.e. $q \neq q(\tilde{\theta})$. Such shirking is detected when auditing the firm.

As the revelation principle applies in our context, we can restrict the search for the
optimal mechanism to the set of direct and incentive-compatible mechanisms without loss of generality. The expected profit of a $\theta$-type firm that announces $\tilde{\theta}$ and emits $q$ is given by

$$V(\theta, \tilde{\theta}, q) = \max_{a, I \geq I_0} \tilde{\pi}(a + q, a, I, \theta) - t(\tilde{\theta}) - \mu(\tilde{\theta})f(\tilde{\theta}, q).$$  \hfill (22)

Denoting by $R(\theta)$ the expected profit of a $\theta$-type firm that truthfully announces its type and pollutes according to the regulator’s allowance, the incentive constraints can be written as

$$R(\theta) \geq \max_{\tilde{\theta}, q} V(\theta, \tilde{\theta}, q) \quad \text{(IC')}$$

for all $\theta$. If the regulator’s mechanism satisfies (IC’), the firm is better off choosing the payment-pollution pair designed for its type than polluting at any other emission level and trying to pay any other amount. To satisfy (IC’), the regulator must be able to inflict severe punishments on the firm. However, the fines the regulator may inflict are usually bounded, i.e.

$$f(\tilde{\theta}, q) \leq \bar{F},$$

where $\bar{F}$ is the exogenous maximum fine, due to, for example, the firm’s limited liability.

Given (22), it becomes evident that firms caught evading taxation should be given the highest fine. A zero fine may also apply when the audit reveals that the firm has complied with the rule. Indeed, with risk-neutral firms, all incentives to truthfully report type can be embedded in payment $t(\theta)$.\footnote{Formally, any $q(\cdot), t(\cdot), \mu(\cdot), f(\cdot)$ scheme with $f(q(\theta), \theta) \geq 0$ may be replaced by a $q(\cdot), \hat{t}(\cdot), \hat{\mu}(\cdot), \hat{f}(\cdot)$ schedule with $\hat{t}(\theta) = t(\theta) + \mu(\theta)f(q(\theta), \theta)$ and $\hat{f}(q(\theta), \theta) = 0$.} Consequently, penalty schedule $f(\cdot)$ takes the form

$$f(\theta, q) = \begin{cases} 0 & \text{if } q = q(\theta) \\ \bar{F} & \text{otherwise} \end{cases}$$

Hence we have a simple penalty scheme: the payment is either 0 if the firm does not evade or the highest penalty otherwise. This simplicity allows us to separate the two aspects of the audit problem: the evasion problem is taken care of by the...
inspection-penalty schedule, whereas the inefficient mimicking problem is deterred by the emission-payment schedule. Therefore, since compliance is not rewarded, the $\theta$-type firm’s expected profit with an incentive-compatible mechanism becomes

$$R(\theta) = \max_{a, I \geq I_0} \tilde{\pi}(a + q(\theta), a, I, \theta) - t(\theta). \quad (25)$$

As mentioned above, auditing firms entails a cost that diminishes the expected social welfare. Assuming that auditing a $\theta$-type firm with $\mu(\theta)$ probability costs $c\mu(\theta)$, social welfare (4) becomes

$$W = \int_\theta^\bar{\theta} \left\{ \tilde{\pi}(\hat{a}(\theta) + q(\theta), \hat{a}(\theta), \hat{I}(\theta), \theta) - D(q(\theta)) - c\mu(\theta) \right\} g(\theta)d\theta \quad (26)$$

where $\hat{a}(\theta)$ and $\hat{I}(\theta)$ satisfy (19) and (20). This objective is constrained by the firm’s participation (IR), and the probability of being audited must satisfy

$$0 \leq \mu(\theta) \leq 1 \quad (27)$$

for all $\theta \in \Theta$. As in the preceding section, incentive constraints (IC’) can be simplified thanks to the following lemma:

**Lemma 2** Incentive constraints (IC’) reduce to (IC1), (MN), and

$$R(\bar{\theta}) \geq K/\bar{\theta} + \bar{\theta}y - I_0 - M \quad (IC2)$$

where $M$ is the expected total payment, which is constant whatever the type, i.e.

$$t(\theta) + \mu(\theta)\bar{F} = M \quad (28)$$

for all $\theta$. Inspection probabilities are given by

$$\mu(\theta) = [M - \tilde{\pi}(\hat{a}(\theta) + q(\theta), \hat{a}(\theta), \hat{I}(\theta), \theta) + R(\theta)]/\bar{F}, \quad (29)$$

and we have $\mu(\theta) > 0$.

---

15Assumption of a constant marginal monitoring cost is made for analytical convenience. Extensions to more general cost functions are straightforward.
**Proof:** See appendix.

