CentER

Discussion Paper

No. 2005–09

POLLUTION STANDARDS, COSTLY MONITORING AND FINES

By Carmen Arguedas

January 2005

ISSN 0924-7815

Tilburg University

Pollution Standards, Costly Monitoring and Fines

Carmen Arguedas*

Departamento de Economía, Universidad Europea de Madrid

Abstract

We investigate the features of optimal regulatory policies composed of pollution standards and probabilities of inspection, where fines for non-compliance depend not only on the degree of violation but also on non-gravity factors. We show that optimal policies can induce either compliance or noncompliance with the standards, the latter being more plausible when monitoring costs are large and, surprisingly, when gravity-based fines are large. Also, both the convexity of the sanctions and the level of the non-gravity-based penalties play a key role as to whether optimal policies induce noncompliance.

JEL classification: D82, K32, K42, L51.

Key words: standards, monitoring, convex fines, gravity-based sanctions, non-gravity-based sancions, noncompliance.

*Address for correspondence: Departamento de Economía, Universidad Europea de Madrid. C/ Tajo, s/n. 28670 Villaviciosa de Odón, Madrid, Spain. Tel. (+34) 91 211 5645. Fax. (+34) 91 616 8265. E-Mail: mcarmen.arguedas@uem.es. This research was conducted when I was visiting the Environmental Economics Group leaded by Aart de Zeeuw at Tilburg University in 2004. I wish to thank them all for their helpful comments and hospitality.

1 Introduction

Environmental regulations often require polluting agents to comply with recommended pollution limits or *standards* while they engage in production activities.[1] Normally, authorities do not observe the performance of firms with respect to such limits, unless they incur in costly monitoring. Regulators then design enforcement policies which consist of frequencies of inspection and sanctions in case firms are found violating the standards.

Generally, penalties for noncompliance have two components.[2] The *gravity-based sanction* is directly related to the degree of noncompliance, accounting for factors such as the seriousness of the violation, the degree of culpability involved, mitigation efforts or any history of prior violations. The *non-gravity-based sanction* considers additional factors such as the economic impact of the penalty on the violator or other matters that justice may require. Generally, there exists flexibility regarding the final structure of the penalty to be imposed, with the limit of the violator's statutory maximum civil penalty liability.

While this structure of fines plays a key role in the likelihood of compliance with environmental rules, surprisingly it has not been yet studied in the literature.[3]

In this paper, we investigate this issue. For that purpose, we consider a model in which the regulator sets a pollution standard and a probability of inspection, considering the optimal behavior of the firm with respect to the policy. If a firm is discovered violating the standard, then it is forced to pay a sanction, which

[1] An example is EPA's National Pollutant Discharge Elimination System Program, where facilities which discharge pollutants into waters of the US are required to obtain a permit to release specific amounts of pollution.

[2] See, for example, the Civil Penalty Policy of the Clean Water Act.

[3] See Heyes (2000) for a complete review.

has both fixed (non-gravity-based) and non-fixed (gravity-based) components, the latter being strictly increasing and convex in the degree of violation.[4]

We find that the optimal policy can induce either compliance or noncompliance with the environmental standards. In any event, the induced pollution level is always larger than the efficient pollution level, that is, the level that maximizes social welfare in the hypothetical case that enforcement costs were zero. Therefore, there exists a trade-off between efficiency and monitoring costs: the larger the monitoring costs, the larger the difference between the efficient and the induced pollution levels.

In the case the gravity-based sanction is linear in the degree of noncompliance, we undoubtedly find that the optimal policy induces compliance, regardless of the level of the non-gravity component This is so because the optimal pollution level does not depend on the degree of noncompliance, and therefore, it is not affected by the standard. Then, it is not possible to save monitoring costs by means of a reduced standard (or, equivalently, a larger fine).

If the gravity-based sanction is strictly convex, we find that the optimal policy can induce either compliance or noncompliance. In the event of noncompliance, we show that the optimal standard is zero. Here, the induced pollution level depends negatively on the degree of noncompliance (then, positively on the standard) and negatively on the probability of inspection, and it is possible to keep a given pollution level decreasing both the standard and the inspection probability accordingly, then saving monitoring costs.

Whether the optimal policy induces compliance or noncompliance depends on the monitoring costs and the level of the sanctions. Regarding the former,

[4] The convexity of the sanction is usually accepted in both efficiency and equity terms (for example, see Shavell (1992)).

we show that it is more likely that the optimal policy entails noncompliance when monitoring costs are high. This result is reasonable, reflecting the fact that when inspections are very costly, monitoring may be insufficient to induce firms' compliance.

With respect to the latter, we find that compliance is more likely when the non-gravity-based sanction is large enough. Surprisingly, however, it is more likely that the optimal policy induces noncompliance when the gravity-based sanction is sufficiently large. In fact, this latter result is in general contrary to the one predicted by the literature on either crime or tax evasion, where an increase in the sanction increases compliance.[5] However, there the standard − or the level which delimits a certain behavior to be punishable or not− is given. Here, the standard is determined endogenously in the model, and therefore, it changes when there is a change in the level of the sanctions.

The explanation of the result is as follows. The social welfare evaluation of each policy is determined by the firm's induced behavior to such a policy and the monitoring costs. In our model, each possible pollution level e can be induced by two kinds of policies. One is a *compliant policy*, which consists of announcing a standard equal to such pollution level and the probability of inspection necessary to induce the agent to comply with the standard. The other one is a *noncompliant policy*, which consists of announcing a standard equal to zero and the probability of inspection that induces the agent to pollute e. The cheapest policy is the one with the lowest inspection probability. When the gravity-based sanction increases, the degree of noncompliance to a given standard decreases, and, given the convexity of the sanction, the smaller the

[5] Some exceptions include Kambhu (1989), Malik (1990), Andreoni (1991), Harrington (1988) or Livernois and McKenna (1999).

standard, the more it decreases. Thus, the probability needed to induce a certain pollution level e decreases in both types of policies, but it decreases more the one associated with the noncompliant policy. Consequently, for sufficiently large gravity-based sanctions, the optimal policy may induce noncompliance.

The literature on environmental regulation has not considered the two components of the penalties analyzed here. Within the closest literature, there are some papers which determine optimal policies constrained to the subset of policies which induce compliance, such as Ellis (1992a) or Amacher and Malik (1996), the latter one in a bargaining context. More recently, Arguedas and Hamoudi (2004) and Arguedas (2005) study optimal regulatory policies, allowing for noncompliance. Both assume strictly convex gravity-based and zero non-gravity-based penalties, and also that firms can decide on technology investment as well. It is obtained that the optimal policy always induces noncompliance, not necessarily to a zero standard in the latter case. Therefore, these previous results are in accordance with the ones presented in this paper, in the sense that noncompliance is more likely when non-gravity-based penalties are small enough.

The remainder of the paper is organized as follows. In the next section, we present the model. In Section 3, we study the optimal behavior of the firm with respect to each feasible policy. In Section 4, we analyze the characteristics of the optimal policy. We conclude in Section 5. All the proofs are in the Appendix.

2 The Model

We consider a firm that generates pollution as a by-product of its production activity. This firm obtains private profits dependent on the pollution level $e \geq 0$ and represented by the function $b(e)$, which is strictly concave with an interior maximum at $\widetilde{e} > 0$. Also, $b(0) = 0$ and $b'''(e) \geq 0$ and sufficiently small.

Pollution generates external damages measured by the function $d(e)$, which is strictly increasing and convex in e, and such that $d(0) = 0$.

