Getting Polluters to Tell the Truth*

Julieta Mercedes

June, 2006

Abstract

We study the problem of a regulator who must control the emissions of a given pollutant from a series of industries when the firms' abatement costs are unknown. We develop a mechanism in which the regulator asks firms to report their abatement costs and implements the most stringent emissions standard consistent with the firms' declarations. He also inspects one of the firms in each industry which declared the cost structure consistent with the least stringent emissions standard and with an arbitrarily small probability, he discovers whether the report was true or not. The firm is punished with an arbitrarily small fine if and only if its report was false.

This mechanism is simple, is implementable in practice, its unique equilibrium is truth telling by firms, it implements the first best pollution standards and shares some features of the regulatory processes actually observed in reality.

Keywords: Efficient Emissions Standards, Command and Control, Truth Telling, Full Nash Implementation.

Journal of Economic Literature **Classification numbers:** D02, D78, D82, Q20, Q52, Q53.

1 Introduction

In this paper we study the problem of a regulator who must control the emissions of a given pollutant from a series of industries. He wants firms to produce the optimal amount of pollution, when both the firms' abatement costs and the costs of pollution to society are considered. Such a regulator faces a fundamental problem faced by every regulator worldwide: that he rarely knows the exact nature of the pollution abatement technology of firms, which of course influences the optimal pollution level to be chosen. The regulator must therefore rely on whatever he can learn

*Hugo Hopenhayn and a referee in this journal noted a bug in a previous version of this paper, and we are grateful for their help. We also thank Atila Abdulkadiroglu, Ezequiel Aguirre, Anil Arya, Jean-Pierre Benoît, Carlos Chavez, Marcelo Cousillas, Federico Echenique, Jeff Ely, Néstor Gandelman, Jonathan Glover, Ana María Ibáñez, Matt Jackson, Larry Kotlikoff, Carlos Lacurcia, Stephen Morris and Francesco Squintani.

about firms' costs from the information they are willing to provide. Given the importance of the problem of regulating polluters, the issue of how to truthfully extract information about their costs has been at the heart of both academic and policymaking discussions for almost three decades.

We posit a model in which the regulator asks firms to declare what their cost functions are and uses these announcements to set an emissions standard for each industry: a maximum allowable level of emissions for every firm in that industry. After receiving the reports, the regulator implements in each industry the most stringent standard consistent with the declarations of the firms in that industry. He also inspects one of the firms in each industry which declared the cost structure consistent with the least stringent emissions standard (the firms most likely to be lying). With an arbitrarily small probability, he discovers whether the report was true or not. A firm which was sampled is punished with an arbitrarily small fine if and only if its report was false.

This mechanism has several important features. First, it is very simple, and therefore applicable in practice. In fact, as we will discuss later in more depth, it is very similar to the mechanism actually used in several countries, including the United States' National Pollutant Discharge Elimination System. Second, it fully implements truth telling by the firms, and results in the regulator setting the efficient standard in each industry. That is, since the unique equilibrium of this game is for firms to tell the truth, the informational asymmetry disappears, and the total welfare of society is maximized. Finally, a third advantage of the mechanism is that it is budget balanced: it implies no costs for the regulator.

There are other studies that have proposed mechanisms that both implement truth telling by the firms and result in an efficient level of pollution. The three most relevant works in this area are Kwerel (1977), Dasgupta, Hammond and Maskin (1980) and Spulber (1988). There are a few problems with these prior studies, the main one being that one does not observe the proposed mechanisms in practice. We believe that there are two main reasons why those mechanisms are not observed in reality. Moreover, our mechanism is free of those problems. The first reason why we don't observe those mechanisms in reality is that they are complicated. This has been a standard criticism about the literature of optimal mechanism design. We believe that another reason why previously proposed mechanisms are not observed is that they are based on taxes, subsidies, or tradeable permits and these types of instruments have several implementation problems as compared to classic "command and control" instruments. Although these types of instruments have been used recently, they have applied only in very specific contexts, and their implementation has been slow for several reasons. For example, regulators are not educated in environmental economics and do not see the advantages of these instruments in terms of cost-effectiveness and efficiency; they see "command-and-control" instruments as stronger statements of support for environmental protection. Moreover, regulators usually think that it is immoral to let firms pollute just because they paid some taxes, or because they purchased pollution permits. Policymakers may also be reluctant to impose further costs on firms because of the impact on employment. Also, incentive-based instruments shift control decisions from regulatory staff to polluting firms, possibly affecting the regulator's job security and prestige.[1]