In addition to the no-mimicking conditions (IC1) and (MN), incentive compatibility when firms may evade regulation brings up constraints (IC2), (28), and (29). When evading taxation, the optimal strategy for any firm is to emit maximum level $\bar{y}$ and invest at minimum level $I_0$. Condition (IC2) states that deterring the highest type of firm from evading regulation is sufficient to ensure that all other firms (with lower pollution revenues) will follow the policy requirements. Condition (29) states that minimal inspection efforts are necessary to maintain these incentives, i.e., the agency has to check compliance to induce lower emission levels rather than selfish ones. Observe that, at the optimum, we may have $\mu = 1$ for low types. Equation (28) defines the expected cost that shirking firms incur by evading regulation, given by $M = t(\theta) + \mu(\theta) \bar{F}$ for all $\theta$. The cost is the same whatever the firms’ stated type. Since the cost is constant, the optimal payment schedule and inspection rate are negatively related: The lower the amount paid by the firm, the higher the probability of inspection. Since $q(\cdot)$ and $t(\cdot)$ are non-decreasing, the inspection rate is thus a non-increasing function of the firm’s type. Consequently, the more openly a firm announces that it pollutes, the less likely it is to be inspected. This can be easily understood: One of the agency’s tasks is to deter firms from lying about their pollution levels. It is tempting for a firm to evade taxation by announcing a low emission level. In order to deter such shirking, the agency has thus to increase the probability of inspection for firms paying low fees.

Thanks to a simple fine schedule, once it is verified that firms are deterred from evading, (by designing an audit policy ensuring that the most profitable firms will not), the agency’s problem is reduced to the standard one of designing a contract in a pure adverse-selection setting: The mimicking incentive constraints are the same as those of a pure contract problem with perfect observability of emission levels. Results on second-order conditions of these models hold.
Using Lemma 2, the agency’s objective simplifies to
\[
W = \int_{\theta}^{\hat{\theta}} \left\{ \left( 1 + \frac{c}{F} \right) \tilde{\pi}(\hat{a}(\theta) + q(\theta), \hat{I}(\theta), \hat{I}(\theta)) - D(q(\theta)) - \frac{c}{F} [R(\theta) + M] \right\} g(\theta) d\theta
\]

Note that since \( R(\cdot) \) and \( M \) affect this objective negatively, (IC2) is binding at the optimum. Focusing on the case where \( \mu(\theta) < 1 \) for all \( \theta \in \Theta \), the optimal policy presents the following attributes:  

**Proposition 4** At the optimum of the agency’s program,

(i) The emission schedule satisfies
\[
\hat{q}(\theta) = (1 + c/F) \theta + \frac{c \cdot G(\theta)}{F \cdot g(\theta)},
\]

(ii) \( R(\theta) > 0 \) for all \( \theta \in \Theta \),

(iii) Firms with a profitability parameter below threshold \( \theta_s \equiv I^{*-1}(I_0) > \theta \) are constrained by technological standard \( I_0 \) satisfying
\[
\int_{\theta}^{\theta_s} \left\{ (1 + c/F) [A_I(\hat{a}(\theta), I_0) + 1] + \frac{c \cdot G(\theta)}{F \cdot g(\theta)} \frac{A_{Ia}(\hat{a}(\theta), I_0)}{A_{aa}(\hat{a}(\theta), I_0)} \right\} g(\theta) d\theta = \frac{c}{F}
\]

**Proof:** See appendix.

The optimal emission schedule (30) presents distortions from the first-best levels due to the monitoring cost and the adverse-selection problem. Let us first abstract from the adverse-selection problem by assuming that the regulator knows the firm’s type but still has to audit it to discover its emission level. Since it is not worth deterring imitation, (IC‘), constraints do not matter. However, the agency still has to deter evasion, i.e., to make sure that firms pay the amounts corresponding to their type and pollute accordingly. With (30), the optimal emission schedule simplifies to
\[
q^A(\theta) = (1 + c/F) \theta
\]

\[\text{Constraints } \mu(\theta) \leq 1 \text{ can be written as } R(\theta) - \tilde{\pi}(\hat{a}(\theta) + q(\theta), \hat{a}(\theta), \hat{I}(\theta), \theta) + M \leq \bar{F} \]

and may bind in the lower tail of the distribution of \( \theta \) when \( \bar{F} \) is small.
for all $\theta$, and thus $q^A(\cdot) > \theta$, i.e., we have under-deterrence compared to first-best levels $q^*(\theta) = \theta$. Indeed, to deter firms from deviating from their assigned emission levels, the agency has to inspect the $\theta$-type firm at minimal rate $\mu(\theta) = (M - \bar{\pi})/\bar{F}$. Compared to the perfect information level, a $dq$ increase in the $\theta$-type emission level induces a decrease in the inspection cost of $d/dq[c\mu] = -\bar{\pi}_q c/\bar{F} = c/\bar{F}$. The corresponding marginal social loss is negligible since $\bar{\pi}_q - D'(q^*(\theta)) = \theta - D'(q^*(\theta)) = 0$. The agency will therefore set the emission schedule above perfect information levels.

