Learning about compliance under asymmetric information[*]

Carmen Arguedas [a] and Sandra Rousseau [b,c]

[a] *Departamento de Análisis Económico: Teoría Económica e Historia Económica, Universidad Autónoma de Madrid*

[b] *CEDON, HUBrussel*

[c] *Center of Economic Studies, K.U.Leuven*

Abstract

Over time, inspection agencies gather information about firms that cause harmful externalities. This information may allow agencies to differentiate their monitoring strategies in the future, since inspections can be influenced by firms' past performance relative to other competitors in the market. If a firm is less successful than its peers in reducing the externality, it faces the risk of being targeted for increased inspections in the next period. This risk of stricter monitoring might induce high-cost firms to mimic low-cost firms, while the latter might try to avoid being mimicked. We show that under certain circumstances, the ongoing signaling game between firm types might reduce socially harmful activities.

Keywords: Monitoring and enforcement; externalities; learning; signaling.

JEL codes: D82, H83, K42.

I. INTRODUCTION

Assume that a new regulation for limiting a negative externality arising from production, such as health damages, noise levels, worker safety, biodiversity loss or environmental pollution, is put into place. When the inspection agency responsible for the compliance of the law and the regulated firms meet for the first time, the agency is obviously less informed than the firms about their specific characteristics. However, the agency can collect information on firms' types by performing audits and measuring externality levels. Once the agency has monitored a fraction of firms and has observed their compliance status, it can *learn* something about the type of firms it is dealing with. Then, the agency's monitoring strategy in the following period is determined not only by the firms' past compliance status, but also by their performance *relative* to other competitors in the market. This relative performance acts as a *signal* of the firms' types, since a firm that is less successful than its peers in reducing its harmful effect on society faces the risk of being targeted for increased inspections in the next period.

In this paper, we claim that firms' behavior with respect to the regulation depends not only on their externality control costs and the regulatory stringency, but also on the interaction between firms' relative externality levels and future regulations. We show that firms with higher control costs are tempted to choose sufficiently low levels of the externality (i.e., mimic low-cost firms), in order to face laxer enforcement in the next period. But, at the same time, we find that low-cost firms might be interested in signaling their true type, trying to prevent being mimicked by high-cost firms. As a result, actual external costs might be considerably lower than those expected in a static

regulatory process. In fact, this process may even lead firms to do better than legally required, i.e. they might over-comply with the regulation.[1]

Since imitation is costly for the imitator and the low-cost firms might try to prevent imitation, we find that pooling of firms' decisions only takes place under specific circumstances, which depend on the agency's budget, the number of firms in the industry, the monitoring costs and the minimal amount of the externality needed for the firms to stay in business. In fact, we show that the only possible pooling equilibrium is such that all firms choose the lowest possible level of the externality. When the costs associated with the pooling equilibrium are prohibitively high, a separating equilibrium is attained in which the low-cost firms reduce the externality just enough to avoid imitation by high-cost firms.

Despite its intuitive appeal, the described signaling game that we propose has not yet been used to explain how low enforcement efforts and high compliance rates coexist in reality, a phenomenon known in the literature as the "Harrington paradox", see Nyborg and Telle (2006) for an overview. This paradox challenges the traditional theoretical prediction by Becker (1968) of widespread non-compliance among firms if expected sanctions are small. However, Becker (1968)'s framework is static. If we instead allow for a dynamic interaction between the inspection agency and the firms in the industry, the agency obtains information over time about the regulated firms through signals and can use this information to define its monitoring strategy in the future. Forward looking firms, knowing how their behavior will affect their future regulatory environment, will take this into consideration and may choose lower externality levels than those expected in a static context.

---

[1] According to Arora and Gangopadhyay (1995), the fear of stricter future regulations is indeed one of the reasons that may lead firms to over-comply with existing regulations.

The closest paper to ours is Harrington (1988), who, based on previous results by Landsberger and Meilijson (1982) and Greenberg (1984), investigates the relationship between firms' compliance costs and the average level of compliance when both enforcement budgets and maximum penalties are limited. Harrington (1988) also considers a dynamic setting, where enforcement can be made more efficient by dividing firms into groups contingent on each firms' past performance, and then subject the recent violators to a stricter monitoring and sanctioning policy than the others. There are, however, two crucial differences between our work and Harrington (1998)'s. First, Harrington (1988) does not consider the signaling game we propose. The reason is that Harrington (1988) treats inspected compliant firms and non-inspected firms identically and, therefore, assumes that the agency does not use all the available information efficiently. As a result, firms cannot gain by "mimicking - avoiding being mimicked" because they cannot affect their future regulatory framework. Second, in our model, firms know that their compliance decisions will influence the inspection probabilities in the next period. By contrast, Harrington (1988) does not include this interaction effect and imposes constant inspection probabilities over time.

The main difference of our model versus the traditional models on signaling, such as Spence (1973), Rothschild and Stiglitz (1976) and Grossman (1981), is that in ours several firms emit simultaneous signals that can be used by the regulator to update its information about the firms and set the regulatory strategy in the subsequent period, as opposed to the case of a single firm emitting a single signal.[2] In our model, the combination of multiple signals allows the regulator to revise its information about the regulated firms. Obviously, if all audited firms send the same signal, the regulatory

---

[2] Signaling games have been previously used in a monitoring and enforcement setting by Reinganum and Wilde (1986) and Sobel (1989). These studies model settlement and prosecution decisions and show that plaintiffs with a strong case demand relatively higher payments in order to settle out of court.

agency gets no useful information out of its inspection activities. Also, in our setting, firms cannot opt out on inspections. The regulation on externality control is mandatory and any firm can potentially be monitored by the agency. Thus the firms cannot choose not to participate in the game and they always have to send a signal.

The mimicking behavior in our context is somehow related to avoidance behavior. Firms have numerous options to avoid apprehension and prosecution (Innes, 2005): they can flee the scene, they can lobby politicians to relax enforcement activities, or they can distance themselves from a violation by using legal means. Malik (1990) shows that one implication of incorporating avoidance behavior is that penalties need not always be set as high as possible. An important strand of literature dealing with avoidance activities is the tax evasion literature, see Slemrod and Yitzhaki (2002) for an overview. However, key to the definition of avoidance is the assumption that avoidance is socially detrimental. Interestingly, this is not the case here. If high-cost firms pretend to be low-cost firms, they reduce the level of the negative externality more than they would otherwise do. Under certain circumstances, mimicking, or even the threat of mimicking, might reduce the socially harmful effects of production.

The remainder of the paper is organized as follows. In section 2, we present the model. In section 3, we analyze the benchmark one-period model. In section 4, we deal with the two-period model. In section 5 we study the likelihood of each possible type of equilibrium, depending on the parameters of the model. We conclude in section 6. All the proofs are in the Appendix.

II. THE MODEL

We consider an industry composed of N firms that face a negative production externality (such as discharges of hazardous substances, smog precursors or noise). The externality level of firm i is denoted as $e_i \in \left[\underline{e}, e_i^0 \right]$, where $\underline{e} \geq 0$ is the lowest possible level of the externality associated with implementing the best available technology (or the absolute minimal level required for the firms to continue to do business) and e_i^o is the externality level without regulation in place. Each firm can reduce this externality at a cost, which depends on the externality level e_i, and also on a parameter θ_i, which defines the firms' type. We assume that the externality control cost function of firm i is $c(\theta_i, e_i)$, with $\theta_i \in \{\theta_H, \theta_L\}$, $\theta_H > \theta_L$, and that it has the usual specification: $c_e(\theta_i, e_i) < 0$ for all $e_i < e_i^0$, $c_e(\theta_i, e_i^0) = 0$, $c_{ee}(\theta_i, e_i) > 0$ for all $e_i \leq e_i^0$ and $c_{eee}(\theta_i, e_i)$ sufficiently small. We also assume that $c_{\theta_i}(\theta_i, e_i) > 0$, $c_{\theta_i e}(\theta_i, e_i) < 0$ and $c_{\theta_i ee}(\theta_i, e_i) \leq 0$. Moreover, there are two types of firms in the industry: high-cost firms (θ_H) and low-cost firms (θ_L). Also, we assume that the number of high-cost (low-cost) firms is N_H (N_L), such that $N_H + N_L = N$. [3]

We assume that there is a regulation in place, which imposes a uniform externality limit or a standard $\bar{e} (> \underline{e})$ on the firms. The stringency of the standard and the associated fine, in case a firm is discovered exceeding the standard, are determined by law. For simplicity, the fine F is assumed to be linear: [4]

---

[3] Our assumptions ensure that the single-crossing property holds. This implies that high-cost firms find reducing the externality more costly than low-cost firms.

