LIBERA UNIVERSITA' INTERNAZIONALE DEGLI STUDI SOCIALI

Guido Carli

Dipartimento di Scienze economiche e aziendali

Environmental Monitoring: A General Equilibrium Approach

di

A. Casagrande – M. Spallone

QUADERNI DPTEA n. 160

Roma, dicembre 2008

Abstract

An firm's decision of obeying environmental regulatory standards depends crucially on its chances of being detected and the consequent costs. We investigate the relationship between the amount of resources devoted to environmental monitoring and the extent of non-compliancy. In our model a population of firms, each of whom decides whether or not to be compliant, and a monitoring agency that can detect non-compliance only by monitoring signals strategically interact. Each firm produces a signal, the distribution of which is correlated with his action. Since the agency has resource constraints, it must choose some (optimal) fraction of the signals to monitor. The key property of this model is that a more aggressive monitoring policy may, in the long run, free the resource constraint of the agency (although in the short-run the demand of resources increases) as long as enough firms, perceiving a higher chance of being detected become compliant. Two main implications follow. First, there may be multiple equilibria with different levels of non-compliancy. Second, any increase in government resources generates a magnification effect which beneficially affects the equilibrium level of compliancy.

Keywords: Environment, monitoring, general equilibrium
JEL codes: Q50, Q58

Environmental Monitoring: A General Equilibrium Approach

A. Casagrande* and M. Spallone†

December 2008

1 Introduction

Environmental protection is a highly considered issue in many countries. In fact, since all economic activities generate negative environmental externalities, such as pollution, goverments usually impose environmental taxes and emission standards to producers. The main problem with this regulatory approach is that firms' compliance is not guaranteed.

We assume that environmental regulation is not fully enforceable because monitoring is costly and only a fraction of the firms can be monitored by a dedicated monitoring agency[1]. We deal only with intentional violations of emission standards and we do not consider accidental violations, due for example to negligence[2]. Moreover, we do not deal with environmental taxes.

This paper addresses the following questions: what is the equilibrium fraction of non-compliant firms within a population of firms with varying propen-

*The Core Consulting, Inc., Rome, Italy

†Corresponding Author: mspallone@luiss.it. G. D'Annunzio University of Pescara and Luiss Guido Carli University of Rome

[1]Environmental compliance topics are extensively reviewed in Cohen (1999)[5].

[2]A classification of environmental violations can be founf in Cropper and Oates (1992)[7].

sities to violate emission standards when the monitoring agency is budget constrained; what impact can an increase in the monitoring agency budget have on the equilibrium fraction of non-compliant firms.

In our model, a population of firms with varying propensities to violate regulatory standards face a discrete choice, either to obey the regulatory standards or to violate them. Each action generates a random signal, and we assume that the higher the signal (for example, a higher concentration of pollutants in the air, or inthe water), the more likely a non-compliant behavior has generated it. The monitoring agency is able to observe all the signals, but can detect non-compliant firms only by monitoring them. Since there are resource constraints and monitoring is costly, the monitoring agency must choose some (optimal) fraction of the firms (signals) to monitor.

The key property of this model is that a more aggressive monitoring policy may, in the long run, free the resource constraint of the monitoring agency (although in the short-run the demand of resources increases) as long as enough firms, perceiving a higher chance of being detected, become compliant. Two main implications follow. First, there may be multiple equilibria with different levels of non-compliancy. Second, as the resources available to the monitoring agency increase, more signals can be monitored and more firms will choose to comply. A general equilibrium effect (we call it "magnification effect") is generated by the above key property of the model. So, if the monitoring agency is provided with additional resources and is therefore allowed to be more aggressive, a larger number of firms will switch to compliancy.

This result is very different from that obtained by Harford (1978)[11], who argues that firms' compliance (with waste / emission standards) does not depend on the monitoring agency budget. The difference with our findings may be explained by the fact that in our model each firm interacts not only with the monitoring agency, but also with all other firms. This complex interaction drives our general equilibrium result. Stadler and Perez-Castrillo (2006)[14] consider a similar interaction among a population of heterogeneous firms and a monitoring agency. However, in our paper the firms' probability of being monitored depends crucially on their signals (i.e., on the level of their emissions), while in their paper it is contingent to the firms' pollution (self) report. Moreover, while they adopt a principal-agent approach assuming perfect commitment of the enforcement agency, we model the monitoring problem as a Bayesian game where the monitoring agency has no commitment capacity.