Let e^w be the efficient pollution level, i.e., $e^w = \arg\max_{e \geq 0} \{b(e) - d(e)\}$. In the absence of a regulation, the firm would not internalize the presence of external damages and pollution would be $\widetilde{e} = \arg\max_{e \geq 0} b(e) > e^w$.

We assume there exists a regulator who sets a standard $s \geq 0$, that is, a maximum permitted pollution level. The regulator cannot observe the pollution level selected by the firm unless it engages in a monitoring activity, which is costly and perfectly accurate. The cost *per* inspection is $c > 0$. Since inspection is costly, it is generally not desirable to inspect the firm in every instance but only occasionally. Therefore, the regulator also sets the probability of inspection, $p \in [0, 1]$. Once inspected, if the firm is discovered to be violating the standard, i.e., if $e > s$, then it is forced to pay a penalty that depends on the degree of noncompliance, $e - s$. Formally, we assume that the sanction takes the following structure:

$$f(e-s) = \begin{cases} a + g(e-s) + h(e-s)^2 & \text{if } e-s > 0 \\ 0 & \text{if } e-s \leq 0 \end{cases} \tag{1}$$

where $a \geq 0$, $g > 0$, $h \geq 0$ and $g^2 > 2ah$. Observe that a is the non-gravity-based sanction. If $a > 0$, then the sanction is discontinuous at 0. We assume that the sanction is fixed by a government entity other than the regulator, for example, the judiciary.[6]

We consider a principal-agent framework in which the regulator chooses the social welfare maximizing regulatory policy considering the optimal response of the firm to a given policy. We consider the subgame perfect equilibrium concept. Therefore, we solve the problem backwards, that is, we first find the firm's optimal pollution level in response to a given policy, and we then find the optimal policy that maximizes social welfare considering the firm's optimal response.

Given a feasible policy $\{s, p\}$, the firm chooses the pollution level that maximizes its expected payoff, that is, its private profits minus its expected penalties, which is represented as follows:

$$P(s,p) = \max_{e \geq 0} \{b(e) - pf(e - s)\} \tag{2}$$

Let $e(s,p)$ be the firm's optimal pollution level given $\{s, p\}$, that is, $e(s,p) = \arg\max_{e \geq 0} \{b(e) - pf(e - s)\}$. Observe that $e(s,p) \leq \widetilde{e}$, the pollution level selected by the firm in the absence of regulation. Given the assumptions of the model, $e(s,p)$ is unique if the sanction is continuous at 0 (that is, if $a =$

[6] This assumption is common in the literature on environmental regulation, except in Heyes (1996) and Arguedas (2005), for example. In other contexts, such as crime, there are several papers devoted to endogenize fines together with probabilities of inspection, such as in Becker (1968), Polinsky and Shavell (1979, 1990) and Bebchuck and Kaplow (1991), among others. In the context of tax evasion, few papers consider endogenous fines. See, for instance, Mookherjee and Png (1989) and Pestieau et al. (1997).

0). Otherwise, $e(s,p)$ is unique except when the firm is indifferent between complying and not complying with the standard. In this latter case, we assume that the firm chooses to comply.

Considering the firm's optimal response, the regulator now selects the policy that maximizes social welfare. We define social welfare to be:

$$SW(s,p) = P(s,p) - d(e(s,p)) + pf(e(s,p) - s) - pc \tag{3}$$

Thus, the regulator is concerned about the expected payoff of the firm, the generated damages, the expected collected fines and the expected monitoring costs. We assume that there are no social costs associated with the collection of the fines, and that fines are redistributed lump-sum, that is, they do not distort behavior. Also, note that we do not impose any budget requirement on the monitoring activity.[7]

An important assumption of our model is that the regulator commits to the announced regulatory policy, that is, it monitors the firm with the announced frequency, even when, in some instances, it would find it profitable to deviate from the announced policy. Commitment can be justified considering that the regulator has to build up a reputation, that is, it has to make his policy announcements credible among the firms to induce the desired behavior.[8]

[7] We could consider consumers' surplus in social welfare. However, since consumers are passive actors in our model, this would not add any qualitative differences to the results. Obviously, there would be quantitative differences, since consumers' consideration might result in more lenient regulations.

[8] A formal justification of this assumption would require the consideration of a dynamic model, which is beyond the scope of this paper. In static models such as ours, the assumption of commitment is common in almost all the literature. Some exceptions include Ellis (1992b) and Grieson and Singh (1990).

In the next section, we study the behavior of the firm with respect to the announced policy.

3 The Optimal Behavior of the Firm

Consider a policy $\{s, p\}$, such that $s \in [0, \widetilde{e}]$ and $p \in [0, 1]$.[9] The firm solves the following problem:

$$\max_{e \geq 0} \{b(e) - pf(e - s)\} \tag{4}$$

where $f(e - s)$ is given by (1). Let $e(s, p)$ be the solution to problem (4).

Given $\{s, p\}$, if the firm complies with the standard, it is not penalized. By the assumptions of the model, the *optimal compliance decision* is s and the firm's payoff is $b(s)$.

If the firm violates the standard, i.e., $e > s$, there is a chance that the firm is inspected and punished. In this case, its expected payoff is $b(e) - p\left(a + g(e - s) + h(e - s)^2\right)$. Let $n(s, p) > s$ be the *optimal noncompliance decision* and let $\pi(s, p)$ be the corresponding noncompliance payoff. Observe that $n(s, p) \leq \widetilde{e}$. By the assumptions of the model, the condition that characterizes the optimal noncompliance decision is the following:

$$b'(e) = p(g + 2h(e - s)) \tag{5}$$

The optimal noncompliance decision and the corresponding firm's payoff are

[9] The regulator is not interested in a standard larger than $\widetilde{e}$, that is, the pollution level that the firm would select in the absence of regulation.

well-defined only if $b'(s) > pg$. For example, consider Figures 1 and 2, where we represent two possible cases of the firm's expected payoff to a given policy $\{s, p\}$, for the different values of the pollution level. In Figure 1, the optimal noncompliance decision exists (in fact, noncompliance is preferred in this case), while it does not exist in Figure 2. In both figures, the sanction is discontinuous at zero, which implies that the firm's expected payoff is discontinuous at $e = s$.

Implicitly differentiating (5), we obtain $n_p(s, p) = \frac{g+2h(n-s)}{b''(n)-2hp}$ and $n_s(s, p) = -\frac{2hp}{b''(n)-2hp}$. By the assumptions of the model, we then have $n_p(s, p) < 0$ and $0 \leq n_s(s, p) < 1$. That is, the pollution level selected by the firm in the event of noncompliance strictly decreases when the probability of inspection increases, and it increases when the standard increases. However, since $n_s(s, p) < 1$, the degree of violation strictly decreases when the standard increases.[10]

Now, given $\{s, p\}$ such that $b'(s) > pg$, the firm selects whether to comply or not depending on the expected payoff of each possibility. Thus, the firm's optimal response is the following:

$$e(s, p) = \begin{cases} s & \text{if } b(s) \geq \pi(s, p) \\ n(s, p) & \text{if } b(s) < \pi(s, p) \end{cases} \tag{6}$$

The firm's expected payoff given by (2) can be further expressed as:

$$P(s, p) = \max \ \{b(s), \pi(s, p)\} \tag{7}$$

For example, considering Figure 1 again, we have $b(s) < \pi(s, p)$ and, in consequence, $e(s, p) = n(s, p)$. Therefore, $P(s, p) = \pi(s, p)$.

[10] Note that $n_s(s, p) = 0$ when either $h = 0$ (i.e., when the sanction is linear), or $p = 0$ or both.

In the following lemma, we present the properties of $P(s,p)$.