Another problem with the existing theorems in the literature, is that they focus on whether truth telling is a Nash equilibrium of the revelation game, and not on whether truth telling is the unique equilibrium. If declaring large abatement costs is an equilibrium that yields higher profits for all firms, one will not observe firms telling the truth, but rather overestimating their costs. Our theorem is free from that problem, since its unique equilibrium is truth telling.

In Section 6 we will discuss the relationships among our mechanism and those in the literature on implementation, but it suffices here to stress two points. First, the implementability of the regulator's rule in our setting does not follow from any of the existing theorems in the literature. Second, and most important, our focus is not on the novelty of the theoretical arguments in the implementation of the regulator's rule, but on the possibility of actually implementing it in real contexts.

We have argued that our mechanism is simple, shares some features of some regulatory practices around the world, implements truth telling and the efficient level of pollution, and is budget balanced. Also, we have argued that one of the reasons why one does not observe in practice alternative mechanisms that have been proposed in the literature is because they were complicated and relied on taxes and subsidies, which may be too difficult to implement for regulators. We now turn to the discussion of our assumptions.

2 Discussion of Assumptions

Our model is very similar to that in Kwerel (1977) and Dasgupta et al. (1980). In some dimensions our model is more general, and the conclusion of the theorem is stronger, but we make two additional assumptions.[2] First, we assume that if the regulator samples one firm, it can find out, with probability ε, for ε arbitrarily small, whether the report of abatement costs was true or not. Second, we assume that in each of m industries there are at least two firms with the same cost functions.

With the first assumption the asymmetry of information between the regulator and the firms ceases to be absolute. The assumption is quite weak for at least three reasons. First, we assume that the regulator inspects and samples just one firm out of a potentially large pool. Second, we assume that in case the inspection is successful and it provides some information, the regulator only learns whether the report was true or not, but in case of a false report, he does not get to know the true cost function. Third, and most important, the regulator only finds out whether the

[1] These and other arguments are well documented in the literature. See for example Bohm and Russell (1985), Russell and Powell (1996), Lewis (1996), Keohane, Revesz and Stavins (1998).

[2] Like both these works, our model can be applied more generally, and not just to the problem of a regulator trying to fix the right level of pollution.

report is true or not with an arbitrarily small chance. That is, we fix any $\varepsilon > 0$, and the regulator only learns whether the report is true with probability ε.

Our assumption that the asymmetry of information is not absolute is also a reasonable one in the context we study. First, regulators worldwide engage in controlling or monitoring the statements of polluters about the abatement technology to be used, so our assumption reflects a common practice. In the US for example, before starting their operations firms are required to present an exhaustive description of their production processes, abatement technology and costs in order to obtain a pollution discharge permit.[3] Second, this common practice is well founded, since the regulators can check each piece of information provided by the firm, and assess its validity, or even in some cases be more proactive by pointing out to firms how other businesses have coped with the same abatement problems. Engineers from the Environmental Protection Agency study the different abatement technologies available to a particular type of industrial activity and then establish effluent standards for each category of polluter and place of discharge (see Field, 1997). Since the regulation, and the standard-setting, occur at a basic "process level" and not at a more complicated "plant-level," the processes involved are standard across industries, and the regulator has a deep knowledge about costs as illustrated, for example, in the following quotation from the Environmental Protection Agency (1992).