In an asymmetric information context, one finds an additional distortion affecting the emission schedule, given by $(c/\bar{F})[G(\theta)/g(\theta)]$. Since the agency is indifferent to the proceeds of the emission-payment schedule, adverse selection effects do not interfere directly with the agency’s objective. But firms’ revenues determine the inspection effort that the agency must produce since from (29) we have $d\mu/dR = 1/\bar{F}$ for all $\theta$. As auditing firms is costly, the agency finds it optimal to increase the emission schedule over $q^A(\cdot)$. This is easily understood when considering the incentives of the $\theta$-type firm. Firms are tempted to reduce their tax payment by understating their type, thus pretending to gain less pollution revenues. A $1 - G(\theta)$ proportion of firms more profitable than $\theta$-type ones extract a rent from this possibility of mimicking $\theta$-type firms. From (IC1) and (29), by increasing the $\theta$-type emission assignment, the agency increases its inspection efforts on these firms, which induces a marginal social cost equal to

$$[1 - G(\theta)]c/\bar{F}. \quad (32)$$

However, a $\tilde{\theta}$-type rent increase allows the agency to reduce $M$, the expected cost of evading, since $R(\tilde{\theta}) = K/\tilde{\theta} + \tilde{\theta}\tilde{y} - I_0 - M$, i.e. $\tilde{\theta}$-type firms are indifferent to whether they evade or pay taxes. In turn, the decrease induces a lower inspection rate on all firms (we have $d\mu/dM = 1/\bar{F}$), hence a marginal inspection cost savings of $c/\bar{F}$. The total effect of the marginal increase of the $\theta$-type emission is thus (32) minus $c/\bar{F}$, which yields marginal cost savings of $G(\theta)c/\bar{F}$. Indeed, inspection efforts are unaffected for
firms with types higher than $\theta$ (increased revenues $R(\cdot)$ and decreased $M$ leave $\mu(\cdot)$ unchanged) and are reduced for firms with types lower than $\theta$ (the decrease in $M$ is not compensated by increased rents).

Point (ii) states that the firms’ net earnings are positive. Although firms’ revenues affect the agency’s program through audit costs, raising public funds is not the objective that motivates the governmental intervention. From (29), we have $\mu(\theta) = [R(\theta) + M - \tilde{\pi}(\cdot)]/F$, and thus it is the sum of the $\theta$-type firm revenue $R(\theta)$ and the expected cost of evading $M$ that determines the cost of auditing. We can thus decrease $M$ and increase $R(\theta)$ from the same constant for all $\theta$ without changing the inspection probability. Consequently, the payment schedule solution of our agency’s problem is known up to a constant, and it is always possible to find a constant such that $R(\theta) > 0$ for all $\theta$. This is also the reason why the emission schedule (30) is simpler than (14)-(16), which was driven by the agency’s objective to raise as much revenue as possible. Observe that to deter evasion, we must have $\mu(\theta)\bar{F} \geq K/\theta + \theta\bar{y} - I_0$, which places a lower bound on $R(\theta) + M$, and is binding for $\bar{\theta}$. Hence, the only constraint on these sums is given by $R(\bar{\theta}) + M = K/\bar{\theta} + \bar{\theta}\bar{y} - I_0$, which depends on the technological standard only.

Determination of this standard is given by (31). To interpret this equation, it is also useful to disentangle costly auditing considerations from the adverse-selection problem. First consider the effect of a costly audit assuming symmetric information about firms’ types. In that case, the marginal cost of incentives is nil and the optimal standard $I_A$ is given by

$$\int_{\theta_A}^{\theta} \left(1 + c/F\right)[A_I(\hat{\theta}(\theta), I_A) + 1]g(\theta)d\theta = c/F$$

with $\theta_A \equiv I_A - 1 > \bar{\theta}$. The intuition is straightforward: Since firms have installed the required pollution-control apparatus, their gain from evading is reduced by the equipment cost. This allows a decrease in the expected payment in case of evasion, and therefore savings on audit cost. Indeed, from (IC2), which is binding at the optimum,
we have

\[ M = K/\bar{\theta} + \bar{\theta}\bar{y} - I_0 \]