Lemma 1 *The function $P(s,p)$ is nondecreasing and concave in s, nonincreasing and convex in p, and has a nonnegative cross partial. Moreover, $P(s,0) = b(\widetilde{e})$ for all $s \in [0,\widetilde{e}]$.*

For the subset of policies such that $b'(s) > pg$, we now characterize the set of policies where the firm is indifferent between complying and noncomplying, which is given by the following equation:

$$b(s) = \pi(s,p) \tag{8}$$

This expression defines an implicit relationship between p and s, i.e., a mapping $p^c(s)$ such that (8) holds. Given s, $p^c(s)$ represents the minimum inspection probability necessary to induce the firm to comply with the standard.[11] In the following lemma, we present the properties of $p^c(s)$.

Lemma 2 *The mapping $p^c(s)$ is strictly decreasing and convex in s. Also, $p^c(\widetilde{e}) = 0$. Moreover, if $\max_{e>0}\left\{b(e) - \left(a + g(e-s) + h(e-s)^2\right)\right\} > 0$, then there exists $\widehat{s} \in (0,\widetilde{e})$ such that $p^c(\widehat{s}) = 1$.*

In Figure 3, we represent the function $p^c(s)$ in the space of feasible policies $\{s,p\}$. We also include the firm's indifference map, where each indifference curve is composed of the set of policies $\{s,p\}$ which give the firm the same expected payoff. On the horizontal axis, we measure the standard and on the vertical axis we measure the probability of inspection.

[11] When the sanction is continuous at 0, (8) never holds since the maximand in (2) is continuous at s. In this case, the firm complies with the standard only if $b'(s) \leq pg$. Therefore, the threshold probability between the compliance and violation decisions in that case is given by $p^c(s) = \frac{b'(s)}{g}$, which is strictly decreasing and convex in s, by the assumptions of $b(e)$. We will refer to this case later on (see Lemma 4, below).

Observe that standards smaller than $\widehat{s}$ cannot be enforced since $p^c(\widehat{s}) = 1$. Concerning the shape of the function $p^c(s)$, enforcing a small pollution level requires more monitoring effort than enforcing a large pollution level, since profits for the firm are strictly increasing and concave in the pollution level; also, the probability of inspection decreases more than proportionally when the standard increases. Above the curve $p^c(s)$, the agent strictly prefers to comply with the regulation; and below $p^c(s)$, the agent strictly prefers not to comply. Consequently, $p^c(s)$ divides the space of feasible policies into two regions, namely the *compliance* and the *noncompliance* regions.

In the compliance region, the firm's expected payoff is $b(s)$, that is, it does not depend on the probability of inspection. Therefore, indifference curves in that region have an infinite slope. In the noncompliance region, the expected payoff is $\pi(s,p)$ and the indifference curves when $p > 0$ have a positive slope.[12] Moreover, by Lemma 1, the expected payoff of the firm increases to the south-east, i.e., whenever the standard is larger and the probability of inspection is smaller.

We now analyze the firm's optimal response to each feasible policy. Consider the case in which fines are strictly convex in the degree of violation (i.e., $h > 0$) and take, for instance, a standard $\overline{s}$. From the previous analysis, there exists a probability of inspection $p^c(\overline{s})$ such that, for all $p \geq p^c(\overline{s})$, the optimal response is $e = \overline{s}$, i.e., to comply with the standard. At $p = p^c(\overline{s})$, the firm

[12] To see the shape of an indifference curve in the noncompliance region, we implicitly differentiate $\pi(s,p) = b(n) - p\left(a + g(n-s) + h(n-s)^2\right) = k$, where $n = n(s,p)$:

$$\frac{dp}{ds}\mid_{\pi=k} = p\frac{g + 2h(n-s)}{a + g(n-s) + h(n-s)^2} > 0$$

is indifferent between complying and not complying, i.e., between choosing either $\overline{s}$ or $n(\overline{s}, p^c(\overline{s})) = \arg\max_{e>s}\left\{b(e) - p^c(\overline{s})\left(a + g(e-s) + h(e-s)^2\right)\right\}$. Therefore, behavior is discontinuous at $p = p^c(\overline{s})$. Finally, at $p < p^c(\overline{s})$, the optimal response is to violate the standard, i.e., $e = n(\overline{s}, p)$, which is strictly decreasing and convex in the probability of inspection. (Note that $n_{pp}(s,p) > 0$ since $b''' \geq 0$). Given $\overline{s}$, $p^n_{\overline{s}}(e)$ denotes the inverse of the function $n(\overline{s}, p)$, and represents the probability necessary to induce the pollution level e, under the standard $\overline{s}$, where $e > \overline{s}$. Since $n(\overline{s}, p)$ is strictly decreasing and convex in p, the inverse function $p^n_{\overline{s}}(e)$ is also strictly decreasing and convex in e. Observe that $p^n_{\overline{s}}(\widetilde{e}) = 0$, since $n(\overline{s}, 0) = \widetilde{e}$.

Consider now an alternative standard $\widetilde{s}$, such that $\widetilde{s} > \overline{s}$. On the one hand, by Lemma 2, the minimum probability necessary to induce compliance with $\widetilde{s}$ is lower than that necessary to induce compliance with $\overline{s}$, i.e., $p^c(\widetilde{s}) < p^c(\overline{s})$. At $p = p^c(\widetilde{s})$, analogously, there exists a discontinuity. And for $p < p^c(\widetilde{s})$, the firm's optimal choice is to violate $\widetilde{s}$, i.e., $e = n(\widetilde{s}, p)$. Finally, since the optimal noncompliance decision increases when the standard level increases, we then have $n(\widetilde{s}, p) > n(\overline{s}, p)$, for all $0 < p < p^c(\widetilde{s})$. Therefore, we have $p^n_{\overline{s}}(e) < p^n_{\widetilde{s}}(e)$, for all e. Obviously, when $p = 0$, we have $n(\widetilde{s}, 0) = n(\overline{s}, 0) = \widetilde{e}$ and, therefore, $p^n_{\overline{s}}(\widetilde{e}) = p^n_{\widetilde{s}}(\widetilde{e}) = 0$.

When sanctions are linear (i.e., $h = 0$), we have $n_s(s,p) = 0$. Therefore, $p^n_{\overline{s}}(e) = p^n_{\widetilde{s}}(e)$, for all e and for all $\overline{s}$, $\widetilde{s}$ such that $\overline{s} \neq \widetilde{s}$.

Once we have studied the firm's optimal response to each feasible policy, we are now ready to analyze the features of the optimal policy.

4 The Optimal Regulatory Policy

The regulator anticipates the firm's optimal response, and takes it into account when selecting the optimal policy. The regulator then solves:

$$\max_{s,p} SW(s,p) = P(s,p) - d(e) + pf(e-s) - cp \tag{9}$$

where $P(s,p)$ is given by (7), $e = e(s,p)$ is given by (6), $s \in [0, \widetilde{e}]$ and $p \in [0,1]$.

Considering (6) and (7), we can reduce (9) to the following problem:

$$\begin{aligned} \max_{s,p} \quad & \{b(e) - d(e) - cp\} \\ \text{s.t.} \quad & e = e(s,p) \end{aligned} \tag{10}$$

where $e(s,p)$ is given by (6). Since fines are costless transfers, they are not in the objective function in (10), although they affect the firm's induced behavior given by $e = e(s,p)$.