> "The document provides a generic process-by-process assessment of pollution prevention opportunities for the Kraft segment of the pulp and paper industry. The process areas covered are: wood yard operations, pulping and chemical recovery, pulp bleaching, pulp drying and papermaking, and wastewater treatment. These process areas are further broken down by specific process (e.g., oxygen delignification as one specific process under the pulping and chemical recovery area). For each specific process there is a description, a cost estimate, a discussion of applicability, and estimate of environmental benefits."[4]

Both the way the regulatory process takes place, and the depth of the knowledge of the regulator about each individual process suggest that the asymmetry of information between firms and the regulators is not absolute, so that our assumption seems appropriate.

Our second assumption, that there are at least two firms in each industry which have identical cost functions also follows from the way the regulatory process works (i.e. setting emissions standards on a process by process basis). If a firm buys cows and delivers leather shoes, it won't

[3] Several countries have copied extensively the US National Pollutant Discharge Elimination System, including our own Uruguay. Most of such systems share the "inspection" features of the US system that we are interested in.

[4] Similar quotations can be found for other industries. See for example EPA (2002) for the iron and steel industries and their process by process regulation.

have the same abatement costs as a firm that buys cows and delivers leather seats for cars. But both firms will first produce raw hides and then tan the leather. Since both firms need to abate its pollution levels at each individual task, each of which is also undertaken in other firms producing different goods, our assumption reflects the fact that even very complicated production processes are based on some elementary processes that are repeated in several firms even across industries. Another reason why the assumption of at least two firms per industry is not so restrictive is that our exact same model would apply if it was common knowledge that costs in the same industry are just "vertical" translations of each other. That is, if firm 1 has a cost function of c, and firm 2 a cost function of $c + k$, those cost functions are "identical" as far as our mechanism is concerned. Therefore, if a firm in California and a firm in New York buy their abatement technology from a firm in New York, and the price in California is just the price in New York plus shipping, those two firms can be modeled as having identical costs. Finally, as we will argue in Section 5, even if there are some firms that have cost functions that no other firm in the whole economy share, our mechanism can still be used. Suppose that the regulator can estimate the cost functions of these firms and produce estimates which are "close" to the truth. Then, the unique equilibrium of our mechanism (when it is applied among the firms in industries with at least two firms) is still truth telling, and the standards set for each industry are "close" to the first-best, complete information, ones.

Another, less disputable, assumption that we make is that the regulator can fine the firms for lying. This is consistent with the practice of pollution regulators worldwide. In Uruguay, for example, as a consequence of "forgery" in the cost declaration, the person in charge of filling the reports about the abatement technology can be imprisoned. Another potential punishment is the temporary closing of the plant. Similar practices are common elsewhere. It is worth emphasizing that for our mechanism to work, the fine can be arbitrarily small. If fines were large, even a small probability of a false report being uncovered would suffice to make truth telling a dominant strategy. In our mechanism the fine is used exclusively for breaking ties.

We also assume that total damages to society are known or can be estimated. Although this has been the standard assumption in this branch of the literature (see Kwerel (1977) and Dasgupta et al., 1980) it is quite strong. As we will argue later, however, our mechanism is robust to whether the regulator knows total damages exactly, or approximately, or just wants to set a total level of emissions for the whole economy. The first extension is relevant if one is able to estimate total damages to society approximately, and is concerned that the emissions standards will be approximately correct. We show that that is indeed the case: our mechanism still fully implements truth telling, and if the regulator's estimate of total damages are close to the true damages, then the emissions standards that result from our mechanism are close to the ones that would be implemented if the regulator knew exactly the damages to society and abatement costs. In a second relaxation of the assumption that the regulator knows damages, we investigate how our mechanism fares when the regulator does not know, or is not interested in, damages to society, but rather on achieving a certain level of emissions for the whole economy. This extension is important because in practice it is common to proceed in that way. Moreover, the adoption of the Kyoto Protocol implies that the regulatory agencies must find the most efficient way to achieve a certain level of emissions for the economy as a whole. We show that our mechanism can be used to determine the standards which minimize the total cost to society of complying with, say, the Kyoto standards.