[4] In practice, a linear specification of fines is often encountered for civil fines, since this structure is easy to understand by firms, citizens and administrations. For example, the EPA's Clean Air Act Stationary

$$F = f \max\{0; e_i - \bar{e}\}, \quad f > 0.$$

Further there exists a regulatory agency whose job it is to minimize total external costs caused by the industry, using the regulation in place. This agency has a budget $B > 0$ *per* period (say, *per* year) to spend on monitoring. We assume that the cost *per* inspection is $m > 0$ and that monitoring is perfectly accurate. We denote by p_{ti} the probability that firm i is inspected in period t, such that $0 \le p_{ti} \le 1$. Also, let e_{ti} be the externality level selected by firm i in period t. For later use, we also define the total costs $C(\theta_i, e_{ti}, p_{ti})$ associated with the regulation for firm i in period t as the sum of the externality control costs and the expected fine:

$$C(\theta_i, e_{ti}, p_{ti}) = c(\theta_i, e_{ti}) + p_{ti} f \max\{0, e_{ti} - \bar{e}\}.$$

For simplicity we assume that first a law announces the standard and the fine, and this announcement is followed by two regulatory periods. In each period, the regulatory agency announces the probability of inspection and then each firm reacts by choosing the level of the externality. Once the agency has observed the behavior of the inspected firms, it can update its information about the firms in the next regulatory period. The chronology of decisions is represented in figure 1.
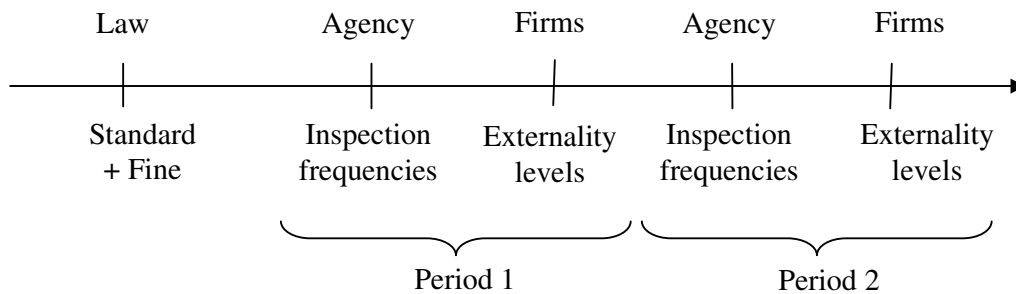

Figure 1: Chronology of decisions

While firms are fully informed about their type (and the types of their partners in the industry), the inspection agency, however, is not[5]. In the first period, the agency is unfamiliar with the regulated firms and it does not have any information on the type of each specific firm. The agency only knows the number of firms of each type in the industry (that is, the distribution of the types). Therefore, initially, the prior probability (or belief) that the agency assigns to the firms to be of the high-cost type is $q_H = \dfrac{N_H}{N}$.

Therefore, the best the agency can do in the first regulatory period is to inspect all the firms with the same frequency, $p_1 = p_{1H} = p_{1L}$ (such that $p_1 mN \leq B$).

We assume that actual externality levels are accurately determined by inspections. Therefore, the agency can update its beliefs, after inspections in the first period have taken place. If e_{1i} is the externality level (signal) of the inspected firm i in period 1, we denote the agency's ex-post belief that an inspected firm i is of the high-cost type as $\mu\left(\theta_i = \theta_H \,\middle|\, e_{1i}, e_{1-i}\right)$. Since there are two types of firms in the industry, the agency can expect at most two different signals e_{1i}, e_{1j}. For the inspected firms, the agency's ex-post beliefs are assumed to be $\mu\left(\theta_i = \theta_H \,\middle|\, e_{1i} = e_{1j}\right) = q_H$, $\mu\left(\theta_i = \theta_H \,\middle|\, e_{1i} < e_{1j}\right) = 0$ and $\mu\left(\theta_i = \theta_H \,\middle|\, e_{1i} > e_{1j}\right) = 1$. That is, if all the inspected firms emit the same signal, the inspection agency cannot update its information (that is, it cannot learn). Therefore, in the second period, the agency will continue to use a uniform inspection frequency, $p_2 = p_{2H} = p_{2L}$, such that $p_2 = p_1$. However, if the inspected firms in period 1 were found to have chosen different externality levels, the agency updates its beliefs: it assigns probability one of being the high-cost type to the firms with the largest

---

[5] Our results are preserved as long as firms have a better (though not necessarily perfect) knowledge about their specific control costs than the agency.

externality level, and it also update its beliefs about the non-inspected firms according to Bayes' rule. Thus, if $n = p_1 N$ is the number of inspected firms in period 1 and n_H (n_L) is the number of firms that happened to be of the high (low)-cost type (such that $n_H + n_L = n$), for the non-inspected firms $k = n+1, n+2, ..., N$, we then have

$$\mu\left(\theta_k = \theta_H \,\middle|\, e_{1i} = e_{1j}\right) = q_H \text{ and } \mu\left(\theta_k = \theta_H \,\middle|\, e_{1i} \neq e_{1j}\right) = \frac{N_H - n_H}{N - n} \text{ , where } i, j = 1, 2, ..., n.$$

If inspected firms perform differently in the first stage, the agency is confronted with three groups of firms in period 2: known high-cost firms, known low-cost firms and firms the agency knows nothing about. The agency can therefore decide to differentiate its inspection strategy and treat each group differently, inspecting with probabilities p_{2H}, p_{2L} and p_{2N}, respectively.

Regarding payoffs, firms choose externality levels that minimize discounted expected costs, composed of control costs and expected fines for non-compliance in each period, where the second period payoff depends on whether the firm was inspected or not in the first period. These discounted expected costs are the following:

$$C\left(\theta_i, e_{1i}, p_1\right) + \delta\left\{p_1 C\left(\theta_i, e_{2i}, p_{2i}\right) + \left(1 - p_1\right) C\left(\theta_i, e_{2i}, p_{2N}\right)\right\}$$

where $0 \leq \delta \leq 1$ is the discount factor.

On the other hand, the agency aims to minimize total external costs in the industry, subject to its financial constraint in each period. The total level of the externality is:

$$E = \left(N_H e_{1H} + N_L e_{1L}\right) + \delta_e\left[n_H e_{2HI} + n_L e_{2LI} + \left(N_H - n_H\right) e_{2HN} + \left(N_L - n_L\right) e_{2LN}\right],$$

where e_{2iI} is the externality level chosen by a firm of type $i = H, L$ in the second period if it was inspected in the first period, and e_{2iN} is the level selected in the second period in case it was not inspected in the first period. The parameter $0 \leq \delta_e \leq 1$ represents the

weight the agency gives to future external costs compared to current externality levels, and can thus be thought of as an externality discount rate.