Remember that $e^w = \arg\max_{e \geq 0} \{b(e) - d(e)\}$ is the efficient pollution level and $e^w < \widetilde{e}$, the pollution level selected by the firm in the absence of regulation. We denote by $p^n(e)$ the probability needed to induce the pollution level e under the standard $s = 0$, that is $p^n(e) \equiv p_0^n(e)$. Also, recall that $p^c(e)$ represents the minimum probability needed to induce the firm to comply with the standard $s = e$. In the next proposition, we characterize each one of the two possibilities for the optimal policy.[13]

[13] In the proposition we implicitly assume that both $p^c(e^w)$ and $p^n(e^w)$ exist (i.e., both are at most 1). If any of them does not exist, we trivially have $e(s^*, p^*) > e^w$.

Proposition 3 *If the optimal policy* (s^*, p^*) *induces compliance, then* $s^* > e^w$ *and* $p^* < p^c(e^w)$. *Conversely, if it induces noncompliance, two cases are possible:*

(i) *If gravity-based sanctions are strictly convex in the degree of violation, then* $s^* = 0$, $p^* < p^n(e^w)$ *and* $e(0, p^*) > e^w$.

(ii) *If gravity-based sanctions are linear, then* $s^* \in [0, \overline{s})$ *where* $\overline{s}$ *is such that* $p^* = p^c(\overline{s})$, $p^* < p^n(e^w)$ *and* $e(0, p^*) > e^w$.

The proposition shows the characteristics of the optimal policy constrained to induce either compliance or noncompliance. The solution to (10) is the possibility that yields the highest expected social welfare.

In Figures 4 and 5, we represent, respectively, the optimal policy inducing either compliance or noncompliance. On the horizontal axis, we measure the pollution level and on the vertical axis we measure the probability of inspection. In both graphs, we represent the map of *iso-welfare* contours for the regulator, where each iso-welfare curve represents the set of pairs pollution level - inspection probability where social welfare is constant. These contours are *inverse-U* shaped, and welfare decreases in all directions from the bliss point $(e^w, 0)$.[14]

From the analysis of the previous section, we can see that each possible pollution level e can be induced by two kinds of policies. On the one hand, the regulator might announce a *compliant policy*, that is, a standard $s = e$ and the minimum probability of inspection necessary to induce the agent to comply with the standard, given by $p^c(e)$. On the other hand, it is possible to

[14] To see this, we implicitly differentiate the equation $b(e) - d(e) - cp = k$ to obtain $\frac{dp}{de}\mid_k = \frac{b'(e) - d'(e)}{c}$, which is positive for $e < e^w$; equal to 0 for $e = e^w$; and negative for $e > e^w$. Moreover, $\frac{d^2p}{de^2}\mid_k = \frac{b''(e) - d''(e)}{c} < 0$.

announce a *noncompliant policy* which, in the case of strictly convex gravity-based fines, consists of a standard $s = 0$ and the probability necessary to induce the pollution level e under $s = 0$, given by $p^n(e)$. Thus, $p^n(e)$ represents the minimum probability necessary to induce the agent to pollute the level e in the event of noncompliance. To see this, note that the pollution level chosen by the firm increases as the prescribed standard increases, and it decreases as the probability of inspection increases. Therefore, if we assume that the policy $\{s, p\}$, for $s > 0$, leads the agent to choose a pollution level e larger than s, then there exists another policy, where both p and s are smaller, and such that it induces the agent to select the same pollution level. Clearly, the second policy is cheaper than the first one, since the probability of inspection is smaller. Thus, the cheapest way to induce the agent to pollute a certain level e is with a standard $s = 0$ and the probability $p^n(e)$.

However, in the linear case, the optimal noncompliant decision does not vary with the standard. In that case, the noncompliant policy to induce a given pollution level e is given by a standard $s \in [0, e)$ and the probability necessary to induce the pollution level e under the standard $s \in [0, e)$. Since $n_s(s, p) = 0$ in this case, we then have $p^n(e) = p_s^n(e)$, for $s < e$. Thus, both cases depicted in Figures 4 and 5 correspond to the case in which gravity-based penalties are strictly convex.

Both functions $p^c(e)$ and $p^n(e)$ are strictly decreasing and convex in e, as shown in Section 3. The feature that both functions cross in the way depicted in Figure 4 is established in Lemma 4, below.

Proposition 3 shows that, whenever monitoring is costly, the optimal policy always results in inducing the agent to overpollute, i.e., to choose a pollution level larger than the efficient level e^w. Therefore, there exists a trade-off between efficiency and monitoring costs; at the optimum, the marginal loss in efficiency is equal to the marginal savings in monitoring costs, independently of whether the policy induces compliance or noncompliance.

It can be easily verified, using both (24) and (26), that when monitoring costs increase, the new optimal policy results in a smaller probability of inspection and a larger chosen pollution level. Graphically, an increase in monitoring costs can be represented by flatter iso-welfare contours and no change in either $p^c(e)$ or $p^n(e)$.

Until now, we have characterized the optimal policy constrained to induce either compliance or noncompliance. The next step is to analyze the factors that influence which is the case. In the following lemma, we show that the situations depicted in Figures 4 and 5 are not the only ones we may have. There can be cases in which the function $p^c(e)$ lies below the function $p^n(e)$, depending on the level and convexity of the sanctions. But if the functions cross, they do so as depicted in Figure 4.

Lemma 4 *(i) If $a = 0$, then $p^c(e) = p^n(e)$ for all e when gravity-based sanctions are linear, and $p^c(e) > p^n(e)$ for all $e < \widetilde{e}$ when gravity-based sanctions are strictly convex.*
(ii) If $a > 0$, $p^c(e) < p^n(e)$ for all $e < \widetilde{e}$ when gravity-based sanctions are linear. When they are strictly convex, there exists $\underline{a}$, $\overline{a}$, $\overline{a} > \underline{a}$, such that for all $a < \underline{a}$, $p^c(e) > p^n(e)$ for all $e < \widetilde{e}$; for all $a \in [\underline{a}, \overline{a}]$, $p^c(e)$ and $p^n(e)$ cross once at a point other than $(\widetilde{e}, 0)$ and, at the point of intersection, $p^{n\prime}(e) < p^{c\prime}(e) < 0$; and for all $a > \overline{a}$, $p^c(e) < p^n(e)$ for all $e < \widetilde{e}$.

Thus, there is no doubt about the type of optimal policy that we obtain in each case, except when penalties include the non-gravity-based component and the gravity-based sanction is strictly convex. In this case, we have shown that the solution depends on the location of the point of intersection between $p^c(e)$ and $p^n(e)$. In the following lemma, we show that the location of this point depends on the level and the convexity of the sanction.

Lemma 5 *Given $h > 0$ and $a \in [\underline{a}, \overline{a}]$, when the gravity-based sanction increases, the new point of intersection between $p^c(e)$ and $p^n(e)$ is such that e is smaller and the probability of inspection is larger.*

Combining Lemmas 4 and 5, we then have the main result:

Proposition 6 *When the gravity-based sanction is linear, the optimal policy always induces compliance. When the gravity-based sanction is strictly convex, three cases are possible:*

(i) If $a < \underline{a}$, the optimal policy induces noncompliance.

(ii) If $a \in [\underline{a}, \overline{a}]$, the optimal policy induces noncompliance when monitoring costs are sufficiently large, and also when the gravity-based sanction is sufficiently large.

(iii) If $a > \overline{a}$, the optimal policy induces compliance.