In this note we are only concerned with the problem of setting the right emissions standards. The enforcement of those standards is a different issue, and we therefore omit its study. Our mechanism does not assume that there is perfect enforcement, only that higher emissions standards are better for firms. If there is perfect enforcement, then our mechanism maximizes total welfare to society. If there isn't, the emissions standards are the correct ones, but if firms violate the standards, welfare is not maximized, and the regulator must try to maximize compliance subject to its enforcement budget (see footnote 5 for more on this issue).

In Section 5 we discuss three variations of our assumptions with which the mechanism still fully implements truth telling.

3 The Model

There are m industries and n^i, for $i = 1, ..., m$, firms in each industry. Firms in $I^1 = \left\{1, ..., n^1\right\}$ are those in industry 1, firms in $I^2 = \left\{n^1 + 1, ..., n^1 + n^2\right\}$ are those in industry 2 and so on. Each industry has at least 2 firms.

Let $\mathcal{C}$ be a set of strictly convex, decreasing and differentiable functions $c : \mathbf{R}_+ \to \mathbf{R}_+$. Each firm in industry i can abate its pollution level using an abatement technology which has a cost of $c^i(\cdot) \in \mathcal{C}$. That is, $c^i(x_j)$ for firm j polluting a level x_j in industry i is the difference in profits from (a) not engaging in abatement, and (b) abating its potential pollution to level x_j.[5] Note that all firms in each industry have the same cost function.

The cost function c^i is unknown to the regulator. He only knows that $c^i \in \mathcal{C}$ for $i = 1, ..., m$ and that the profile $c = \left(c^1, c^2, ..., c^m\right)$ is drawn from $\mathcal{C}^m$ using some probability distribution P which is common knowledge. In the mechanism of this paper, the regulator asks firms to report their cost functions. In spite of the informational asymmetry, the regulator can inspect one firm. With probability $\varepsilon > 0$ he finds out whether the report was truthful or not, with probability $1 - \varepsilon$ the inspection is inconclusive. In case the regulator discovers that the report was not true, he does not find out the true c^i, but only that the report was false.

[5] If $c^i(x_j)$ is interpreted as the cost of abating pollution to x_j, one is implicitly assuming that there is perfect enforcement, and therefore our mechanism will maximize total welfare. If $c^i(x_j)$ is interpreted as the cost of having a standard of x_j, one is not assuming perfect enforcement, only that higher standards are better. In that case, our mechanism sets the right standard, but eschews the issue of whether they will be enforced.

The total damages to society coming from pollution are a convex and differentiable function $D : \mathbf{R}_+ \to \mathbf{R}_+$ where total damages are given by $D(X)$ and X is the total pollution from every firm in every industry:

$$X^i = \sum_{j \in I^i} x_j, i = 1, ..., m \quad \text{and} \quad X = \sum_1^m X^i$$

We assume that the regulator knows or is able to estimate $D(X)$, but we relax this assumption in Section 5.2. Also, this definition of damages also assumes that what matters is the total level of pollution, and not its geographic distribution. Although this assumption is not essential for our mechanism to work, it can be justified on the grounds that the pollutant to be regulated is "uniformly mixed" in the sense that only the amounts emitted are relevant, and not their place of generation.[6]

In this context, a social choice function is a function $f : \mathcal{C}^m \to \mathbf{R}_+^m$ that specifies for each possible profile of cost functions (one for each industry) the pollution level that each firm must produce. The regulator wishes to implement the social choice function that minimizes the total cost of pollution. Technically, f is a selection from the correspondence $F : \mathcal{C}^m \rightrightarrows \mathbf{R}_+^m$ defined by

$$F(c) = \arg \min_{(x^1,...,x^m)} \left[D\left(\sum n^i x^i \right) + \sum n^i c^i \left(x^i \right) \right], \tag{1}$$

for all $c = \left(c^1, ..., c^m\right) \in \mathcal{C}^m$, where x^i is the standard set for industry i, with which all firms in the industry must comply. We will also assume that $\mathcal{C}$ is the set of all functions c such that for all x, $D'(x) + c'(0) < 0$.