An increase in \( I_0 \) allows \( M \), the expected evasion payment, to be diminished, and thus from (29), allows that the probability of inspection \( \mu(\theta) \) to be reduced for all \( \theta \). This induces a marginal social gain equal to \( c/\bar{F} \), the right-hand side of (33). However, this entails an over-investment for low profitability firms, i.e. for firms of types lower than \( \theta_A \), a proportion \( G(\theta_A) = \int_{\theta_A}^{\theta} g(\theta)d\theta \) of firms. The marginal increase in \( I_0 \) costs \( A_I + 1 \) for these firms. In addition to this direct effect on firms, it also requires the agency to monitor them more closely (at the associated marginal cost \( c \)) since their revenues are decreased. Indeed, from (29), which simplifies to

\[ \mu(\theta) = \{ M - [K/\theta + \theta(q(\theta) + \hat{a}(\theta)) - A(\hat{a}(\theta), I_A) - I_A]\}/\bar{F}, \]

definitions of the direct effect on low-type firms’ revenues and the indirect cost for the administration of this over-investment is given by (33)’s left-hand term.

In an asymmetric information context, over-investment induces an additional effect embedded in the term

\[ \frac{c}{\bar{F}}G(\theta)A_{ta}(\hat{a}(\theta), I_0)/A_{aa}(\hat{a}(\theta), I_0) \]

for each \( \theta < \theta_s \), which accounts for the decrease in the monitoring cost due to reduced information rents. Indeed, from (20), \( \theta \)-type firms increase abatement according to \( d\hat{a}(\theta)/dI_0 = -A_{ta}/A_{aa} \) as \( I_0 \) increases, which from (18) affects marginal revenues (we have \( d\hat{R}/dI_0 = d\hat{a}/dI_0 \)). As revenues are decreasing in the lower tail of the distribution of firms, profits of the \( G(\theta) \) firms with a type lower than \( \theta \) are reduced (we have \( dR/dI_0 = -d\hat{R}/dI_0 = A_{ta}/A_{aa} < 0 \)). Since their emission payments are increased, the agency can reduce its monitoring efforts on these firms: From (29), we have \( d\mu/dI_0 = (dR/dI_0)/\bar{F} \).
It is left to the reader to verify that the emission schedule (30) is unchanged if the agency does not impose permitting to firms. In this context again, requiring the less polluting firms to over-invest in pollution-control equipment is not intended to enhance the environment but only to diminish the operating cost of the agency.

6 Conclusion

We have studied the optimal regulation of pollution while reconsidering the regularity assumption that firms’ revenues vary monotonically with the private information parameter, an assumption usually found in the literature on environmental policy under adverse selection. By considering separately first that emissions are costlessly observable and second that monitoring pollution is possible only through costly audits, we have shown that policy requirements are very sensitive to this assumption. In particular, when monitoring is not costly, the optimal instrument resembles more an “adaptable” pollution standard than the economic instrument of an environmental tax. We have also shown that permitting, which actually requires firms to over-invest in pollution-control equipment, does not serve the objective of improving the environment but rather allows the agency to increase the proceeds of the policy or to diminish its monitoring cost.

Note that, under complete information, a technological standard would be useless because once net emissions are taxed at the Pigovian level, firms adopt the abatement technologies corresponding to socially efficient equipment. This is no longer true under incomplete information because the private and social values of a new abatement technology are no longer the same, due to informational rents and audit costs. When adding a minimum technological standard to the regulatory arsenal, both the informational rents left to firms and the incentives to evade environmental regulation are reduced.
References


Figure 1: Optimal emission schedule

\[
(1 + \lambda)\bar{\theta} + \lambda / g(\bar{\theta})
\]

\[
(1 + \lambda)\bar{\theta}
\]

\[
q^{PI}(\theta) + \lambda \frac{G(\theta)}{g(\theta)}
\]

\[
q^{PI}(\theta)
\]

\[
q^{PI}(\theta) - \lambda \left[1 - \frac{G(\theta)}{g(\theta)} \right]
\]

\[
K/\theta - a^*(\theta)
\]

\[
(1 + \lambda)\bar{\theta} - \lambda / g(\bar{\theta})
\]
Appendix

A Resolution of program I

Monotonicity constraint (MN) can be expressed as

\[ \dot{q}(\theta) = \delta(\theta) \]  \hspace{1cm} (34)

with \( \delta(\theta) \geq 0 \) for all \( \theta \in \Theta \). We denote by \( [\theta_0, \theta_1] \) the subset of \( \Theta \) where \( \delta(\theta) = 0 \), i.e. where (MN) is binding. Over this set, we have \( q(\theta) = \bar{q} \). Denoting by \( \nu \) and \( \gamma \) the Lagrange multipliers corresponding to (IC1) and (34), the Lagrangian of this program is given by