In Figure 6, we present a summary of the results. Therefore, noncompliance to the optimal policy depends crucially on the convexity of the gravity-based sanctions. As we have shown, when sanctions are linear, the optimal policy always induces compliance.[15] However, if gravity-based sanctions are strictly

[15] This result clarifies a point made in the literature that concentrates only on the subset of policies which induce compliance. For example, Amacher and Malik (1996) assume linear fines and argue that restricting attention to this subset of policies may be overstating enforcement costs, that is, there may be a noncompliant policy that induces the firm to choose the same pollution level with a lower probability of inspection. However, we have shown that when fines

convex, the result depends both on the monitoring costs and the level and convexity of the sanctions. Then, it is more likely that we have noncompliance when non-gravity-based sanctions are small. This result goes in accordance with the literature on crime, which generally considers constant sanctions, that is, independent on the degree of violation.

The interesting case is when $a \in [\underline{a}, \overline{a}]$, that is, when there exists a point of intersection between $p^c(e)$ and $p^n(e)$, as in Figure 4. There, the optimal policy induces noncompliance when monitoring costs are large, that is, when the iso-welfare contours are flat enough. This result is reasonable since, when inspections are very costly, enforcement may not be enough to induce the firm to comply with the regulation.

Finally, and somewhat surprisingly, the optimal policy may induce noncompliance when the gravity-based sanctions are large (that is, when g and/or h are large). As discussed in the Introduction, this result is in general contrary to the one predicted in the literature, which states that an increase in the level of the sanctions increases compliance. However, as we have pointed out already, the main difference between our approach and the one followed in the literature concerns the endogenization of the standard.

5 Conclusions

In this paper, we have studied the characteristics of optimal regulatory policies composed of pollution standards and probabilities of inspecting firms' behavior, where sanctions depend both on gravity-based and non-gravity-based compo-

are linear, the optimal policy always entails compliance. Therefore, there is no overstatement of enforcement costs in that case, since the set of policies that induce compliance is the appropriate set to consider.

nents.

We have demonstrated that the optimal policy can induce either compliance or noncompliance with the standards, depending on the monitoring costs, and the level and convexity of the sanctions. In contrast to previous literature, we have shown that when gravity-based sanctions are strictly convex, noncompliance with the standards is more likely when sanctions are large than when sanctions are small.

Several extensions of our study are possible. For instance, we have assumed that inspections are perfectly accurate. However, the degree of accuracy of a monitoring process generally depends on its cost: the more resources devoted to monitoring, the more accuracy in measurement. This may affect the optimal policy as well as firms' response to the policy, specially under risk aversion, and it may be interesting to endogenize the optimal degree of accuracy of the monitoring policy.

The problem of fully determining the optimal standards, probabilities of inspection and fines has not been solved. This may be, indeed, an interesting problem to look at, considering the full range of policies, since this may help us to study the different types of penalties from a welfare perspective.

Also, we have obtained that, whenever there is noncompliance, the optimal standard is zero. This suggests, under some circumstances, the superiority of probabilistic environmental taxes to pollution standards. A possible explanation for this result is that we have assumed that there are no social costs associated with the imposition and the redistribution of fines. If we assumed, more realistically, that these costs are strictly positive, this would not affect firms' optimal behavior, but it would decrease social welfare in the event sanctions were actually imposed. Therefore, it may be more likely that the optimal policy induces compliance in that case. But, if it continues to induce noncompliance, it may be the case that the optimal standard is positive, since this implies a reduction of the expected fines and, thus, a social cost savings.

All these issues are left for future research.

6 Appendix

Proof of Lemma 1.

To prove the first part, when $P(s,p) = b(s)$, the function is strictly increasing and concave in s, but it does not depend on p. Conversely, when $P(s,p) = \pi(s,p) = \max_{e>s}\left\{b(e) - p\left(a + g(e-s) + h(e-s)^2\right)\right\}$, we have:

$$\pi_s(s,p) = p(g + 2h(n(s,p) - s)) \geq 0 \tag{11}$$

$$\pi_{ss}(s,p) = 2hp(n_s(s,p) - 1) = -2hp\frac{b''(n)}{b''(n) - 2hp} \leq 0 \tag{12}$$

$$\pi_p(s,p) = -\left(a + g(n(s,p) - s) + h(n(s,p) - s)^2\right) < 0 \tag{13}$$

$$\pi_{pp}(s,p) = -(g + 2h(n(s,p) - s))\,n_p(s,p) = -\frac{(g + 2h(n(s,p) - s))^2}{b''(n) - 2hp} > 0 \tag{14}$$

$$\pi_{sp} = g + 2h(n(s,p) - s) + 2hpn_p(s,p) = \frac{b''(n)(g + 2h(n(s,p) - s))}{b''(n) - 2hp} > 0 \tag{15}$$

Summing up both possibilities we obtain the desired result.

To prove the second part, note that, for all $s \geq 0$, we have $\pi(s,0) = \max_{e>0} b(e) = b(\widetilde{e})$. Thus, $P(s,0) = b(\widetilde{e})$. ■

Proof of Lemma 2.

Implicitly differentiating (8), we have:

$$p^{c\prime}(s) = \frac{b'(s) - \pi_s(s,p)}{\pi_p(s,p)} < 0 \tag{16}$$

The denominator is negative (by (13)). The numerator is positive since $\pi_s(s,p) = b'(n)$ (by (5) and (11)), $n(s,p) > s$ and $b(e)$ is strictly concave, which implies that $b'(s) > b'(n)$.

To prove the strict convexity, we differentiate (16) with respect to s:

$$p^{c\prime\prime}(s) = \frac{[b''(s) - \pi_{ss} - \pi_{sp}p^{c\prime}]\,\pi_p - [b'(s) - \pi_s]\,[\pi_{sp} + \pi_{pp}p^{c\prime}]}{(\pi_p)^2} \tag{17}$$

Observe that the sign of (17) is determined by the sign of the numerator. The first expression in brackets of the numerator can be expressed, using (12), (15), and (16), as follows:

$$\begin{aligned} & b''(s) - \pi_{ss} - \pi_{sp}p^{c\prime}(s) = \\ = \; & A[b''(s)\,F\,(b''(n) - 2ph + 2phb''(n)) + b''(n)\,(g + 2h\,(n-s))\,(b'(s) - b'(n))] \end{aligned} \tag{18}$$

where $A = \frac{1}{(b''(n)-2ph)F} < 0$ and $F = a + g\,(n-s) + h\,(n-s)^2 > 0$.

The second expression in brackets of the numerator of (17) can be expressed, using (14), (15) and (16) as follows:

$$\pi_{sp} + \pi_{pp} p^{c\prime}(s) = A\left[b''(n)(g + 2h(n-s))F + (g + 2h(n-s))^2 (b'(s) - b'(n))\right] \tag{19}$$

Now, considering (5), (11), (13), (18), (19) and the assumption $b''' \geq 0$, (17) reduces to:

$$p^{c\prime\prime}(s) \geq -\frac{A}{(\pi_p)^2}\left[b''(n)F + (g + 2h(n-s))(b'(s) - b'(n))\right]^2 \tag{20}$$

The right hand side of (20) is strictly positive, since $A < 0$. Thus, $p^{c\prime\prime}(s) > 0$, as desired.

The second part of the lemma is trivial since $e(\widetilde{e}, p) = \widetilde{e}$, for all $p \in [0, 1]$.

Finally, if $\max_{e>0}\{b(e) - F(e)\} > b(0) = 0$, then $s = 0$ cannot be enforced with any $p \in [0, 1]$. Therefore, by (16) and the fact that $p^c(\widetilde{e}) = 0$, there exists $\widehat{s} \in (0, \widetilde{e})$ such that $p^c(\widehat{s}) = 1$. ■

Proof of Proposition 3.