This assumption rules out the possibility that firms declare a cost function that would make the optimal standard for that industry equal to 0. It is a reasonable assumption for regulation of industries or processes that are already functioning, since it just reflects the fact that regulators have chosen not to prohibit those industries or processes.

When c in equation (1) is the true profile of cost functions, the correspondence F yields the first best emission levels: the emission levels that the regulator would choose if he knew the true cost functions. In this paper we will show that our mechanism allows the regulator to find out the true profile of cost functions c, and therefore find the first best emission levels. We will *not*, however, deal with the problem of finding the best allocations for the whole economy, when firms pay to consumers the damage caused. In the problem of finding this optimal allocation when firms have to pay the damage caused, some polluting firms could be forced to close down due to losses.[7] This difference is relevant because, among other things, regulatory agencies in some countries care about the impact of their regulation on the probability of inducing firms to close down. Nevertheless, our take on this problem is the standard one in the literature on Environmental Economics (including the papers most related to ours).

[6] Less importantly, it is the standard assumption in this strand of the literature.

[7] We thank a referee in this journal for bringing this problem to our attention.

It is also worth noting that since our model is static, we are eschewing the problem of collusion among firms. In our static model, the unique equilibrium is truth telling, but if the game of "standard setting" were repeated an infinite number of times, other equilibria (including a collusive outcome in which firms claim high abatement costs) could arise. Since collusion is a widespread problem, it is a drawback of our model. But because we lack a decent theory of equilibrium selection for infinitely repeated games, the same can be said of any static mechanism. Therefore, if collusion is strongly suspected in the regulation of some pollutant (if there are few firms, for example) the best alternative may be the method that has been used the most in the past: estimation of cost functions by the regulator.

4 The Mechanism and the Theorem

We now present our mechanism, and then show that it fully implements f. That is, we will show that in the unique equilibrium of the game designed by the regulator, firms truthfully disclose their cost functions.

For our direct revelation mechanism, the strategy space for each firm is $\mathcal{C}$. Firms must announce their cost functions, and thereby, the cost function of the industry. For each profile of announcements $C = \left(C^1, ..., C^m\right)$, C^i will represent the profile of announcements of firms in industry i, so that

$$C = \left(C^1, ..., C^m\right) = \left(\underbrace{c^1, ..., c^{n_1}}_{\text{industry 1}}, \underbrace{c^{n_1}, ..., c^{n_1+n_2}}_{\text{industry 2}}, c^{n_1+n_2+1}, ..., c^{n_1+n_2+...+n_m}\right). \tag{2}$$

For each profile C let

$$x_j^1 = \min\left\{f_1\left(c^j, c^{p_2}, ..., c^{p_m}\right) : p_i \in I^i, i = 2, ..., m\right\}.$$

The number x_j^1 is the emissions standard that would result for industry 1 if the regulator believed the announcement of firm j in this industry, and chose the announcement of a firm p_i in each remaining industry $i \neq j$ which would result in the most stringent standard for industry 1. A firm with a low x_j^1 is most likely telling the truth, since it is announcing a cost function that could result in a harsh environmental policy. Similarly, define x_j^i for $i = 2, ..., m$ and $j \in I^i$ to be the standard that would be implemented for industry i if the regulator believed the announcement of firm j in that industry. Also, define

$$\underline{x}^i = \min_{j \in I^i} x_j^i \qquad \text{and} \qquad \overline{x}^i = \max_{j \in I^i} x_j^i \tag{3}$$

to be, respectively, the most (least) stringent standard consistent with the announcements of firms in industry i.

Our mechanism is as follows:

1. Firms announce their types
2. If in industry i announcements coincide, the regulator samples randomly one of the firms and inspects it. If the announcements do not all coincide, the regulator: identifies the firms, or firm, which announced the cost functions which are consistent with $\overline{x}^i$; randomly selects one of them and inspects this firm with probability $\pi > \left(n^i - 1\right)/n^i$, and some other firm with probability $1 - \pi$. The idea is to monitor more firms which are most likely lying. A firm is fined if and only if: it is sampled; its report is false; the inspection discovers (with probability ε) that the report was false. The size of the fine does not matter, it can be as small as one wants.
3. The emissions standards $\left(\underline{x}^1, ..., \underline{x}^m\right)$ are implemented.