The relevant equilibrium concept to solve this dynamic game with incomplete information is the perfect Bayesian equilibrium (see Fudenberg and Tirole, 1991a). In this equilibrium, the strategies of the players are optimal for given beliefs and strategies of the other players, and beliefs on the equilibrium path are updated from observed actions according to Bayes' rule, while beliefs off the equilibrium path are updated according to Bayes' rule where possible.

As a reference case, in the next section we consider the one-period model, under both the situations of complete and asymmetric information.

III. ONE-PERIOD REGULATION

After the standard $\overline{e}$ and the fine f are made public knowledge, the agency announces an inspection probability p_i for each firm i, which afterwards responds with an externality level e_i.[6] We solve the problem backwards, and first study the optimal behavior of the firms.

3.1 Firms' behavior

Given $\{\overline{e}, f, p_i\}$, firm i solves the following problem:

$$\min_{e_i} C(\theta_i, e_i, p_i) = \min_{e_i} \left[c(\theta_i, e_i) + p_i f \max\{0, e_i - \overline{e}\} \right].$$

---

[6] Acknowledging a slight abuse of notation, in this section we avoid to use the subscript *t* (which refers to period), since we are just considering one-period regulation.

Lemma 1. *Given* $\{\bar{e}, f, p_i\}$, *firm i's optimal externality level,* e_i^*, *is given by the conditions:*

$$c_e\left(\theta_i, e_i^*\right) + p_i f \geq 0,$$

$$e_i^0 \geq e_i^* \geq \bar{e},$$

$$\left[c_e\left(\theta_i, e_i^*\right) + p_i f\right]\left[e_i^* - \bar{e}\right] = 0.$$

The intuition of this result is straightforward. Given the policy $\{\bar{e}, f, p_i\}$, the firm can decide to either comply with the standard, or not. The optimal strategy is to comply when the marginal expected penalty for non-compliance is larger than the marginal control costs savings of exceeding the standard; that is, when $p_i f \geq -c_e\left(\theta_i, \bar{e}\right)$. In that case, the optimal strategy is $e_i^* = \bar{e}$.[7] However, the optimal strategy is to exceed the standard if the marginal expected penalty is below the marginal control cost savings at the standard. In that case, the firm will choose the externality level such that marginal control cost savings and marginal expected fines are equal. Therefore, $e_i^* > \bar{e}$ and $c_e\left(\theta_i, e_i^*\right) + p_i f = 0$. Note that $e_i^* < e_i^o$ as long as $p_i > 0$.

From lemma 1 we can immediately see that there exists a threshold inspection probability for each type, such that compliance is ensured above that threshold. That minimum probability required is:

$$\bar{p}_i = -\frac{c_e\left(\theta_i, \bar{e}\right)}{f}$$

Obviously, $\bar{p}_H > \bar{p}_L$, since $\theta_H > \theta_L$ and $c_{\theta_i e}\left(\theta_i, e_i\right) < 0$.

---

[7] In the one-period model, the firm never chooses an externality level strictly below the standard: it just increases control costs, but there are no penalty savings. However, this does not longer hold in a multi-period regulation, as we will see later on.

Thus, $p_i > \bar{p}_i$ ensures that type i complies with the standard, i.e., $e_i^* = \bar{e}$. However, the contrary ensures that type i exceeds the standard, i.e., $e_i^* > \bar{e}$, where e_i^* is such that $c_e\left(\theta_i, e_i^*\right) + p_i f = 0.$ This expression defines an implicit relationship between the inspection probability and the induced externality level. Using the implicit function theorem, we have:

$$\frac{\partial e_i}{\partial p_i}\left(\theta_i, e_i\right) = -\frac{f}{c_{ee}\left(\theta_i, e_i\right)} < 0, \qquad (1)$$

which defines the effect on the externality of a marginal increase in the inspection probability; the larger the probability, the lower the externality level. Note that the denominator is 'almost' constant in our model, thus implying that the impact of the inspection frequency on the externality level is 'almost' constant. From (1) and our assumptions $c_{\theta_i ee}\left(\theta_i, e_i\right) \leq 0$ and $c_{eee}\left(\theta_i, e_i\right)$ sufficiently small, we have that the high-cost firms react more than the low-cost firms to an increase in the probability of inspection (in terms of the externality reduction), independently of the externality level. That is[8]:

$$\frac{\partial e_i}{\partial p_i}\left(\theta_H, e_H\right) < \frac{\partial e_i}{\partial p_i}\left(\theta_L, e_L\right). \qquad (2)$$

For later reference, we define the minimum cost function $\mathbb{C}\left(\theta_i, p_i\right) \equiv \min_{e_i} C\left(\theta_i, e_i, p_i\right).$ By the envelope theorem, this function is increasing in the probability of inspection. That is, $\dfrac{\partial \mathbb{C}\left(\theta_i, p_i\right)}{\partial p_i} = f \max\left\{0, e_i^* - \bar{e}\right\} \geq 0.$

---

[8] Alternatively, the relative impact of low-cost firms on the externality is higher than that of high-cost firms if $c_{\theta_i ee}\left(\theta_i, e_i\right) > 0$. All our results are then reversed: the agency targets known low-cost firms more than known high-cost firms, low-cost firms have incentives to mimic high-cost firms and high-cost firms try to deter imitation. Successful imitation and successful active deterrence of imitation both lead to higher externality levels.

3.2 Agency's behavior

The behavior of the agency depends on the information available about the firms. With perfect information, the inspection agency knows the specific type of each firm, and is thus able to differentiate its monitoring strategy depending on the type. Since the agency takes into account the optimal response of the firm, presented in Lemma 1 above, the agency's optimization problem under complete information is the following:

$$\min_{p_H, p_L} \quad N_H e_H + N_L e_L$$
$$s.t. \quad c_e(\theta_i, e_i) + p_i f \ge 0, \quad i = H, L$$
$$e_i \ge \bar{e}, \quad\quad\quad\quad i = H, L \quad\quad\quad (3)$$
$$e_i^0 \ge e_i, \quad\quad\quad\quad i = H, L$$
$$m[N_H p_H + N_L p_L] \le B$$

The first three constraints represent the firms' best responses, as established in Lemma 1. The last condition is the agency's budgetary constraint.

In the following lemma, we present the optimal policy in this case.

Lemma 2. *In the one-period game under perfect information, the inspection agency's optimal policy* $\left(p_H^*, p_L^* \right)$ *is the following:*

(i) If $m[N_H \bar{p}_H + N_L \bar{p}_L] \le B$, *then* $p_H^* \ge \bar{p}_H$, $p_L^* \ge \bar{p}_L$ *such that* $m\left[N_H p_H^* + N_L p_L^* \right] \le B$.

(ii) If $mN_H \bar{p}_H \le B < m[N_H \bar{p}_H + N_L \bar{p}_L]$, *then* $p_H^* = \bar{p}_H$, $p_L^* = \dfrac{B - mN_H \bar{p}_H}{mN_L} < \bar{p}_L$.

(iii) If $B < mN_H \bar{p}_H$, *then* $p_H^* = \dfrac{B}{mN_H} < \bar{p}_H$ *and* $p_L^* = 0$.

The results of Lemma 2 are very intuitive. Case (i) refers to the situation where the available budget of the inspection agency is sufficient to deter all violations in the

13

industry. Therefore, all firms comply with the regulation: $e_H^* = e_L^* = \bar{e}$. This case is, however, trivial, as well as unrealistic, so in the remainder of the paper we assume that

$$B < m\left[N_H \bar{p}_H + N_L \bar{p}_L\right].$$

In case (ii), the high-cost firms comply with the standard, while the low-cost firms exceed it: $e_H^* = \bar{e}$ and $e_L^* > \bar{e}$. By (2), the marginal impact of the inspection probability on the externality levels is larger (in absolute terms) for the high-cost firms than for the low-cost firms. Therefore, it is more efficient (in terms of the global externality reduction) to spend the budget on monitoring the high-cost firms until they are induced to comply, and then use the remaining budget (if any) to monitor the low-cost firms.