Finally, the existence of a magnification effect is consistent with a robust empirical evidence (Dasgupta *et al* (2001)[8], Foulon *et al* (2002)[10]) that proves that not only monitoring, but also the threat of monitoring are useful in reducing the level of pollution.

Two more sections and an Appendix follow this introduction. Section 2 describes the theoretical model and the results. Section 3 concludes the paper. The Appendix contains the stability analysis performed over the multiple equilibria of the model.

2 The Model

There is a monitoring agency willing to apprehend as many non-compliant firms as possible given the amount of resources available for monitoring. For simplicity, it is assumed that only a fixed fraction of the firms R can be monitored, with $0 < R < 1$. It is also assumed that every non-compliant firm is apprehended if monitored. Hence, the specific objective of the monitoring agency is to maximize the probability of apprehending non-compliant firms by choosing a subset of firms for monitoring. This role of the monitoring agency is consistent with the view of Garvie and Keeler (1994)[9] who assume that the agency's goal is to achieve the highest level of compliance given its budget.

There is a continuum of firms. Each firm has a firm-specific cost of getting apprehended, $c \in \Re$, which is private information and is independently distributed across the population of firms with cumulative distribution function F[3]. Each firm can choose whether to be non-compliant (D) and break the law or to be compliant (H). The firms who decide to be compliant receive a positive payoff equal to B_H with certainty. If non-compliant, they get a higher benefit , $B_D > B_H$, but they suffer their cost c if apprehended[4]. The firm-specific cost may be thought as the sum of a common pecuniary fine and a non-pecuniary penalty that depends on community pressure and social norms (Pargal and Wheeler (1996)[13], Hettige et al. (1996)[12], Arora and Cason (1996)[1], Brooks and Sethi (1997)[4]).

In either case, firms send a one dimensional signal that is observable by the monitoring agency. However, the agency cannot establish *a priori* whether signals are generated by a willfull violation of the regulatory standards, or by a random act of nature[5]. It is assumed that the signals can take any value s in some connected subset of $\Re$. Let $g(s \mid D)$ and $g(s \mid H)$ be the density functions of the signals, given compliant and non-compliant behavior respectively. It is assumed that $g(s \mid D)$ and $g(s \mid H)$ are positively defined over all s. Denote by $G(s \mid D)$ and $G(s \mid H)$ the respective cumulative distribution functions.

We assume that *ex ante* each firm doesn't know what its signal is going to be. Nonetheless, each firm knows that when it does not comply, its signal is likely to be relatively higher than when she does not, as *Assumption 1* below

[3]Our results do not depend crucially on the assumption of the underlying distribution of non-compliance costs; it is possible to show that they hold even by assuming that the population is almost homogeneous.

[4]We are assuming that the net benefit of non-compliancy, $B_D - B_H$, is the same for all firms. This assumption is based on the idea that firms willing to obey the regulatory standard bear some fixed cost (for example, the cost of anti-pollution devices) that non-compliant firms do not bear.

[5]Stochastic pollution was examined by Beavis and Walker (1983)[3], Beavis and Dobbs (1987)[2], and Cohen (1987)[6].

clarifies:

Assumption 1 *(Monotone likelihood ratio property). The likelihood ratio* $\frac{g(s|H)}{g(s|D)}$ *is decreasing in s.*

The following *Lemma* is an immediate implication of *Assumption 1* and fully defines the monitoring strategy of the agency.

Lemma 1 *Let* $P(D)$ *be the fraction of the population of firms that does not comply. Then, the marginal density of the signals* $h(s)$ *may be defined as:*

$$h(s) = g(s \mid D) P(D) + g(s \mid H)(1 - P(D))$$

If $0 < P(D) < 1$*, then the government monitors the highest signals and chooses the unique* s^* *satisfying:*

$$(1 - G(s^* \mid D)) P(D) + (1 - G(s^* \mid H))(1 - P(D)) = R \tag{1}$$

In view of Lemma 1, a monitoring agency policy may be defined by s^*.