A strategy for a firm in the game that this mechanism defines is a continuous function $s : \mathcal{C} \to \mathcal{C}$ that announces a cost function for each possible type (real cost function) of the firm.

Theorem 1. *Any efficient social choice function f defined by equation (1) is fully implementable. That is, the unique equilibrium of the direct revelation mechanism, is truth telling.*

The proof is in the appendix.

Remark 1. *It is worth emphasizing that our mechanism has a unique equilibrium.*[8] *As has been argued in the literature on mechanism design the issue of multiplicity is very relevant, especially if the equilibria arising from the game can be Pareto ranked, and it becomes focal to lie. See Moore (1992) p. 186, fn. 5 and the references therein.*

Remark 2. *The central idea of the proof is very simple (which may also help in the actual implementation of the mechanism). First, if one firm is telling the truth, it is optimal for all to tell the truth, since they will relax the standard, and reduce the probability of a fine (to 0). That shows that truth telling is an equilibrium. Second, if all firms are lying, one will have an expected probability of inspection which is larger than the rest. That firm has an incentive to slightly undercut all other firms in its announcement, since it will strictly reduce the chance of an inspection and change the standard only slightly.*

From Remark 2 one can see that if one is willing to use a stronger equilibrium concept, in particular, trembling hand perfection, then one can simplify the mechanism even further by eliminating the probability of the inspection of the firms which are not declaring $\overline{x}^i$. That inspection is

[8] Kwerel (1977) and Dasgupta et al. (1980) study the issue of whether truth telling is an equilibrium. The issue of uniqueness and whether truth telling is the only equilibrium is not analyzed.

used to get rid of equilibria in which in some industry one or more firms declare the truth, and two or more firms declare cost functions which yield standards that are more stringent than the ones corresponding to the truth. Without these inspections, those firms have no incentive to deviate, since the standard will be very low (stringent) even if they declare the truth, and they are not fined if they lie. If one took a finite type space, and all firms mixed on all of their types, there would always be a chance that a firm would be the only one declaring the cost function consistent with the low standard, and it would then be a best response to play the truth.[9]

From a practical point of view, the application of the mechanism as it is may present two difficulties. First, the type space may be too large, and it may be too hard for firms to estimate exactly which is its cost function. Second, and related to the previous point, two firms "trying" to declare the truth may not declare the exact same cost function, and it would not make much sense to punish them in that case. A solution to both of these problems is to present the firms with a fairly large (but finite) menu of cost functions that can be declared, and the authority deems the statement to be true if it is close enough to the truth (the inspection, instead of declaring truth or not, would declare whether the statement is close to the truth or not, which is even easier for the regulator).[10]

It is worth noting that our mechanism can also be used to elicit the optimal Pigouvian taxes.[11] In that case, the regulator still wishes to implement F from equation (1). If he knew the true cost functions c, he would calculate $x = \left(x^1, ..., x^m\right)$, then he would set $t = D'\left(\sum n^i x^i\right)$. Then, the firms' problem in industry i would be to choose x to minimize $c(x) + tx$, so that the optimal emission would be characterized by

$$-c^{i\prime}(x) = t = D'\left(\sum n^i x^i\right).$$

Since this is the first order condition of the regulator's problem when choosing the optimal x, we see that the firm's problem yields the first best levels of pollution.

In order to use our mechanism to implement F via taxes, the regulator would calculate (as before) for each firm j in industry i, the highest tax rate consistent with the firm's declaration. Then, he would calculate for each industry i, $\overline{t}^i = \max t^i_j$ and $\underline{t}^i = \min t^i_j$. Then, in the mechanism, $\overline{t}^i$ would play the role of $\underline{x}^i$ and $\underline{t}^i$ the role of $\overline{x}^i$.