Finally, case (iii) is the situation of full non-compliance, since the agency does not have enough means to induce the high-cost firms to comply. Here, the budget is fully allocated to monitor the high-cost firms. As a result, the induced externality levels in this case are $e_H^* \geq \bar{e}$ and $e_L^* = e_L^0$.

Summarizing, the inspection agency devotes an extra unit of monitoring resources to the firm type that is more efficient in reducing the externality. Since by our assumptions, this is type θ_H, independently on the externality levels, the agency first uses its resources to deter the known high-cost firms and then uses any remaining resources to deter the known low-cost firms.

Now assume, by contrast, that the agency does not know the specific type of the firms it is dealing with. Here, the agency only has one possibility: it randomly inspects as many firms as possible within the budget restriction, since there is no rational basis for using a separating strategy. This is trivially stated in the following lemma.

Lemma 3. *In the one-period game under imperfect information, the inspection agency's optimal policy is* $p = \dfrac{B}{mN}$.

Combining Lemmas 2 and 3, it is easy to see that the uniform inspection frequency is such that $p_L^* \leq p \leq p_H^*$. This uniform inspection strategy is clearly inefficient, since, compared to the solution under complete information, the low-cost firms will be inspected too often and the high-cost firms will be insufficiently monitored. Nevertheless, the agency is not able to improve upon it, since it does not have the necessary information. [9]

IV. TWO-PERIOD REGULATION

We now turn to the two-period case illustrated earlier in figure 1. We first look at the behavior of the agency in the second period, once she has updated her information

---

[9] The assumption that expression (1) is "almost" a constant is crucial in obtaining corner solutions (in Lemma 2 above, but also later in Proposition 1). However, we prefer to keep this assumption for expositional purposes, since the spirit of the results does not change under a more general specification of abatement costs, such that the marginal impact of the inspection probability on the externality level varies with the inspection probability. In fact, it is quite intuitive to see what happens under a more realistic case where $\lim_{p \to 0} |\partial e_i / \partial p_i| = \infty$ and $\partial^2 e_i / \partial p_i^2 \geq 0$, which is equivalent to $\lim_{e_i \to \infty} c_{ee}(\theta_i, e_i) = 0$ and $c_{eee}(\theta_i, e_i) < 0$. Here, the lower the inspection probability, the larger (in absolute terms) the marginal impact on the externality level of an increase the inspection probability. In this case, an interior solution would be possible, since $c_{ee}(\theta_i, \overline{e}) > \lim_{e_j \to \infty} c_{ee}(\theta_j, e_j) = 0$. Optimality would then imply $\partial e_i / \partial p_i = \partial e_j / \partial p_j$: an extra euro spent on monitoring should be used in such a way that the marginal effects on the firms' externality levels are equal. In other words, it might be convenient to spend some monitoring effort on the low-cost firms to prevent excessive pollution, even if the high-cost firms cannot be induced to comply. In any case, it is easy to check that the ranking of inspection probabilities $p_L^* \leq p \leq p_H^*$ remains valid. The spirit of the remaining results of the paper is also the same under this alternative specification of abatement costs.

based on the inspections carried on in the first period. Then, in a later stage, we analyze

the firms' strategies in the first period and the inspection agency's beliefs that sustain

those strategies as part of the perfect Bayesian equilibrium.

4.1 Second period

In the second period the inspection agency chooses the inspection probabilities, and

then firms respond as in Lemma 1. Therefore, we study the agency's strategy, taking the

firms' responses into account. On the one hand, if all the inspected firms in the previous

period choose the same amount of the externality (pooling equilibrium), they all face the

same inspection probability in the second period, $p = \dfrac{B}{mN}$, as established in Lemma 3.

On the other hand, if the previous inspections allow the agency to divide the audited

firms into high-cost and low-cost firms (separating equilibrium), then it can differentiate

its optimal inspection strategy. In this case, the problem faced by the inspection agency

in the second period is:

$$
\min_{p_{2H},p_{2L},p_{2N}} \quad n_H e_{2HI} + n_L e_{2LI} + \left(N_H - n_H\right)e_{2HN} + \left(N_L - n_L\right)e_{2LN}
$$

$$
s.t. \quad m\left[n_H p_{2H} + n_L p_{2L} + \left[N - n_H - n_L\right]p_{2N}\right] \le B
$$

$$
c_e\left(\theta_i, e_{2iI}\right) + p_{2i} f \ge 0, \qquad i = H, L
$$

$$
c_e\left(\theta_i, e_{2iN}\right) + p_{2N} f \ge 0, \qquad i = H, L
$$

$$
e_{2iI} \ge \overline{e}, \quad e_{2iN} \ge \overline{e}, \qquad i = H, L,
$$

where e_{2iI} is the externality level chosen by a firm of type i that was inspected in the

first period, and e_{2iN} is the level selected in case a firm of type i was not inspected.

Therefore, the agency now has perfect information about two types of firms (the high-

cost and low-costs inspected in the first period), and no information about the exact type

of the remaining firms, even though it knows the proportion of each firm type in the

group of non-inspected firms.

The agency's optimal strategy in this case combines features of the one-period regulation optimal policy under complete and incomplete information analyzed in the previous section.

From (2), we already know that monitoring resources must be first devoted to known high-cost firms rather than to known low-cost firms, since the marginal impact of the inspection probability on the externality level of the former group is larger (in absolute terms). Since the group of non-inspected firms is formed by high-cost and low-cost firms (in proportions $\frac{N_H - n_H}{N - n}$ and $\frac{N_L - n_L}{N - n}$, respectively), the expected marginal impact of the probability of inspection on the externality level of the non-inspected firms is between the marginal impacts of the high-cost and low-cost firms. That is,

$$\frac{\partial e_{HI}}{\partial p_{2H}} \leq \frac{N_H - n_H}{N - n}\frac{\partial e_{HN}}{\partial p_{2N}} + \frac{N_L - n_L}{N - n}\frac{\partial e_{LN}}{\partial p_{2N}} \leq \frac{\partial e_{LI}}{\partial p_{2L}} \tag{4}$$

Here, the optimal inspection strategy follows the same structure as that of lemma 2. Therefore, the agency first devotes monitoring resources to the firms that are most effective in reducing external costs (i.e. the firms that were inspected in the first period and that turned out to be type θ_H); if the agency has enough resources to induce compliance in this group, it devotes the money left to the next most effective group (i.e. the non-inspected firms in the first period); and then, if there is still money left, to the least effective group in reducing the externality problem (i.e. the inspected firms that turned out to be of type θ_L). The optimal inspection strategy is formally presented next.

Proposition 1. *The inspection agency's optimal policy* $\left(p_{2H}^*, p_{2L}^*, p_{2N}^* \right)$ *in the second period satisfies the following conditions:*

(i)   If $m(N - n_L)\bar{p}_H \leq B$, *then* $p_{2H}^* = p_{2N}^* = \bar{p}_H$ *and* $p_{2L}^* = \dfrac{B - m(N - n_L)\bar{p}_H}{mn_L} < \bar{p}_L.$

(ii) If $mn_H \overline{p}_H \leq B < m(N - n_L)\overline{p}_H$, then $p^*_{2H} = \overline{p}_H$, $p^*_{2N} = \dfrac{B - mn_H \overline{p}_H}{m(N-n)} < \overline{p}_H$ and

$p^*_{2L} = 0$.