Proof *Define* $P(D \mid s)$ *as the probability of non-compliant behavior associated with each signal* s*. We show first that* $P(D \mid s)$ *is increasing in* s*.*

Let s_L *and* s_H *be two signals with* $s_L < s_H$*. Then:*

$$\frac{P(D \mid s_L)}{P(D \mid s_H)} = \frac{g(s_H \mid D)}{h(s_H)} \frac{h(s_L)}{g(s_L \mid D)} > \frac{g(s_H \mid H)}{g(s_L \mid H)} \frac{h(s_L)}{h(s_H)} = \frac{P(H \mid s_H)}{P(H \mid s_L)}$$

t*he strict inequality being a consequence of Assumption 1. Since for all* $s \in S$*,* $P(H \mid s) + P(D \mid s) = 1$*, the result follows. Since* $P(D \mid s)$ *is an increasing function of* s*, then the monitoring agency is going to monitor the* *highest signals. By monotonicity of* $G(s^* \mid D)$ *and* $G(s^* \mid H)$ *there is a unique* s^* *satisfying equation 1.* ∎

Even if the monitoring policy is fully defined by s^*, each firm does not know whether her signal is going to be higher or lower than s^*. Then, for any choice s^* by the monitoring agency, the probability of apprehension for those cheating is $1 - G(s^* \mid D)$.

Each firm must decide whether to be compliant or not. A strategy σ maps each firm-specific cost into the binary set $\{D, H\}$. Let $B_N = B_D - B_H > 0$ denote the net benefit from being non-compliant. For any s^*, let $c^*(s^*)$ be such that:

$$c^*(s^*) = \frac{B_N}{1 - G(s^* \mid D)} \tag{2}$$

Then, the best strategy for each individual with cost c is given by:

$$\sigma(c) = \begin{cases} H, & \forall c \geq c^*(s^*) \\ D, & \forall c \leq c^*(s^*) \end{cases}$$

As it is obvious, the strategy of each firm depends on the monitoring policy as firms perceive any monitoring policy into their probability of getting apprehended.

From now on, c^* will denote $c^*(s^*)$. Hence, the fraction of non-compliant firms in the population, $P(D)$, is endogenous and equal to $F(c^*)$.

The Equilibria

Roughly speaking, an equilibrium is a monitoring policy and a strategy for all firms such that each firm does not want to change his action, given the monitoring agency policy, while the monitoring agency makes an efficient use of all its resources, given the firms' behavior. An equilibrium is defined as a pair of $\{s^*, c^*\}$ which satisfies both equations (1) and (2). To eliminate equilibria in which the entire population of firms chooses either to be compliant or non-compliant, we introduce the following assumption:

Assumption 2 $0 < F(c) < 1$ for all $c > B_N$.

Assumption 2 implies:

Lemma 2 *In equilibrium* $0 < P(D) < 1$.

Since $P(D) = F(c^*)$, we may rewrite equation (1) as:

$$(1 - G(s^* \mid D)) F(c^*) + (1 - G(s^* \mid H))(1 - F(c^*)) = R \tag{3}$$

Then, equation (2) may be seen as the reaction function of the firms to any monitoring agency policy, while equation (3) is the response of the monitoring agency to any c^*. It is possible to solve for c^* in equation (2) and substitute into equation (3) in order to obtain the following equation in s^*:

$$J(s^*; R) = F\left(\frac{B_N}{1 - G(s^* \mid D)}\right)(1 - G(s^* \mid D)) +$$

$$+ \left(1 - F\left(\frac{B_N}{1 - G(s^* \mid D)}\right)\right)(1 - G(s^* \mid H)) - R = 0 \tag{4}$$

We can interpret $J(s^*;R)$ as the excess demand of resources by the monitoring agency. Given that firms are choosing their optimal c^*, whenever $J(s^*;R) > 0$, the monitoring agency exceeds its budget constraint, while $J(s^*;R) < 0$ implies that the monitoring agency is not using all the available resources.

Multiplicity of Equilibria

Depending on the properties of F and G, equation (4) may be satisfied by more than one s^* for some feasible R. Note that the limits of $J(s^*;R)$ as s^* approaches its upper and lower bound are $-R$ and $1-R$ respectively. Then, uniqueness of solution for any possible R requires $J(s^*;R)$ being monotone increasing in s^*. Let f be the density function of the firm-specific costs. Then, by differentiating, and using equation (2) to replace $\frac{B_N}{G(s^*|D)}$ with c^*, we have:

$$J'(s^*;R) = F(c^*)(-g(s^* \mid D)) + (1-F(c^*))(-g(s^* \mid H)) +$$

$$+\frac{g(s^* \mid D)\,B_N}{(1-G(s^* \mid D))^2} f(c^*)(G(s^* \mid H) - G(s^* \mid D)) \tag{5}$$

We can interpret the first two terms in equation (5) as the additional resources needed by the monitoring agency to monitor an increasing fraction of firms (both compliant and non-compliant), given the original cutoff cost c^*. The third term in equation (5) is the source of multiple equilibria, since under the usual monotonicity conditions it is always positive. It measures the amount of resources available to the monitoring agency since a fraction of the population of firms becomes compliant. It has two components.