[9] The size of the type space would have to be fairly large for the small ε and small fine to be enough incentive for firms to undercut each other.

[10] One choice of a finite type space for which the unique equilibrium is telling the closest "declarable type" to the truth is the following. Partition the interval $[0, M]$, for large M into intervals of length 1. The menu of cost functions that can be declared is that of costs which have constant derivative in those intervals, and the derivative is a multiple of $1/K$ (for large K). The mechanism is the same, only that the regulator declares a firm to be lying if its declaration is not close enough to the truth (with the metric of the supremum).

[11] We do not pursue this route here, because of their reduced applicability, discussed in the introduction.

5 Different Assumptions

In this Section we briefly discuss three variants of our assumptions and of the mechanism that still fully implement truth telling. This is relevant since the institutional settings may vary from country to country, making some versions impossible to implement, while rendering others feasible.

Before turning to the variations of the model, we note that our two main assumptions are necessary for the mechanism to fully implement truthtelling. If the regulator had no way of finding out whether the firms are lying, the following would be an equilibrium. Suppose there is a maximum potential pollution level in each industry $\overline{X}_i$ (when firms do not engage in abatement), and that firms in industry i report a cost function c_i such that

$$D'\left(\sum n_i \overline{X}_i\right) = -c_i'\left(\overline{X}_i\right)$$

and the regulator then sets the non binding standard $\overline{X}_i$ in industry i. Also, note that even if the regulator could find out whether a report was true with probability ε, as in our model, if a firm were alone in the industry, it would maximize profits by declaring a cost function that yields $\overline{X}_i$ as its standard, provided ε and the fine are sufficiently small.

There are several variants of the mechanism that also yield truthtelling as the unique equilibrium. Here we analyze two. The first variant is concerned with our main assumption: that there are at least two firms in each industry. The second analyzes the case where damages to society are unknown, or there is no interest in determining them.

5.1 Industries with one firm.

Suppose industries 1 through k have just one firm, $n^1 = n^2 = ... = n^k = 1$, and that industries $k+1$ through $k+m$ have at least two firms, as has been our assumption so far. As before, we let I^i be the set of indexes of firms in industry i, even for industries with 1 firm. Again, the regulator wishes to implement the social choice function that minimizes the net cost of pollution. Technically, he wishes to implement a selection f from the correspondence $F : \mathcal{C}^{k+m} \rightrightarrows \mathbf{R}_+^{k+m}$ defined by

$$F(c) = \arg \min_{\left(x^1,...,x^{k+m}\right)} \left[D\left(\sum n^i x^i\right) + \sum n^i c^i\left(x^i\right)\right], \tag{4}$$

for all $c = \left(c^1, ..., c^{k+m}\right) \in \mathcal{C}^{k+m}$. We endow $\mathcal{C}$ with the sup norm.

Suppose that the regulator can estimate, not necessarily exactly, the cost functions of industries 1 through k and call $\widehat{c}^i$ those estimates. As before, the regulator will ask firms in industries $k+1$ through $k+m$ to report their cost structures. For each profile of announcements,

$$C = \left(C^{k+1}, ..., C^{k+m}\right) = \left(\underbrace{c^{k+1}, ..., c^{k+n^{k+1}}}_{\text{industry } k+1}, ..., c^{k+n^{k+1}+n^{k+2}+...+n^{k+m}}\right)$$

let

$$\widehat{x}^1 = \min\left\{f_1\left(\widehat{c}^1, ..., \widehat{c}^k, c^{p_1}, ..., c^{p_m}\right) : p_i \in I^{k+i}, i = 1, ..., m\right\}$$

and similarly for industries $2, ..., k$. As before, define

$$x_j^{k+1} = \min\left\{f_{k+1}\left(\widehat{c}_1, ..., \widehat{c}_k, c^j, ..., c^{p_m}\right) : p_i \in I_{k+i}, i = 2, ..., m\right\}$$

and similarly for industries $k+2$ through $k+m$. The definitions of $\overline{x}^i$ and $\underline{x}^i$ are as before, from equation (3).