\[
\mathcal{L} = \int_{\theta_0}^{\theta_1} \left\{ [(1 + \lambda)\pi(q, \theta) - D(q) - \lambda R]g + \nu(-K/\theta^2 + a^* + q - \dot{R}) + \gamma(\delta - \dot{q}) \right\}d\theta
\]  \hspace{1cm} (35)

where \( \gamma = 0 \) whenever \( \delta > 0 \). Integration by parts gives

\[
\int_{\theta_0}^{\theta_1} \nu\dot{R}d\theta = \nu(\bar{\theta})R(\bar{\theta}) - \nu(\theta_1)R(\theta_1) - \int_{\theta_0}^{\theta_1} \dot{\nu}Rd\theta
\]

and

\[
\int_{\theta_0}^{\theta_1} \gamma\dot{q}d\theta = \gamma(\bar{\theta})q(\bar{\theta}) - \gamma(\theta_1)q(\theta_1) - \int_{\theta_0}^{\theta_1} \dot{\gamma}qd\theta
\]

Lagrangian (35) becomes

\[
\mathcal{L} = \int_{\theta_0}^{\theta_1} Hd\theta - R(\bar{\theta})\nu(\bar{\theta}) + R(\theta_1)\nu(\theta_1) - \gamma(\bar{\theta})q(\bar{\theta}) + \gamma(\theta_1)q(\theta_1)
\]

where

\[
H \equiv [(1 + \lambda)\pi(q, \theta) - D(q)]g + R[\dot{\nu} - \lambda g] + \nu(-K/\theta^2 + a^* + q) + \dot{\gamma}q + \gamma\delta
\]

Transversality conditions are deduced from

\[
\frac{\partial \mathcal{L}}{\partial R(\bar{\theta})} = -\nu(\bar{\theta}) \leq 0
\]  \hspace{1cm} (36)

and

\[
\frac{\partial \mathcal{L}}{\partial R(\theta_1)} = \nu(\theta_1) \leq 0
\]  \hspace{1cm} (37)
and since there is no condition on emission levels \( q(\bar{\theta}) \) and \( q(\bar{\theta}) \) (other than non-negativity), we have \( \gamma(\bar{\theta}) = \gamma(\bar{\theta}) = 0 \).

Pointwise maximizations give

\[
\frac{\partial H}{\partial q} = [(1 + \lambda)\theta - q]g + \nu + \dot{\gamma} = 0 \quad (38)
\]

\[
\frac{\partial H}{\partial \delta} = \gamma \leq 0 \quad (\dot{\gamma} \geq 0) \quad (39)
\]

\[
\frac{\partial H}{\partial R} = \dot{\nu} - \lambda g \leq 0 \quad (R \geq 0) \quad (40)
\]

Integrating (40) over \([\theta, \bar{\theta}]\) gives

\[
\nu(\bar{\theta}) - \nu(\theta) = \int_{\theta}^{\bar{\theta}} \dot{\nu}du \leq \lambda \int_{\theta}^{\bar{\theta}} g(u)du = \lambda [1 - G(\theta)]
\]

and thus, using (36),

\[
\nu(\theta) \geq -\lambda [1 - G(\theta)]
\]

Integrating (40) over \([\bar{\theta}, \theta]\) also gives

\[
\nu(\theta) - \nu(\bar{\theta}) = \int_{\bar{\theta}}^{\theta} \dot{\nu}(u)du \leq \lambda \int_{\bar{\theta}}^{\theta} g(u)du = \lambda G(\theta)
\]

Consequently, using (37), we have

\[
\lambda G(\theta) \geq \nu(\theta) \geq -\lambda [1 - G(\theta)]
\]

for all \( \theta \in \Theta \). To diminish the rents, having (IR) constraints binding on a \( \Theta \) subset, i.e., \( R(\theta) = 0 \) for all \( \theta \in [\alpha, \beta] \subset \Theta \), would be optimal absent the monotonicity constraint on quantities. Indeed, we would have \( \dot{R}(\theta) = 0 \) for all \( \theta \in (\alpha, \beta) \), i.e., \( q(\theta) = K/\theta^2 - a^*(\theta) \), which is decreasing. We thus have to take account of (MN), implying that \( q(\theta) = \bar{q} \) over some interval \([\theta_0, \theta_1]\), with \( \dot{R}(\theta) < 0 \) for all \( \theta < \theta_0 \) and \( \dot{R}(\theta) > 0 \) for all \( \theta > \theta_1 \). Denote by \( \theta^* \in (\theta_0, \theta_1) \) the threshold type verifying