The first component $\frac{g(s^*|D)B_N}{(1-G(s^*|D))^2}$ represents the adjustment in the cutoff cost of the firms due to the change of the monitoring agency policy s*. The second component $f(c^*)(G(s^* \mid H) - G(s^* \mid D))$ measures the corresponding budget relief for the monitoring agency. So, as more signals are monitored, an increasing fraction of the firms, by perceiving tougher controls, becomes compliant. The effect is a relief of the budget constraint of the monitoring agency. So, as more signals are monitored, an increasing fraction of the firms, by perceiving tougher controls, becomes compliant. The effect is a relief of the budget constraint of the monitoring agency.

Up to now our analysis has been static. We address the stability issues about the equilibria described above in the appendix.

Comparative Statics

Within the theoretical framework described above, it is also possible to analyze the effects of a change in R on the equilibrium fraction of non-compliant firms. In particular, it is possible to isolate an impact or short-run effect as well as a complementary magnification effect, which is the general equilibrium effect generated by our model.

To compute these, we show again the system of equations that characterizes the model. Equations (2) and (3) are rewritten by splitting the endogenous (c^* and s^*) and the exogenous variables (R and B_N) of the system. We get:

$$\begin{cases} (1 - G(s^* \mid D))\, c^* = B_N \\ (1 - G(s^* \mid D))\, F(c^*) + (1 - G(s^* \mid H))\,(1 - F(c^*)) = R \end{cases}$$

The *total effect* (TE) of the change in R on the fraction of non-compliant firms can be measured, using total derivatives over the above system of equations:

$$TE = \frac{dF(c^*)}{dc^*}\frac{dc^*(R)}{dR} = f(c^*)\frac{B_N g(s^* \mid D)}{(1-G(s^* \mid D))^2}\left(\frac{1}{J'(s^*;R)}\right)$$

As mentioned before, the *total effect* can be splitted into the *impact effect* and the *magnification effect*. By the *impact effect* (IE) it is meant the change in the fraction of non-compliant firms following a change in R if no further reactions by firms were considered. A monitoring agency, which does not take into account general equilibrium considerations, would think of it as the total effect:

> *After an increase in R, given the cutoff level cost c^*, the monitoring agency is able to decrease s^*; as a consequence, people decrease c^*; since the probability of apprehension increases, more firms decide to be compliant, and the fraction of non-compliant firms decreases.*

The impact effect is given by:

$$IE = \frac{f(c^*)}{F(c^*)(-g(s^* \mid D)) + (1-F(c^*))(-g(s^* \mid H))}\left(\frac{B_N g(s^* \mid D)}{(1-G(s^* \mid D))^2}\right)$$

The *magnification effect* (ME) is defined as the further decrease in the fraction of non-compliant firms once the impact effect has taken place:

> *Because of the impact effect, the signals sent by the firms have shifted downward, and less signals are now above the actual policy; so, the monitoring agency is able to further decrease s^*; hence, firms fur-*

ther decrease c^*, and the fraction of non-compliant firms further decreases.

We can measure this magnification effect as follows:

$$ME = \frac{F\left(c^*\right)\left(-g\left(s^* \mid D\right)\right) + \left(1 - F\left(c^*\right)\right)\left(-g\left(s^* \mid H\right)\right)}{J'\left(s^*; R\right)} = \frac{1}{1-m}$$

where

$$m = \frac{\left(\frac{g(s^*|D)B_N}{(1-G(s^*|D))^2}\right) f\left(c^*\right)\left(G\left(s^* \mid H\right) - G\left(s^* \mid D\right)\right)}{F\left(c^*\right)\left(-g\left(s^* \mid D\right)\right) + \left(1 - F\left(c^*\right)\right)\left(-g\left(s^* \mid H\right)\right)}$$

Notice that the magnification effect takes values always greater than 1, as long as $m < 1$. We can interpret m as a general equilibrium multiplier. The numerator and the denominator of m are the positive and the negative component of $J\left(s^*; R\right)$ respectively. So, the higher the budget relief by the monitoring agency following any decrease in s^*, the lower the additional resources needed to monitor more signals, the higher m. Furthermore, as m approaches 1, the magnification effect increases and becomes very relevant.