(iii) If $B < mn_H \overline{p}_H$, then $p^*_{2H} = \dfrac{B}{mn_H} < \overline{p}_H$ and $p^*_{2N} = p^*_{2L} = 0$.

Note that cases (i) and (ii) represent situations in which there are not enough monitoring resources to induce compliance of the known low-cost firms (in case i), or the known low-cost firms and non-inspected firms (in case ii). In these cases, the induced pollution levels are, respectively, ($e^*_{2H} = e^*_{2N} = \overline{e}$, $e^0_L \geq e^*_{2L} > \overline{e}$) and ($e^*_{2H} = \overline{e}$, $e^0_H \geq e^*_{2N} > \overline{e}$ and $e^*_{2L} = e^0_L$). Case (iii) represents the case where no group can be induced to comply, such that $e^0_H \geq e^*_{2H} > \overline{e}$, $e^*_{2N} \in \{e^0_H, e^0_L\}$ and $e^*_{2L} = e^0_L$.

For later purposes, it is now interesting to see which is the ranking of inspection probabilities compared to the uniform inspection probability $p = \dfrac{B}{mN}$. From Proposition 1, we have $p^*_{2L} \leq p^*_{2N} \leq p^*_{2H}$. It is also clear that $p^*_{2L} \leq p \leq p^*_{2H}$, which means that the uniform inspection strategy monitors the known low-cost firms too often, while the known high-cost firms are insufficiently inspected. However, whether $p^*_{2N} \leq p$ or $p^*_{2N} \geq p$ crucially depends on the available monitoring budget, as we show next.

Proposition 2. *The ranking of inspection probabilities* $\left(p^*_{2H}, p^*_{2L}, p^*_{2N}\right)$ *and* $p = \dfrac{B}{mN}$ *is the following:*

(i)   If $m[N - n_L]\overline{p}_H \leq B$, then $p^*_{2H} = p^*_{2N} = \overline{p}_H \geq p > p^*_{2L} \geq 0$.

18

(ii)   If $m\dfrac{n_H N}{n}\overline{p}_H \le B < m[N-n_L]\overline{p}_H$, then $p_{2H}^* = \overline{p}_H > p_{2N}^* \ge p > p_{2L}^* = 0.$

(iii)   If $mn_H\overline{p}_H \le B < m\dfrac{n_H N}{n}\overline{p}_H$, then $p_{2H}^* = \overline{p}_H > p \ge p_{2N}^* > p_{2L}^* = 0.$

(iv)   If $B < mn_H\overline{p}_H$, then $\overline{p}_H > p_{2H}^* > p > p_{2N}^* = p_{2L}^* = 0.$

In case (i), the available funds are sufficient to enforce $e_{2H} = \overline{e}$ and $e_{2N} = \overline{e}$ (in correspondence with case (i) of Proposition 1). Thus, the inspection agency can successfully induce all firms, except the low-cost firms that were inspected in the previous period, to comply with the regulation (but see footnote 9). The rest of the monitoring resources are used to inspect the known low-cost firms and, therefore, the differentiated inspection frequencies are trivially ranked $p_{2H}^* = p_{2N}^* = \overline{p}_H \ge p > p_{2L}^* \ge 0.$ In cases (ii) and (iii), both linked to case (ii) of Proposition 1, the available funds are sufficient to enforce $e_{2H} = \overline{e}$ but not to enforce $e_{2N} = \overline{e}$, which then implies $p_{2L}^* = 0$ and $e_{2L}^* = e_L^0.$ Since $p = \dfrac{B}{mN}$ and $p_{2N}^* = \dfrac{B - mn_H\overline{p}_H}{m(N-n)}$, both the numerator and the denominator of p_{2N}^* are lower than those of p (there are fewer available funds to deter unknown firms in the second period, but also a smaller number of unknown firms). Therefore, $p_{2N}^* \ge p$ if, and only if, the percentage of decrease in the available funds to deter unknown firms is lower than the percentage of decrease in the number of unknown firms. That is, if and only if $\dfrac{n_H\, m\, \overline{p}_H}{B} \le \dfrac{n}{N}$, which holds under case (ii). The opposite holds under case (iii), and therefore, $p_{2N}^* \le p$ in that scenario. Finally, in case (iv), in correspondence with case (iii) of Proposition 1, there are not enough resources to induce the known high-cost firms to comply, and therefore, we trivially have $\overline{p}_H > p_{2H}^* > p > p_{2N}^* = p_{2L}^* = 0.$

4.2. First period

As we know from lemma 3, in the first period the agency has no information and cannot

do better than to randomly inspect firms: $p = \dfrac{B}{mN}$.

In this period, the relevant issue is to analyze the behavior of the firms, who realize that

their actions have an effect on next period's monitoring. Knowing that inspection

probabilities in period 2 will be as in Proposition 1, we can distinguish two types of

equilibria: 1) separating equilibria in which both types of firms choose a different

externality level and 2) pooling equilibria in which both types of firms select an

identical level of the externality.[10]

Firstly, in a separating equilibrium, the high-cost firms select an externality level so as

to minimize its expected costs: $e_{1H} = e_{1H}^*$. The low-cost firms equilibrium strategy is to

select an externality level e_{1L} no larger than the one that minimizes its expected costs,

$e_{1L} \leq e_{1L}^*$. In order for $\left(e_{1H} = e_{1H}^*, e_{1L} \right)$ to be part of a separating equilibrium, it must be

the case that the high-cost firms do not prefer e_{1L} to e_H^* (since, obviously, low-cost

firms do not prefer e_{1H}^* to e_{1L}):

$$C\left(\theta_H, e_{1L}, p\right) + \delta \mathbb{C}\left(\theta_H, p\right) \geq \mathbb{C}\left(\theta_H, p\right) + \delta \left\{ p\mathbb{C}\left(\theta_H, p_{2H}\right) + \left(1 - p\right)\mathbb{C}\left(\theta_H, p_{2N}\right) \right\} \qquad (5)$$

where the left hand side of (5) are the discounted expected costs of choosing e_{1L} in the

first period (and then, face a uniform inspection probability in the second period), and

---

[10] Note that we assume that all players expect an opponent to play according to the equilibrium strategies
even if that opponent deviates from the equilibrium path (Fudenberg and Tirole, 1991b).

the right hand side of (5) represent the discounted expected costs of choosing e_H^* , and then face a differentiated monitoring strategy in the second period.

Expression (5) holds whenever $e_{1L} \leq \tilde{e}_H$, where $\tilde{e}_H \geq \underline{e}$ is the level for which (5) holds with equality. So, in order to have the possibility of a separating equilibrium, the externality level chosen by the low-cost type must be sufficiently low so as to make it unattractive for the high-cost firms to deviate. Therefore, the low-cost firms choose $e_{1L} = \min\{\tilde{e}_H, e_{1L}^*\}$. This is indeed the *Riley* outcome (Riley, 1979).[11] Since $e_{1L} < e_{1H}^*$, the agency learns that the firm that chooses e_{1H}^* is a high-cost type and that the firm that selects e_{1L} is a low-cost type. Based on this knowledge, the agency can also update its beliefs about the subgroup of non-inspected firms according to Bayes' rule. Therefore, in the second period, the agency inspects the known high-cost firms, the known low-cost firms and the subgroup of non-inspected firms with probabilities $\left(p_{2H}^*, p_{2L}^*, p_{2N}^*\right)$, respectively, as defined in Proposition 1, and firms react accordingly, as in Lemma 1.