\[
\bar{q} = K/\theta^{*2} - a^*(\theta^*)
\]
For all $\theta \in (\theta_0, \theta^*)$, we have
\[
\dot{R}(\theta) = -K/\theta^2 + a^*(\theta) + \bar{q}
\]
\[
= K(1/\theta^* - 1/\theta^2) + a^*(\theta) - a^*(\theta^*)
\]
\[
< 0
\]
hence $R(\theta) > 0$ for all $\theta \in (\theta_0, \theta^*)$, and thus $\nu(\theta) = \lambda G(\theta)$ on this set. For all $\theta \in (\theta^*, \theta_1)$, the same computations give $\dot{R}(\theta) > 0$, hence $R(\theta) > 0$ and thus $\nu(\theta) = -\lambda[1 - G(\theta)]$ on this set. Consequently, $R(\theta) = 0$ for $\theta = \theta^*$ only, and co-state $\nu$ “falls down” at $\theta^*$ from $\lambda G(\theta^*)$ to $-\lambda[1 - G(\theta^*)]$. Schedules $q_1$ and $q_2$ are deduced from (38), with $\dot{\gamma} = 0$ using $\nu(\theta) = \lambda G(\theta)$ and $\nu(\theta) = -\lambda[1 - G(\theta)]$ respectively. Since they are both increasing under our assumptions on the distribution of $\theta$, we have $\ddot{q}(\theta) = q_1(\theta)$ for all $\theta \leq \theta_0$ and $\ddot{q}(\theta) = q_2(\theta)$ for all $\theta \geq \theta_1$. For $\theta \in (\theta_0, \theta^*)$, (38) gives
\[
\dot{\gamma}(\theta) = -[\bar{q} - q_1(\theta)]g(\theta)
\]
and, for $\theta \in (\theta^*, \theta_1)$,
\[
\dot{\gamma}(\theta) = [q_2(\theta) - \bar{q}]g(\theta).
\]
Integrating over $(\theta_0, \theta^*)$ and $(\theta^*, \theta_1)$ respectively yields
\[
\gamma(\theta_1) - \gamma(\theta_0) = -\int_{\theta_0}^{\theta^*} [\bar{q} - q_1(\theta)]g(\theta)d\theta + \int_{\theta^*}^{\theta_1} [q_2(\theta) - \bar{q}]g(\theta)d\theta
\]
As $\gamma(\theta) = 0$ for all $\theta \notin [\theta_0, \theta_1]$, continuity imposes $\gamma(\theta_0) = \gamma(\theta_1) = 0$, which gives (15).

\section*{B \ Proof of proposition 3}

The condition $I(\theta) = I_0$ for all $\theta < \theta_s$ can also be written as $I(\theta) = I_0$ and $\dot{I}(\theta) = 0$ for all $\theta \in (\underline{\theta}, \theta_s)$. Denoting by $\omega(\cdot)$ the corresponding Lagrange multiplier and integrating to obtain
\[
\int_{\underline{\theta}}^{\theta} \omega(\theta)\dot{I}(\theta)d\theta = \omega(\theta)I(\theta) - \omega(\theta)I_0 - \int_{\underline{\theta}}^{\theta} \dot{\omega}(\theta)I(\theta)d\theta
\]
where \( \omega(\theta) = \dot{\omega}(\theta) = 0 \) for all \( \theta > \theta_s \), the Lagrangian of the agency’s program is given by

\[
\mathcal{L}^I = \int_{\bar{\theta}}^{\theta} H^I d\theta - R(\bar{\theta})\nu(\bar{\theta}) + R(\theta)\nu(\theta) - \gamma(\bar{\theta})q(\bar{\theta}) + \gamma(\theta)q(\theta) + \omega(\bar{\theta})\dot{I}(\bar{\theta}) - \omega(\theta)I_0
\]

with

\[
H^I \equiv [(1 + \lambda)\tilde{\pi}(q + a, a, I, \theta) - D(q)]g + R[\dot{\nu} - \lambda g] + \nu(-K/\theta^2 + a + q) + \dot{\gamma}q + \gamma\delta - \dot{\omega}(\bar{\theta})I(\bar{\theta}) + \eta(\theta - A_a) + \tau(A_I + 1)
\]

and where \( \eta(\cdot) \) and \( \tau(\cdot) \) are the Lagrange multipliers corresponding to (20) and (19) respectively, with \( \tau = 0 \) whenever \( I = I_0 \) and thus whenever \( \omega(\theta) = 0 \) (and \( \dot{\omega}(\theta) = 0 \)).