3 Concluding Remarks

The main result of this paper is that, although there may be multiple equilibria with different levels of non-compliancy, in order to enforce the law, it is beneficial to increase monitoring.

Our results rely on two main assumptions: first, on the hypothesis that the monitoring agency spends its money efficiently; second, on the assumption that the agency itself is not corrupt, or that beyond corrupt officers there exists a clean public authority responsible for monitoring them. If more public resources do not imply either better or higher numbers of controls at the same (high) quality, and/or if the agency is itself corrupt, there is obviously no advantage in increasing public budgets.

Furthermore, we assume that the government has an exogenous amount of resources that can be spent to reduce non-compliancy. These resources may be thought of as the result of some political process.

This work can be extended in several other ways. First, one could allow social networks as a channel for propagating either honest or dishonest behavior, using frameworks such as local interaction analysis. Second, one can look at different methodologies or technologies of monitoring. Third, use this framework to analyze other phenomena such as direct or indirect tax evasion, or corruption.

Appendix: Stability Analysis

We suppose that both the government and the firms act in a myopic incremental perspective. More specifically we suppose that, for given c^*, the government decreases its policy s^* over time in order to correct any budget surplus and increases it to adjust any budget deficit. Likewise we assume that, for given s^*, firms revise c^* to keep track of any government policy change. They increase c^* if net benefits exceed the expected costs at the actual threshold.

The above assumptions ensure that all stable equilibria are plausible. The government does not have a first mover advantage and cannot commit to the highest stable equilibrium policy. The motivations behind these assumptions are two. On the one hand, usually firms are conservative in their behavior and monitoring agencies do not dramatically alter their policies. On the other hand, it would be difficult, even for the monitoring agency, to infer the distributions of the reported incomes and of the non-compliancy costs.

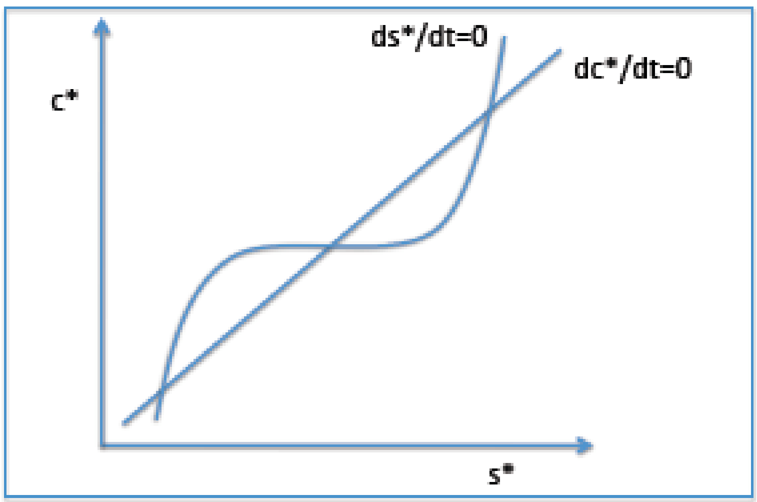
Let $\dot{s}^*$ and $\dot{c}^*$ be the derivative of s^* and c^* with respect to time respectively. The following system of equations describes the above behavior:

$$\begin{cases} \dot{s}^* = \left(1 - G\left(s^*\left(t\right) \mid D\right)\right) F\left(c^*\left(t\right)\right) + \left(1 - G\left(s^*\left(t\right) \mid H\right)\right)\left(1 - F\left(c^*\left(t\right)\right)\right) - R \\ \dot{c}^* = \frac{B_N}{1 - G(\dot{s}^*(t) \mid D)} - \dot{c}^*\left(t\right) \end{cases}$$

The government increases s^* as non-compliancy increases, and firms decrease c^* as monitoring policy gets more aggressive. At any locus (s^*, c^*) below $\dot{s}^* = 0$ the government decreases s^* to correct its budget surplus, whereas below $\dot{c}^* = 0$ firms increase c^* since expected costs at the current s^* are less than net benefits.