To show the first part, we solve (10) constrained to the subset of policies that induce compliance:

$$\begin{aligned} \max_{s,p} \quad & \{b(e) - d(e) - cp\} \\ \text{s.t.} \quad & b(s) \geq \pi(s, p) \end{aligned} \tag{21}$$

The optimal response of the firm in this case is $e(s, p) = s$, i.e., to comply with the standard. Thus, an equivalent formulation for (21) is:

$$\max_{s,p} \quad \{b(s) - d(s) - cp\}$$
$$\text{s.t.} \quad p \geq p^c(s) \tag{22}$$

At the optimum, the restriction is binding, since it is enough to set the minimum necessary probability level to enforce a given standard, which is the cheapest one. Thus, problem (22) can be further reduced to:

$$\max_{s} \{b(s) - d(s) - cp^c(s)\} \tag{23}$$

The maximand in (23) is strictly concave, by the assumptions of the model and Lemma 2. Thus, the first order condition in (23) characterizes the interior optimum in the compliance region:

$$b'(s^*) - d'(s^*) = cp^{c\prime}(s^*) \tag{24}$$

By Lemma 2, we know that $p^{c\prime}(s) < 0$. Therefore, we have $b'(s^*) - d'(s^*) < 0$, which implies $s^* > e^w$ and, therefore, $p^* < p^c(e^w)$.[16]

To show the second part, we now solve (10) constrained to the subset of policies that induce noncompliance:

$$\max_{s,p} \quad b(e) - d(e) - cp$$
$$\text{s.t.} \quad e = n(s,p); \; s \geq 0 \tag{25}$$

[16] Since we assume that $p^c(e^w)$ exists, we then have that the solution to (24) is $p^* < 1$. The only possibility for a corner solution is when $d'(\widetilde{e}) + cp^{c\prime}(\widetilde{e}) \geq 0$, which leads to $s^* = \widetilde{e}$ and $p^* = 0$.

Now, the Kuhn-Tucker optimality conditions imply the following:[17]

$$(b'(e) - d'(e))\, n_p(s,p) = c \tag{26}$$

$$(b'(e) - d'(e))\, n_s(s,p) \leq 0 \tag{27}$$

Since $n_p(s,p) < 0$, an interior solution implies that $b'(e) - d'(e) < 0$.

If sanctions are strictly convex in the degree of violation, $n_s(s,p) > 0$ and (27) is strictly negative. Consequently, we obtain $s^* = 0$. And, since $b'(e) - d'(e) < 0$, we have $e(0,p^*) > e^w$ and $p^* < p^n(e^w)$, as desired.

If sanctions are linear, $n_s(s,p) = 0$ and (27) is satisfied with equality. Also, in the linear case we have that $n_{sp}(s,p) = 0$. Therefore, (26) does not depend on s. In consequence, the optimal policy is given by p^* such that (26) holds and any standard that implies noncompliance, that is, $s^* \in [0, \overline{s})$ where $\overline{s}$ is such that $p^* = p^c(\overline{s})$.[18] ■

Proof of Lemma 4.

(i) If $a = 0$, the maximand in (2) is continuous at s. Since it is also strictly concave, it has a unique maximum and the threshold probability between the compliance and noncompliance decisions is given by $p^c(e) = \frac{b'(e)}{g}$. On the other hand, the optimal noncompliance decision when $s = 0$ is given implicitly by $b'(e) = p(g + 2he)$ and, therefore, $p^n(e) = \frac{b'(e)}{g+2he}$. When sanctions are linear, we have $h = 0$ and, consequently, $p^c(e) = p^n(e)$ for all e. However, if sanctions are strictly convex (i.e., when $h > 0$), we have $g < g + 2he$ and $p^c(e) > p^n(e)$, for all $0 < e < \widetilde{e}$.

[17] These conditions are necessary and sufficient for characterizing the optimum, since $n_{pp}(s,p) > 0$ and $n_{ss}(s,p) > 0$, given $b''' \geq 0$. Again, $p < 1$ is guaranteed since we assume that $p^n(e^w)$ exists.

[18] In this case, we would have a corner solution when $d'(\widetilde{e})\, n_p(s,p) + c \geq 0$, and the optimal policy would be such that $s^* \in [0, \widetilde{e})$ and $p^* = 0$.

(ii) If $a > 0$, the maximand in (2) is discontinuous at s and $p^c(e)$ is given implicitly by (8). Given $\{s, p\}$, an increase in a decreases $\pi(s, p)$ while $b(s)$ remains constant. Therefore, $p^c(e)$ decreases when a increases. However, $p^n(e) = \frac{b'(e)}{g+2he}$ does not depend on a. Consequently, if $h = 0$, $a > 0$ implies that $p^c(e) < p^n(e)$ for all $e < \widetilde{e}$.

When sanctions are strictly convex, an increase in a implies that $p^c(e)$ decreases while $p^n(e)$ remains the same. Since at $a = 0$, $p^c(e) > p^n(e)$, there exists $\underline{a} > 0$ such that $p^c(e) > p^n(e)$ for all $a < \underline{a}$. At $a = \underline{a}$, we claim that both $p^c(e)$ and $p^n(e)$ intersect at a point other than $(\widetilde{e}, 0)$, where $p = 1$. Let $(\overline{e}, 1)$ be a point of intersection between $p^c(e)$ and $p^n(e)$, which is characterized by the following conditions:

$$b(\overline{e}) = b(n(\overline{e}, \overline{p})) - pF(n(\overline{e}, \overline{p}) - \overline{e}) \tag{28}$$

$$b'(n(\overline{e}, \overline{p})) = pF'(n(\overline{e}, \overline{p}) - \overline{e}) \tag{29}$$

$$b'(\overline{e}) = pF'(\overline{e}) \tag{30}$$

where (28) implicitly defines $p^c(\overline{e})$, as in (8); (29) characterizes the interior optimal noncompliance decision when the standard is $\overline{e}$, that is, $n(\overline{e}, p)$; and, finally, (30) characterizes the interior optimal noncompliance decision when the standard is 0, i.e., $n(0, p)$.

At $(\overline{e}, \overline{p})$, considering (11), (13), (16) and (29), the slope of $p^c(e)$ is as follows:

$$p^{c\prime}(\overline{e}) = -\frac{b'(\overline{e}) - b'(n(\overline{e}, \overline{p}))}{F(n(\overline{e}, \overline{p}) - \overline{e})} \tag{31}$$

By (28) and (29), $\frac{F(n-\overline{e})}{n-\overline{e}} > F'(n - \overline{e})$. Also, since $n_s < 1$, $n(\overline{e}, \overline{p}) - \overline{e} < \overline{e}$.

Since $F'' \leq \frac{F'}{n-e}$, we then have $\frac{F'(n-\overline{e})}{n-\overline{e}} \geq \frac{F'(\overline{e})}{\overline{e}}$. Considering this and also the fact that $b''' \geq 0$, we then have:

$$p^{c\prime}\left(\overline{e}\right) > \frac{b''\left(\overline{e}\right)}{F'\left(\overline{e}\right)} \frac{\overline{e}}{n\left(\overline{e},\overline{p}\right)-\overline{e}} \tag{32}$$

Now, consider (30), which characterizes $p^n\left(e\right)$ at $\overline{e}$. Differentiating (30), the slope of $p^n\left(e\right)$ at $\overline{e}$ is the following:

$$p^{n\prime}\left(\overline{e}\right) = \frac{b''\left(\overline{e}\right) - \overline{p}F''\left(\overline{e}\right)}{F'\left(\overline{e}\right)} \tag{33}$$

Considering $b''' \geq 0$ and $F''' = 0$, (33) can be reduced to the following:

$$p^{n}\left(\overline{e}\right) < \frac{b''\left(\overline{e}\right)\left(2\overline{e}-n\left(\overline{e},\overline{p}\right)\right) + b''\left(n\left(\overline{e},\overline{p}\right)\right)\left(n\left(\overline{e},\overline{p}\right)-\overline{e}\right)}{\left(2\overline{e}-n\left(\overline{e},\overline{p}\right)\right)F'\left(\overline{e}\right)} \tag{34}$$

To prove that $p^{c\prime}\left(\overline{e}\right) > p^{n\prime}\left(\overline{e}\right)$ and, therefore, that the two functions cross only once, it is sufficient to have:

$$b''\left(\overline{e}\right) + b''\left(n\left(\overline{e},\overline{p}\right)\right)\frac{\left(n\left(\overline{e},\overline{p}\right)-\overline{e}\right)}{\left(2\overline{e}-n\left(\overline{e},\overline{p}\right)\right)} \leq b''\left(\overline{e}\right)\frac{\overline{e}}{n\left(\overline{e},\overline{p}\right)-\overline{e}} \tag{35}$$

by (32) and (34).