Secondly, if low-cost firms cannot select a sufficiently low level of the externality such that high-cost firms are not deterred from mimicking, the equilibrium is pooling. This happens when $\underline{e} > \tilde{e}_H$, where again, $\tilde{e}_H$ is the threshold level in equation (5), and therefore, $e_{1L} = e_{1H} = \underline{e}$. Low-cost firms would have to decrease their externality levels below $\underline{e}$ to avoid mimicking, but this is not possible since $\underline{e}$ is the absolute minimum

---

[11] In order to obtain the *Riley* outcome as a unique equilibrium in a model of two types, we have used the 'intuitive criterion' of Cho and Kreps (1987), which imposes certain conditions on the agency's out-of-equilibrium beliefs, in the sense that certain types of firms should not be expected to use certain strategies. This means, for example, that for any pair of externality levels chosen by two particular firms, the agency assigns probability zero of being a high-cost type to the firm that selected the lowest externality level. Of course, other strategies can be sustained in equilibrium with other beliefs, but those beliefs might not be reasonable.

amount of the externality required for the firms to continue doing business. Then, it is always profitable for the high-cost firm to imitate the low-cost firms and the low-cost firms will not be able to prevent this. Thus, as Kreps and Sobel (1994) show, pooling is only possible at the lowest possible level of the externality, that is, the amount of the externality associated with implementing the best available technology. Any other pooling strategy is not stable (any firm might be tempted to infinitesimally reduce its own level to pretend to be of the low-cost type). Therefore, the agency's inspections in the first period will not provide the necessary information for differentiating its inspection strategy in the second period.

All this reasoning lead us to a unique perfect Bayesian equilibrium, that can be either separating or pooling crucially depending on $\underline{e} \leq \tilde{e}_H$ or $\underline{e} > \tilde{e}_H$. We summarize the result in the following:

Proposition 3. *The following strategies constitute the unique perfect Bayesian equilibrium:*

(i) If $\underline{e} \leq \tilde{e}_H$, then (i.1) in the first period, the regulator inspects all firms with probability $p = \dfrac{B}{mN}$; (i.2) firms choose $e_{1H} = e_{1H}^*$ and $e_{1L} = \min\{\tilde{e}_H, e_{1L}^*\}$; (i.3) the regulator's ex-post beliefs are $\mu\left(\theta_i = \theta_H \middle| e_{1H}\right) = 1$ and $\mu\left(\theta_i = \theta_H \middle| e_{1L}\right) = 0$ for inspected firms, and $\mu\left(\theta_k = \theta_H\right) = \dfrac{N_H - n_H}{N - n}$ for non-inspected firms; (i.4) in the second period, the regulator inspect firms with probabilities $\left(p_{2H}^*, p_{2L}^*, p_{2N}^*\right)$; and (i.5) firms choose $e_{2HI}^*, e_{2LI}^*, e_{2HN}^*, e_{2LN}^*$.

(ii) If $\underline{e} > \tilde{e}_H$, then (ii.1) in the first period, the regulator inspects all firms with probability $p = \dfrac{B}{mN}$; (ii.2) firms choose $e_{1H} = e_{1L} = \underline{e}$; (ii.3) the regulator's ex-post beliefs are $\mu\left(\theta_i = \theta_H \,\middle|\, e_{1i}\right) = q_H$ for all firms; (ii.4) in the second period, the regulator inspects all firms with probability $p = \dfrac{B}{mN}$; and (ii.5) firms choose $e^*_{2L} = e^*_{1L}$ and $e^*_{2H} = e^*_{1H}$.

For the reader's convenience, table 1 summarizes the sequence of choices in equilibrium, depending on the thresholds $\underline{e}$ and $\tilde{e}_H$.

	Period 1		Period 2	
	Agency	Firms	Agency	Firms
Case (ia) $\underline{e} < \tilde{e}_H < e^*_{1L}$	Uniform inspections: p	Separating: $e_{1H} = e^*_{1H} > e_{1L} = \tilde{e}_H$	Differentiated inspections: p_{2H}, p_{2L}, p_{2N}	Separating: $e^*_{2HI}, e^*_{2LI}, e^*_{2HN}, e^*_{2LN}$
Case (ib) $\underline{e} < e^*_{1L} \leq \tilde{e}_H$	Uniform inspections: p	Separating: $e_{1H} = e^*_{1H} > e_{1L} = e^*_{1L}$	Differentiated inspections: p_{2H}, p_{2L}, p_{2N}	Separating: $e^*_{2HI}, e^*_{2LI}, e^*_{2HN}, e^*_{2LN}$
Case (ii) $\underline{e} > \tilde{e}_H$	Uniform inspections: p	Pooling: $e_{1L} = e_{1H} = \underline{e}$	Uniform inspections: p	Separating: $e^*_{2H} = e^*_{1H} > e^*_{2L} = e^*_{1L}$

Table 1: Summary of firms' and agency's decisions

V. DISCUSSION

The purpose of this section is twofold. First, we compute the total externality levels associated with each type of equilibrium, and select the *best* equilibrium, in terms of the lowest total externality level induced. Second, we study the likelihood of obtaining each of the two equilibria depending on the parameters of the model. In particular, we analyze the influence of the agency's budget, the number of firms and the monitoring costs on the possibility that firms choose the same or different strategies and, therefore, whether the agency can learn the firms' types in the first period and differentiate its strategy in the second regulatory period.

5.1. Impact on total externality levels

The strategies chosen by the firms and the agency will influence the total externality levels, resulting from monitoring and enforcing the standard $\bar{e}$. Depending on the case (see table 1), total external costs will differ. As a reference point, we use the case where the agency cannot learn anything from inspecting firms, even if externality levels differ between firms. So, the reference scenario is analogous to playing the static game with imperfect information (see section III) twice. This implies that total discounted externality levels over the two periods equal:

$$(1+\delta_e)\left(N_H e_{1H}^* + N_L e_{1L}^*\right)$$

From proposition 3 and table 1, we calculate the resulting externality levels for each possible case in our signaling game. This gives:

For case (ia): $\quad E = \left[N_H e_{1H}^* + N_L \tilde{e}_H\right] + \delta_e E_2$

For case (ib): $\quad E = \left[N_H e_{1H}^* + N_L e_{1L}^*\right] + \delta_e E_2$

For case (ii): $E = N\underline{e} + \delta_e \left(N_H e_{1H}^* + N_L e_{1L}^* \right)$

with $E_2 = n_H e_{2HI} + n_L e_{2LI} + \left(N_H - n_H \right) e_{2HN} + \left(N_L - n_L \right) e_{2LN}$, which is the total externality

in period 2 for a differentiated inspection policy. Since the agency can always choose a

uniform inspection strategy in the second period if that would be better, the level E_2

never exceeds $N_H e_{1H}^* + N_L e_{1L}^*$.

Trivially, this shows that learning is always beneficial. In any case, externality levels in

a situation where the agency can acquire information through auditing firms will be

lower than in the reference scenario, where such learning was non-implementable.

Comparing cases (ia) and (ib), we also find that the threat of mimicking is a good thing,

since the low-cost firms will reduce their externality levels more than in the static case -

they might even over-comply with the standard - in order to deter high-costs firms from

imitating them. So, when firms choose a separating equilibrium in the first period,

resulting total external costs over the two periods are lower if mimicking is actively

prevented.

Comparing case (ii) with cases (ia) and (ib), we find that successful imitation (that is,

the firms are pooled) can reduce total discounted externality levels when the discount

rate δ_e is sufficiently low and, for case (ia), also if the number of high-cost firms N_H is

sufficiently high. In fact, externality costs in the first period can be very low, since firms

select externality levels below the standard, i.e., they over-comply. By contrast, for a

sufficiently high discount rate, pooling and mimicking by firms might worsen the

externality problem, since the second period external costs associated with a uniform

inspection probability will exceed those resulting from a differentiated inspection policy

$\left(E_2 \right)$.