We define an equilibrium $(\bar{s}^*, \bar{c}^*)$ to be stable if there exists a sufficiently small neighborhood of it such that any initial pair (s^*, c^*) in the neighborhood asymptotically converges to $(\bar{s}^*, c^*)$. Generically speaking, extreme equilibria are stable, and between any two stable equilibria there is an unstable one, as it is clear from Figure 1.

Figure 1



In fact, through the linearization of our system of equations around any steady state $(\bar{s}^*, \bar{c}^*)$, we get:

$$\begin{vmatrix} \dot{s}^* \\ \dot{c}^* \end{vmatrix} - \begin{vmatrix} A & B \\ C & -1 \end{vmatrix}_{(\bar{s}^*,\bar{c}^*)} \begin{vmatrix} s^* \\ c^* \end{vmatrix} = \begin{vmatrix} 0 \\ 0 \end{vmatrix}$$

where

$$A = -g(s^* \mid D)\, F(c^*) - g(s^* \mid H)\,(1 - F(c^*))$$

$$B = f(c^*)\,(G(s^* \mid D) - G(s^* \mid H))$$

$$C = \frac{B_N g(s^* \mid D)}{(1 - G(\dot{s}^* \mid D))^2}$$

The trace and the determinant of the Jacobian matrix is all we need for inferring over the stability of any equilibrium. It easily turns out that the trace is always negative, and the sign of the determinant is positive if and only if $J'(s^*;R) > 0$. Then we may conclude that whenever $J'(s^*;R) > 0$ the equilibrium is stable, and whenever $J'(s^*;R) < 0$ the equilibrium is unstable. Since, as we said above, the limits of $J'(s^*;R)$ as s^* approaches its upper and lower bound are $-R$ and $1-R$ respectively, extreme equilibria are stable, and between any two stable equilibria there is an unstable one.

References

[1] Arora, S. and T. Cason (1996). “Why do Firms Volunteer to Exceed Environmental Regulations? Understanding Participation in EPA’s 33/50 Program”, Land Economics, 413-32.

[2] Beavis, B. and I. Dobbs. (1987). "Firm Behavior Under Regulatory Control of Stochastic Environmental Wastes by Probabilistic Constraints", Journal of Environmental Economics and Management, 112-127.

[3] Beavis, B. and M. Walker. (1983). "Random Wastes, Imperfect Monitoring and Environmental quality Standards", Journal of Public Economics, 377-87.

[4] Brooks, N. and R. Sethi. (1997). “The Distribution of Pollution: Community Characteristics and Exposure to Air Toxics”, Journal of Environmental Economics and Management, 233-250.

[5] Cohen, M. A. (1999). “Monitoring and Enforcement of Environmental Policy”, International Yearbook of Environmental and Resource Economics, 44-106.

[6] Cohen, M. A. (1987)."Optimal Enforcement Strategy to Prevent Oil Spills: An Application of a Principal-Agent Model with ’Moral Hazard’", Journal of Law and Economics, 23-51.

[7] Cropper, M.L. and W.E. Oates (1992). “Environmental Economics: A Survey”, Journal of Economic Literature, 675-740.

[8] Dasgupta, S., Laplante, B., Maningi, N. and H. Wang (2001). “Inspections Pollution Prices, and Environmental Performance: Evidence from China”, Ecological Economics, 487–498.

[9] Garvie, D. and A. Keeler (1994). "Incomplete Enforcement with Endogenous Regulatory Choice", Journal of Public Economics, 141-162.

[10] Foulon J., Lanoie, P., and B. Laplante (2002). "Incentives for pollution control: regulation or information?", Journal of Environmental Economics and Management, pp. 169–187.

[11] Harford, J.D. (1978). "Firm Behavior under Imperfectly Enforceable Pollution Standards and Taxes", Journal of Environmental Economics and Management, 26-43.

[12] Hettige, M., S. Pargal, M. Singh, and D. Wheeler. (1996). "Formal and Informal Regulation of Industrial Pollution: Comparative Evidence from Indonesia and the U.S.", World Bank Policy Research Working Paper.

[13] Pargal, S. and D. Wheeler. (1996). "Informal Regulation in Developing Countries: Evidence from Indonesia", Journal of Political Economy, 1314-34.

[14] Macho-Stadler, I. and D. Perez-Castrillo (2006). "Optimal Enforcement Policy and Firms' Emissions and Compliance with Environmental Taxes", Journal of Environmental Economics and Management, pp. 110-131.