In addition to first-order conditions (38)-(40) which are unchanged, we also have to account for

\[
\frac{\partial H^I}{\partial \dot{a}} = \nu - \eta A_{aa} + \tau A_{Ia} = 0 \quad (41)
\]

\[
\frac{\partial H^I}{\partial \dot{I}} = -(1 + \lambda)(A_I + 1)g - \eta A_{Ia} + \tau A_{II} - \dot{\omega}(\theta) = 0 \quad (42)
\]

and

\[
\frac{\partial \mathcal{L}}{\partial I_0} = -\omega(\theta) = 0 \quad (43)
\]

For \( \theta > \theta_s \), as \( \dot{\omega}(\theta) = 0 \), (41) and (42) allow us to deduce \( \eta(\theta) \) and \( \tau(\theta) \), while binding (19) and (20) give \( \hat{a}(\theta) = a^*(\theta) \) and \( \hat{I}(\theta) = I^*(\theta) \). For \( \theta < \theta_s \), we have \( \tau(\theta) = 0 \), which gives \( \eta = \nu/A_{aa} \) and thus

\[
\dot{\omega}(\theta) = -(1 + \lambda)[A_I(\hat{a}(\theta), I_0) + 1]g(\theta) - \nu(\theta)\frac{A_{Ia}(\hat{a}(\theta), I_0)}{A_{aa}(\hat{a}(\theta), I_0)}
\]

Integrating over \( (\theta, \theta_s) \) and using \( \omega(\theta_s) = 0 \) (by continuity) yields

\[
\int_{\theta_s}^{\theta} \left\{ (1 + \lambda)[A_I(\hat{a}(\theta), I_0) + 1] + \frac{A_{Ia}(\hat{a}(\theta), I_0) \nu(\theta)}{A_{aa}(\hat{a}(\theta), I_0) g(\theta)} \right\} g(\theta)d\theta = 0 \quad (44)
\]

where \( \nu(\theta) = \lambda G(\theta) \) for \( \theta < \theta^* \) and \( \nu(\theta) = -\lambda[1 - G(\theta)] \) for \( \theta > \theta^* \). Assume \( \theta_s > \theta^* \).

We can rewrite (44) as

\[
\int_{\theta_s}^{\theta^*} f_1(\theta)g(\theta)d\theta + \int_{\theta_s}^{\theta_s} f_2(\theta)g(\theta)d\theta = 0
\]
where
\[ f_1(\theta) \equiv (1 + \lambda)[A_I(\hat{a}(\theta), I_0) + 1] + \frac{\lambda A_{Ia}(\hat{a}(\theta), I_0) G(\theta)}{A_{aa}(\hat{a}(\theta), I_0)} g(\theta) \]
and
\[ f_2(\theta) \equiv (1 + \lambda)[A_I(\hat{a}(\theta), I_0) + 1] - \lambda \frac{A_{Ia}(\hat{a}(\theta), I_0) 1 - G(\theta)}{A_{aa}(\hat{a}(\theta), I_0)} g(\theta) \]

As \( f_2(\theta) > 0 \) for all \( \theta > \theta^* \), we must have \( \int_{\theta}^{\theta^*} f_1(\theta) g(\theta) d\theta < 0 \). However, as \( f_1(\theta) > 0 \) and \( f_1 \) and \( g \) are continuous, there exists \( \theta_2 < \theta^* \) such that \( \int_{\theta_2}^{\theta^*} f_1(\theta) g(\theta) d\theta = 0 \), i.e. \( \theta_2 = \theta_s \), hence a contradiction. We thus have \( \theta_s \leq \theta^* \), and (44) simplifies to (21).

C  Proof of Lemma 2

(IC1) and (MN) are derived from standard arguments (see, e.g. Guesnerie and Laffont, 1984). Define \( M \equiv \min_\theta \{ t(\bar{\theta}) + \mu(\bar{\theta}) \bar{F} \} \). From (IC'), to deter tax evasion, we must have
\[ M + I_0 \geq K/\theta + \theta \bar{y} - R(\theta). \]

(IC2) follows from
\[ \frac{d}{d\theta} [K/\theta + \theta \bar{y} - R(\theta)] = \bar{y} - [q(\theta) + \hat{a}(\theta)] \geq 0 \]

We also have
\[ M \leq t(\theta) + \mu(\theta) \bar{F} \]
\[ = [\bar{\pi}(\hat{a}(\theta) + q(\theta), \hat{a}(\theta), \bar{I}(\theta), \theta) - R(\theta)] + \mu(\theta) \bar{F} \]
using (13). Since the agency’s objective decreases in \( \mu(\theta) \), the constraint is binding at the optimum, which gives (29), and \( M = t(\theta) + \mu(\theta) \bar{F} \) for all \( \theta \). Finally, we have
\[ \mu(\theta) \bar{F} = M + R(\theta) - \bar{\pi}(\hat{a}(\theta) + q(\theta), \hat{a}(\theta), \bar{I}(\theta), \theta) \]
\[ > M + R(\theta) - (K/\theta + \theta \bar{y} - I_0) \geq 0 \]
where the last inequality comes from (IC').
D Proof of Proposition 4