Observe that (35) holds since $b''\left(\overline{e}\right)\overline{e} \leq b''\left(n\left(\overline{e},\overline{p}\right)\right)n\left(\overline{e},\overline{p}\right)$ and $\overline{e} > n-\overline{e}$.

Since $p^c\left(e\right)$ decreases when a increases while $p^n\left(e\right)$ remains the same, there exists a range of values, $a \in [\underline{a},\overline{a})$, for which the functions $p^c\left(e\right)$ and $p^n\left(e\right)$ continue to cross once at a point other than $\left(\widetilde{e},0\right)$. Finally, for all $a > \overline{a}$, we have that $p^c\left(e\right) < p^n\left(e\right)$ for all $e < \widetilde{e}$. ■

Proof of Lemma 5.

Given $h > 0$, $a \in [\underline{a}, \overline{a}]$ implies that there exists a point of intersection between $p^c(e)$ and $p^n(e)$ other than $(\widetilde{e}, 0)$, namely $(\overline{e}, \overline{p})$. We also have $n_s(\overline{e}, \overline{p}) = -\frac{2hp}{b''-2hp} > 0$ when $p > 0$. An increase in the level of the sanction means that either g, h or both increase, which means that $n_s(\overline{e}, \overline{p})$ does not decrease. Also, both $n(e, \overline{p})$ and $n(0, \overline{p})$ decrease, by (5). But the decrease of $n(0, \overline{p})$ is at least as great as the decrease of $n(e, \overline{p})$, since $n_s(\overline{e}, \overline{p})$ is nondecreasing in the level of the sanction. Also, $n(\overline{e}, \overline{p}) - \overline{e}$ decreases. As a result, both $p^c(e)$ and $p^n(e)$ decrease, but the variation of $p^n(e)$ at $\overline{e}$ is at least as larger as the variation of $p^c(e)$ at $\overline{e}$. Using Lemma 4, this implies that the new point of intersection entails a smaller standard and a larger probability of inspection, as desired. ■

References

[1] Amacher,G.S. and A.S. Malik (1996), "Bargaining in Environmental Regulation and the Ideal Regulator", Journal of Environmental Economics and Management 30, 233-253.

[2] Arguedas, C. and H. Hamoudi (2004), "Controlling Pollution with Relaxed Regulations", Journal of Regulatory Economics 26, 85-104.

[3] Arguedas, C. (2005), "Bargaining in Environmental Regulation Revisited", *forthcoming*, Journal of Environmental Economics and Management.

[4] Bebchuk, L.A. and L. Kaplow (1991), "Optimal Sanctions when Individuals are Imperfectly Informed about the Probability of Apprehension", Discussion Paper 4/91, Harvard Law School.

[5] Becker, G.S. (1968), "Crime and Punishment: an Economic Approach", Journal of Political Economy 76, 169-217.

[6] Ellis, G.M. (1992a), "Incentive Compatible Environmental Regulations", Natural Resource Modelling 6, 225-255.

[7] Ellis, G.M. (1992b), "Environmental Regulation with Incomplete Information and Imperfect Monitoring", Ph. D Dissertation, University of California, Berkeley.

[8] Grieson, R.E. and N. Singh (1990), "Regulating Externalities Through Testing", Journal of Public Economics 41, 369-387.

[9] Harrington, W. (1988), "Enforcement Leverage When Penalties Are Restricted", Journal of Public Economics 37, 29-53.

[10] Heyes, A.G. (1996), "Cutting Environmental Policies to Protect the Environment", Journal of Public Economics 60, 251-265.

[11] Heyes, A.G. (2000), "Implementing Environmental Regulation: Enforcement and Compliance", Journal of Regulatory Economics 17, 107-129.

[12] Kambhu, J. (1989), "Regulatory Standards, Noncompliance and Enforcement", Journal of Regulatory Economics 1, 103-114.

[13] Livernois, J. and C.J. McKenna (1999), "Truth or Consequences: Enforcing Pollution Standards with Self-Reporting", Journal of Public Economics 71, 415-440.

[14] Malik, A.S. (1990), "Avoidance, Screening and Optimum Enforcement", The RAND Journal of Economics 21, 341-353.

[15] Mookherjee, D. and I. Png (1989), "Optimal Auditing, Insurance and Redistribution", Quarterly Journal of Economics.

[16] Pestieau, P., U.M. Possen and S.M. Slutsky (1997), "Jointly Optimal Taxes and Enforcement Policies in Response to Tax Evasion", mimeo.

[17] Polinsky, A.M. and S. Shavell (1979), "The Optimal Trade-Off Between the Probability and the Magnitude of the Fines", American Economic Review 69, 880-891.

[18] Polinsky, A.M. and S. Shavell (1990), "Enforcement Costs and the Optimal Magnitude and Probability of Fines", Discussion Paper 67, Stanford Law School.

[19] Polinsky, A.M. and S. Shavell (1991), "A Note on Optimal Fines when Wealth Varies Among Individuals", American Economic Review 81, 618-621.

Shavell, S. (1992), "A Note on Marginal Deterrence", International Review of Law and Economics 12, 345-355.