5.2. Likelihood of each equilibrium

The likelihood of obtaining a particular equilibrium crucially depends on the complete

ranking of inspection probabilities, presented in Proposition 2, which is critically

affected by the available budget B. Clearly, the larger p_{2H}^* and p_{2N}^* (compared to p),

the larger the incentives of the high-cost firms to mimic the low-cost firms (in other

words, the lower $\tilde{e}_H$, the threshold externality level of the high-cost firms, defined in

the previous section). Conversely, the lower p_{2L}^* and p_{2N}^* (compared to p) the larger the

incentives of the low-cost firms to prevent imitation by the high-cost firms.

We now explore the four cases of Proposition 2, to see under which possibility we have

the largest incentives for mimicking (by the high-cost firms) and the largest incentives

to deter mimicking (by the low-cost firms). Clearly, on the one hand, imitation is most

attractive to the high-cost firms in case (i) of Proposition 2, since in that case

$p_{2H}^* = p_{2N}^* = \bar{p}_H > p$. We also see that these incentives for mimicking are successively

lower under cases (ii), (iii) and (iv) of Proposition 2. On the other hand, case (iv)

provides the low-cost firms with the largest incentives to avoid mimicking, since

$p_{2L}^* = p_{2N}^* = 0$, and cases (iii), (ii) and (i) provide successively lower incentives to deter

imitation.

Since case (i) of Proposition 2 provides the largest likelihood of mimicking and the

lowest likelihood of mimicking deterrence, this case then provides the largest likelihood

of pooling behavior among firms in the first period. So, interestingly, pooling behavior

is more likely under a large inspection budget.

We are now able to comment on the influence of the available budget, the number of

firms in the industry and the level of the monitoring costs on the motivation for

mimicking. As we have already shown, an increase in the available budget, *ceteris paribus*, provides more incentives for imitation. Indeed, a higher budget implies more inspections in any scenario. Thus the benefits from mimicking increase, since they depend, amongst other things, on the likelihood *p* of being inspected in the first period. The more likely it is that a high-cost firm can be detected in the first period, the more it stands to gain from hiding among the low-cost firms. For this reason, we also make the following two observations. Firstly, the lower the number of firms affected by the regulation, the higher the incentives for pooling, *ceteris paribus.* Secondly, if inspections become less expensive, *ceteris paribus,* high-cost firms will be induced mimic low-cost firms.

VI. CONCLUSIONS

This paper shows that incorporating learning in regulatory enforcement has implications for the agency and the firms' strategies, as well as for total external costs. We assume that the regulatory agency has the possibility to learn about the true types of the firms it is confronted with through inspection, but only if it finds subgroups of firms performing differently. This type of learning is used afterwards to target known types in the subsequent regulatory period. Since the agency has a fixed enforcement budget (which sometimes can be quite small), we show that it will devote more enforcement resources in the next period to auditing the known high-cost firms (whose reactions to a change in the inspection probability are larger). Only if the agency can successfully induce compliance of the known high-cost firms and it has money left, will it devote effort (resources) to try to improve compliance in the next group, i.e. those firms that where not inspected in the first period. Money left (if any) will then be devoted to the least-efficient group, i.e. the known low-cost firms.

In principle, this targeting strategy can be detrimental for high-cost firms but beneficial for low-cost firms. Therefore, high-cost firms may have an incentive to avoid this situation by trying to mimic the low-cost firms; although low-cost firms may try to avoid being mimicked.

We show that the likelihood of a pooling equilibrium (that is, one in which high-cost firms successfully mimic low-cost firms) depends positively on the agency's budgetary constraint, and negatively on the number of firms in the industry and the monitoring costs. These three factors crucially determine the probability of being inspected in the first period; the larger this probability, the more prone the agency is to target high (low)-cost firms with larger (lower) inspection frequencies in the next period. Also, note that stimulating the dispersion of available technologies, e.g. by subsidizing best available technologies, makes externality reductions cheaper and thus pooling behavior more attractive.

To prevent mimicking, our results suggest that the agency should not be given a large budget. But interestingly, if the agency significantly discounts future externality reductions and thus focuses on present gains and if the proportion of high-costs firms in the industry is sufficiently large, mimicking, or even the threat of mimicking, might be good for society. These social benefits arise when we compare the externality levels with the separating equilibrium where high-cost firms do not imitate low-cost firms, because it is just not profitable for them to do so (even if low-cost types do not actively avoid being copied). In any case, this avoidance behavior results in a lower amount of global external costs in the first period. On the one hand, the high-cost types may try to reduce the externality in the first period to avoid future tighter monitoring. On the other hand, the low-cost types may try to reduce the externality as well, in order to differentiate themselves from the high-cost types and thus prevent being pooled.

Sometimes they might even avoid being imitated by doing better than the standard in the first period (i.e. by over-complying).

Several extensions of our model are possible. For example, we have assumed that inspections are accurate, that is, there are no measurement errors. However, in reality, it is sometimes the case that a high-cost firm is wrongly thought to be a low-cost firm (*false positives*). This then implies that the information obtained by the agency from the first period is less valuable than with perfect inspections, and therefore benefits from mimicking are reduced.

Interestingly, pooling is not the likely outcome when the regulatory time horizon is sufficiently large. There would be many periods with imitation costs for the high-cost firms, while the benefit of mimicking would only be reached asymptotically (in the 'last' period). This observation also implies that a reversal in strategy (from pooling to separating or from separating to pooling) is never optimal either. Thus, in an infinitely repeated game, the agency is always able to learn the firms' true type through inspections, since high-cost firms never pose as low-cost firms. Finally, the infinitely repeated game converges to a steady-state equilibrium, which involves a cost efficient allocation of the available funds. This implies that the agency's resources are used where they cause the greatest reduction in the level of the externality. As a result, it might be optimal for the agency to exclusively target one group of firms and to completely ignore the other firms, if the available budget is sufficiently small.

APPENDIX

Proof of Lemma 1.

The first order conditions of this optimization problem are:[12]

$$c_e\left(\theta_i,e\right)+pf-\lambda=0$$
$$\lambda\geq0, e-\overline{e}\geq0,\ \lambda\left[e-\overline{e}\right]=0$$

where $\lambda\geq0$ is the Kuhn-Tucker multiplier associated to the inequality restriction $e-\overline{e}\geq0$. Easily combining these conditions, we obtain the desired result.

Proof of Lemma 2.

The first order conditions of the problem are:

$$-\gamma_i f+\lambda mN_i=0,$$
$$N_i-\gamma_i c_{ee}\left(\theta_i,e_i\right)-\beta_i+\alpha_i=0,\tag{6}$$

where $\gamma_i\geq0$ are the Kuhn-Tucker multipliers associated with the firm types' optimal responses $\left(c_e\left(\theta_i,e\right)+pf\geq0\right)$, $\beta_i\geq0$ are associated with $e_i-\overline{e}\geq0$, $\alpha_i\geq0$ are associated with $e_i^0\geq e_i$, and $\lambda\geq0$ is the one associated to the agency's budgetary constraint.

Consider first the possibility of an interior solution, $\overline{e}<e_i<e_i^0$. This implies $\alpha_i=\beta_i=0$ for all i. Combining the optimality conditions, we then obtain the condition $c_{ee}\left(\theta_H,e_H\right)=c_{ee}\left(\theta_L,e_L\right)$. However, our assumptions $c_{\theta ee}\left(\theta_i,e\right)<0$ for all $e\leq e_i^0$ and $c_{eee}\left(\theta_i,e\right)$ sufficiently small ensure $c_{ee}\left(\theta_H,e_H\right)<c_{ee}\left(\theta_L,e_L\right)$ for all $\left(e_H,e_L\right)$. Therefore, only corner solutions are possible.