Using binding (IC2), the Lagrangian of the agency’s program is given by

\[ \mathcal{L} = \int_{\bar{\theta}}^{\bar{\theta}} H \, d\theta - [K/\bar{\theta} + \bar{\theta}\bar{y} - I_0 - M]\nu(\bar{\theta}) + R(\bar{\theta})\nu(\bar{\theta}) - \gamma(\theta)q(\theta) + \gamma(\bar{\theta})q(\bar{\theta}) + \omega(\bar{\theta})I(\bar{\theta}) - \omega(\theta)I_0 - Mc/\bar{F} \]

where

\[ H \equiv [(1 + c/\bar{F})\pi(q + a, a, I, \theta) - D(q)]g + R(\dot{\nu} - gc/\bar{F}) + \nu(-K/\theta^2 + a + q) \]

\[ + \dot{\gamma}q + \gamma\delta - \dot{\omega}(\theta)I(\theta) + \eta(\theta - A_a + \tau(A_I + 1) \]

using the same notations as above. Transversality conditions are deduced from

\[ \frac{\partial \mathcal{L}}{\partial M} = \nu(\bar{\theta}) - c/\bar{F} = 0 \quad (45) \]

\[ \frac{\partial \mathcal{L}}{\partial R(\bar{\theta})} = \nu(\bar{\theta}) \leq 0 \quad (46) \]

\[ \frac{\partial \mathcal{L}}{\partial I_0} = \nu(\bar{\theta}) - \omega(\bar{\theta}) \leq 0 \quad (47) \]

and we have \( \gamma(\bar{\theta}) = \gamma(\theta) = 0 \). Pointwise maximizations give

\[ \frac{\partial H}{\partial q} = (1 + c/\bar{F})\theta - q]g + \nu + \dot{\gamma} = 0 \quad (48) \]

\[ \frac{\partial H}{\partial \delta} = \gamma \leq 0 \quad (\dot{q} \geq 0) \quad (49) \]

\[ \frac{\partial H}{\partial \dot{R}} = \dot{\nu} - gc/\bar{F} \leq 0 \quad (R \geq 0) \quad (50) \]

\[ \frac{\partial H}{\partial A_a} = \nu - \eta A_{aa} + \tau A_{la} = 0 \quad (51) \]

\[ \frac{\partial H}{\partial I} = -(1 + c/F)(A_i + 1)g - \eta A_{Ia} + \tau A_{II} - \dot{\omega}(\theta) = 0 \quad (52) \]

Integrating (50) over \([\theta, \bar{\theta}]\) yields

\[ \nu(\bar{\theta}) - \nu(\theta) = \int_{\theta}^{\bar{\theta}} \dot{\nu} \, du \leq (c/\bar{F}) \int_{\theta}^{\bar{\theta}} g(u) \, du = [1 - G(\theta)]c/\bar{F} \]

and thus, using (45),

\[ \nu(\theta) \geq G(\theta)c/\bar{F} \]
Integrating (50) over \([\underline{\theta}, \theta]\) also gives
\[
\nu(\theta) - \nu(\underline{\theta}) = \int_{\underline{\theta}}^{\theta} \dot{\nu}(u) du \leq (c/F) \int_{\underline{\theta}}^{\theta} g(u) du = G(\theta)c/F
\]

Consequently, using (46), we have
\[
G(\theta)c/F \geq \nu(\theta) \geq G(\theta)c/F
\]
hence
\[
\nu(\theta) = G(\theta)c/F
\]
for all \(\theta \in \Theta\). Since \(\dot{\nu}(\theta) = g(\theta)c/F\) for all \(\theta \in \Theta\), we have \(R(\theta) > 0\) for all \(\theta \in \Theta\). Plugging into (48) and assuming \(\dot{\gamma} = 0\) yields (30), which is increasing. Hence \(\gamma(\theta) = 0\) for all \(\theta\), and (30) corresponds to the optimal emission schedule.

For \(\theta > \theta_s\), as \(\dot{\omega}(\theta) = 0\), (51) and (52) allow us to deduce \(\eta(\theta)\) and \(\tau(\theta)\), while binding (19) and (20) give \(\hat{a}(\theta) = a^*(\theta)\) and \(\hat{I}(\theta) = I^*(\theta)\). For \(\theta < \theta_s\), we have \(\tau(\theta) = 0\), which gives \(\eta = \nu/A_{aa}\) and thus
\[
\dot{\omega}(\theta) = -(1 + c/F)[A_I(\hat{a}(\theta), I_0) + 1]g(\theta) - \frac{c}{F} G(\theta) \frac{A_{ta}(\hat{a}(\theta), I_0)}{A_{aa}(\hat{a}(\theta), I_0)}
\]
Integrating over \([\underline{\theta}, \theta_s]\), using \(\omega(\theta_s) = 0\) (by continuity), (47) and (45) yield (31).