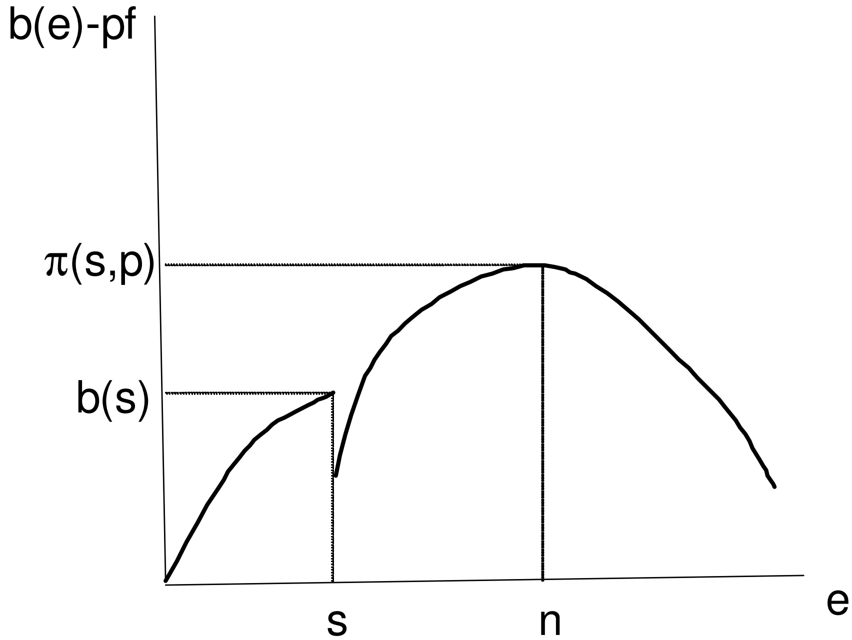

Figure 1: The optimal noncompliance decision exists

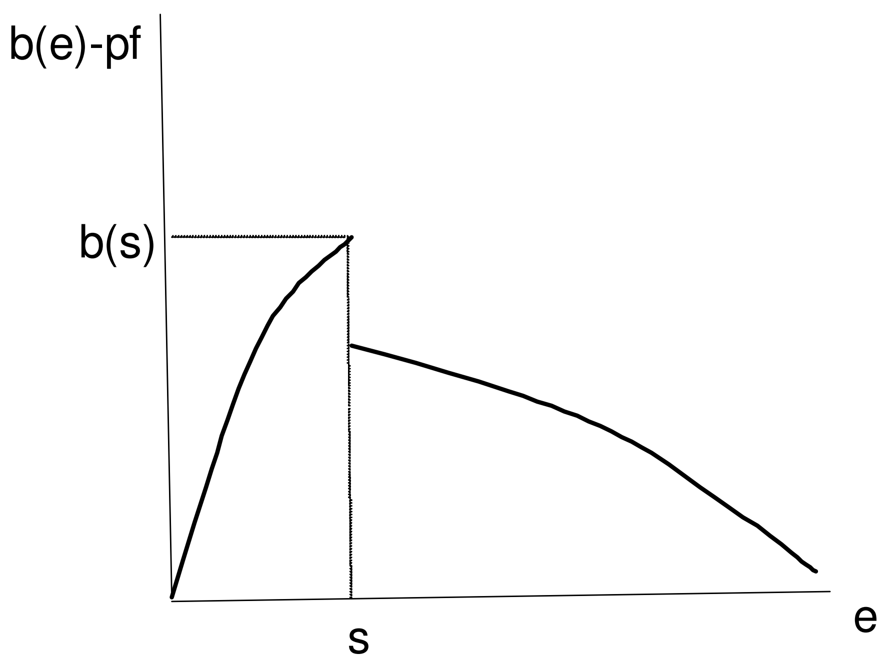

Figure 2: The optimal noncompliance decision does not exist

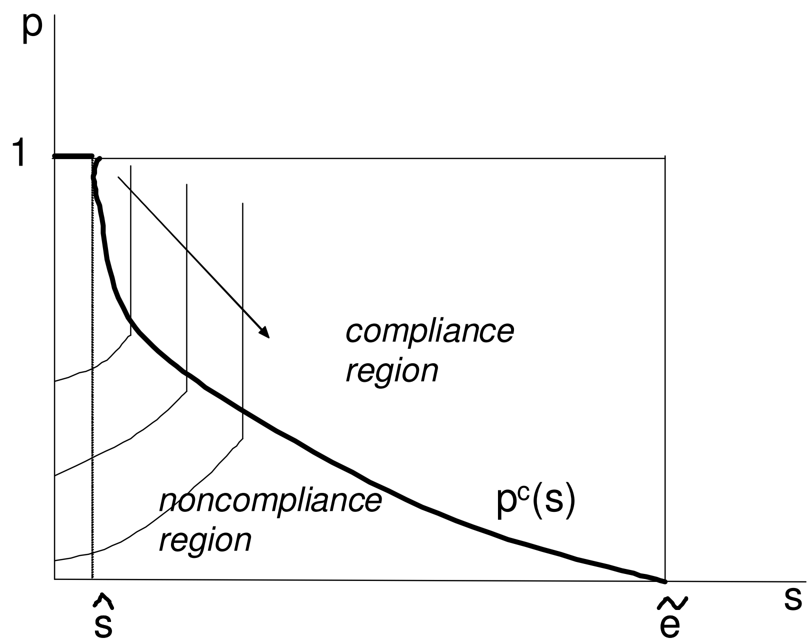

Figure 3: The compliance and noncompliance regions

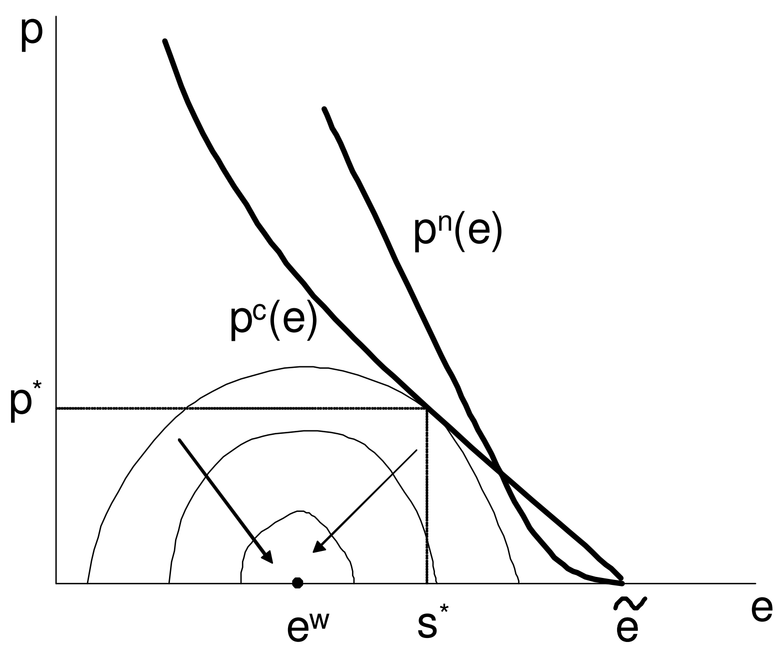

Figure 4: The optimal policy induces compliance

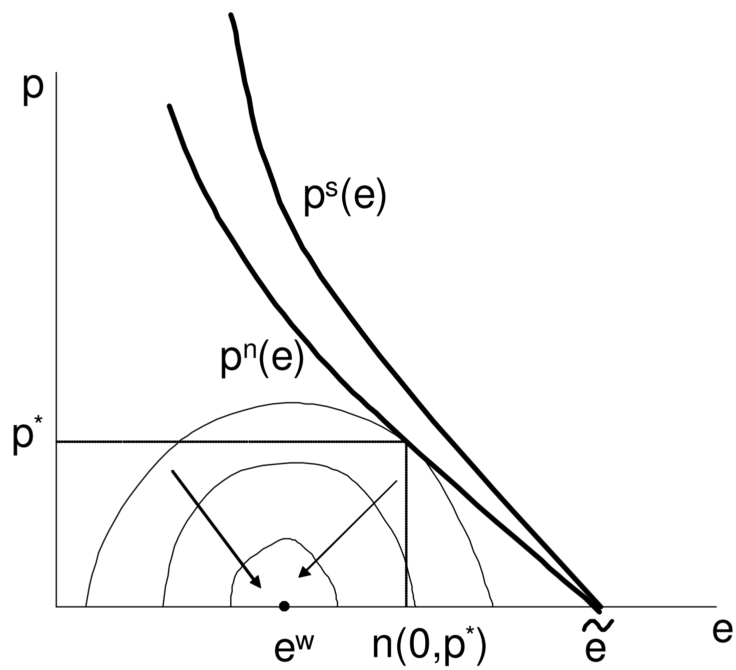

Figure 5: The optimal policy induces noncompliance

non-gravity basd sanction	gravity-based sanction: linear	gravity-based sanction: convex
zero	compliance	noncompliance
positive	compliance	noncompliance under: -“a” small -“c” large -“g” and/or “h” large

Figure 6: The optimal policy inducement dependant on the structure of the fine