Case (i). $\beta_i\geq0$. This implies $e_i=\overline{e}$ for all i (and consequently, $e_i<\overline{e}$ if $\overline{e}<e_i^0$, which

---

[12] Given the assumptions of our model, these are necessary and sufficient conditions for an optimum. The same applies for the remaining optimization problems in the paper.

then implies $\alpha_i = 0$). Thus, $p_i^* = \bar{p}_i$ and $B \geq m[N_H \bar{p}_H + N_L \bar{p}_L]$.

Case (ii). $\beta_H \geq 0$ and $\beta_L = 0$. This implies $e_H = \bar{e}$ and $e_L \geq \bar{e}$, and $\alpha_H = 0$, $\alpha_L \geq 0$.

Considering first that $e_L < e_L^0$ (and therefore, $\alpha_L = 0$), we then have

$$N_H - \gamma_H c_{ee}(\theta_H, e_i) = \beta_H \geq 0 \quad \text{and} \quad N_L - \gamma_L c_{ee}(\theta_L, e_L) = -\alpha_L \leq 0, \quad \text{which imply,}$$

respectively, $\dfrac{N_H}{\gamma_H} \geq c_{ee}(\theta_H, \bar{e})$ and $\dfrac{N_L}{\gamma_L} \leq c_{ee}(\theta_L, e_L)$. Since $\dfrac{N_H}{\gamma_H} = \dfrac{N_L}{\gamma_L} = \dfrac{f}{\lambda m} > 0$, we

then have $c_{ee}(\theta_H, \bar{e}) \leq c_{ee}(\theta_L, e_L)$, which holds under our assumptions $c_{ee}(\theta_i, e) > 0$

for all $e \leq e_i^0$ and $c_{eee}(\theta_i, e)$ sufficiently small.[13] Then, in this case we have $p_H^* = \bar{p}_H$

and $0 < p_L^* < \bar{p}_L$, and the condition on the budget is $mN_H \bar{p}_H \leq B \leq m[N_H \bar{p}_H + N_L \bar{p}_L]$.

Case (iii). $\beta_H \geq 0$, $\beta_L = 0$ and $\alpha_L \geq 0$, which then implies $e_H = \bar{e}$ and $e_L = e_L^0$.

Following a similar reasoning as that of case (ii), we obtain the condition

$c_{ee}(\theta_H, \bar{e}) \leq c_{ee}(\theta_L, e_L^0)$, which holds under our assumptions. Thus, we have $p_H^* = \bar{p}_H$

and $p_L^* = 0$ and the condition on the budget is $B \leq mN_H \bar{p}_H$.

Proof of Proposition 2.

Case (i). $B \geq m[N - n_L]\bar{p}_H$. By case (i) of Proposition 1, we have $p_{2H}^* = p_{2N}^* = \bar{p}_H > p$

and $p_{2L}^* = \dfrac{B - m(N - n_L)\bar{p}_H}{mn_L} \geq 0$, which implies $e_{2H} = e_{2N} = \bar{e}$, and $e_{2L} > \bar{e}$. Since

$p > p_{2L}^*$, we trivially obtain $p_{2H}^* = p_{2N}^* = \bar{p}_H \geq p > p_{2L}^* \geq 0$.

---

[13] Following a similar reasoning in the alternative case $\beta_H = 0$, $\beta_L \geq 0$ leads to the condition $c_{ee}(\theta_L, \bar{e}) \leq c_{ee}(\theta_H, e_H)$, which does not hold under our assumptions.

Case (ii). $m\dfrac{n_H N}{n}\overline{p}_H \leq B < m\left[N-n_L\right]\overline{p}_H$. Since $n_H < \dfrac{n_H N}{n}$, by case (ii) of Proposition

1 we then have $p^*_{2H} = \overline{p}_H$, $p^*_{2N} = \dfrac{B - mn_H\overline{p}_H}{m(N-n)} < \overline{p}_H$ and $p^*_{2L} = 0$. Since $p = \dfrac{B}{mN}$,

$p^*_{2N} \geq p$ if, and only if, the percentage of decrease in the available funds to deter

unknown firms is lower than the percentage of decrease in the number of unknown

firms. That is, if and only if $\dfrac{n_H\, m\, \overline{p}_H}{B} \leq \dfrac{n}{N}$. This inequality holds, since $m\dfrac{n_H N}{n}\overline{p}_H \leq B$.

Thus, the complete ranking in this case is $p^*_{2H} = \overline{p}_H > p^*_{2N} \geq p > p^*_{2L} = 0$.

Case (iii). $mn_H\overline{p}_H \leq B < m\dfrac{n_H N}{n}\overline{p}_H$. Since $\dfrac{n_H N}{n} \leq N - n_L$, we are again under case (ii)

of Proposition 1. Following a similar reasoning as above, we now obtain $p^*_{2N} \leq p$

because $m\dfrac{n_H N}{n}\overline{p}_H \geq B$, and the ranking is $p^*_{2H} = \overline{p}_H > p \geq p^*_{2N} > p^*_{2L} = 0$.

Case (iv). $B < mn_H\overline{p}_H$. This corresponds to case (iii) of Proposition 1 and, since

$\overline{p}_H > p^*_{2H}$ and $p^*_{2N} = p^*_{2L} = 0$, we trivially obtain $\overline{p}_H > p^*_{2H} > p > p^*_{2N} = p^*_{2L} = 0$.

REFERENCES

Arora, S., Gangopadhyay S., 1995. Toward a theoretical model of voluntary overcompliance. Journal of Economic Behavior and Organization 28, 289-309

Becker, G. S., 1968. Crime and punishment: An economic approach. Journal of Political Economy 76(2), 169-217.

Cho, I., Kreps, D.M., 1987. Signaling games and stable equilibria. The Quarterly Journal of Economics 102(2), 179-221

Fudenberg, D., Tirole, J., 1991a. Perfect bayesian and sequential equilibrium. Journal of Economic Theory 53, 236-260.

Fudenberg, D., Tirole , J., 1991b. Game theory. MIT Press.

Greenberg, J., 1984. Avoiding tax avoidance: A (repeated) game-theoretic approach. Journal of Economic Theory 32, 1-13.

Grossman, S., 1981. The role of warranties and private disclosure about product quality. Journal of Law and Economics 24, 461-483

Harrington, W., 1988. Enforcement leverage when penalties are restricted. Journal of Public Economics 37, 29-53.

Kreps, D.M., Sobel, J., 1994. Signalling. In: Aumann, R.J., Hart, S. (Eds.). Handbook of Game Theory, vol.2. Amsterdam: Elsevier.

Innes, R., 2005. Violators avoidance activities and self-reporting in optimal law enforcement. Journal of Law, Economics, and Organization 17(1), 239-256.

Landsberger, M., Meilijson, I., 1982. Incentive generating state dependent penalty system. Journal of Public Economics 19, 333-352.

Malik, A. S., 1990. Avoidance, screening and optimum enforcement. The RAND

Journal of Economics 21(3), 341-353.

Nyborg K, Telle, K., 2006. A dissolving paradox: Firms' compliance to environmental regulation. Environmental and Resource Economics 35(1), 1-18

Reinganum, J., Wilde, L., 1986. Settlement, litigation, and the allocation of litigation costs. The RAND Journal of Economics 17, 557-566.

Riley, J.G., 1979. Informational equilibrium. Econometrica 47(2), 331-359.

Rothschild, M., Stiglitz, J.E., 1976. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. Quarterly Journal of Economics 80, 629-649.

Slemrod, J., Yitzhaki, S., 2002. Tax avoidance, evasion, and administration. In: Auerbach, A.J., Feldstein, M. (Eds.). Handbook of Public Economics. Amsterdam: Elsevier Science.

Sobel, J., 1989. An analysis of discovery rules. Law and Contemporary Problems 52, 133-159

Spence, M., 1973. Job market signaling. Quarterly Journal of Economics 87, 355-374.