Consider the following mechanism:

1. The regulator estimates a cost function $\widehat{c}^i$ for firms in industries $i = 1, ..., k$.

2. Firms in industries $k+1$ through $k+m$ announce their types

3. If in industry $i = k+1, ..., k+m$ announcements coincide, the regulator samples randomly one of the firms and inspects it. If the announcements do not all coincide, the regulator: identifies the firms, or firm, which announced the cost functions which are consistent with $\overline{x}^i$; randomly selects one of them and inspects this firm with probability $\pi > \left(n^i - 1\right)/n^i$, and some other firm with probability $1 - \pi$. A firm is fined if and only if: it is sampled; its report is false; the inspection discovers (with probability ε) that the report was false..

4. The emissions standards $\left(\widehat{x}^1, ..., \widehat{x}^k, \underline{x}^{k+1}, ..., \underline{x}^{k+m}\right)$ are implemented.

Theorem 2. *For any estimates $\left(\widehat{c}^1, ..., \widehat{c}^k\right)$ the unique equilibrium of the direct revelation mechanism, is truth telling. Moreover, the standards $\left(\widehat{x}^1, ..., \widehat{x}^k, \underline{x}^{k+1}, ..., \underline{x}^{k+m}\right)$ are continuous in $\left(\widehat{c}^1, ..., \widehat{c}^k\right)$ so that if the estimated $\left(\widehat{c}^1, ..., \widehat{c}^k\right)$ are close to the truth, the standards in all industries will be close to the first best standards.*

Proof. The proof that the unique equilibrium is truth telling mirrors exactly the proof of Theorem 1, and is therefore omitted.

Continuity of the standards follows from applying Berge's Maximum Theorem (see Aliprantis and Border (1999), p. 539) to $F(c)$ in equation (4): when $\left(c^{k+1}, ..., c^{k+m}\right)$ are fixed in their true levels,

$$D\left(\sum n^i x^i\right) + \sum_1^k \widehat{c}^i\left(x^i\right) + \sum_{k+1}^{k+m} n^i c^i\left(x^i\right) \tag{5}$$

is a function of $\left(\widehat{c}^1, ..., \widehat{c}^k\right)$ and $x = \left(x^1, ..., x^{k+m}\right)$. Then, the set $x\left(\widehat{c}\right)$ of minimizers of (5) is upper hemicontinuous, and therefore continuous, as was to be shown. ∎

5.2 Unknown Damages

In this section we consider two extensions to our basic model that address the question of whether our mechanism works when either D is unknown, or irrelevant.

Suppose first that the regulator is able to estimate D. Then, as in the previous section, we have that the mechanism works, and that if the estimate of D is accurate, the emissions standards will be close to the complete information ones.

Theorem 3. *For any estimate $\widehat{D}$ the unique equilibrium of the direct revelation mechanism of Section 4, is truth telling. Moreover, the standards are continuous in $\widehat{D}$ so that if the estimated $\widehat{D}$ is close to the truth, the standards in all industries will be close to the first best standards.*

The proof of Theorem 3 is similar to that of Theorem 2, and is therefore omitted.

Another extension of the model that is relevant is one in which total damages to society are irrelevant. Consider the case of a country that wants to achieve a certain level of pollution $\overline{X}$ in the most efficient way. This could be the case, for example, of countries that adopted the Kyoto Protocol: they have committed to achieving by 2012 a certain level of emissions. Europe, for instance, must abate its 1990 levels of green house gases by 8%. The problem of the regulator is therefore to find the standards for each industry that minimize the total costs of abatement, and that achieve the desired level of emissions. Formally, suppose that the Kyoto standard is $\overline{X}$, and let

$$\Gamma\left(\overline{X}\right) = \left\{\left(x^1, ..., x^m\right) : \sum n^i x^i \leq \overline{X}\right\}.$$

Then, the regulator wants to implement a selection f from

$$F(c) = \arg \min_{\left(x^1,...,x^m\right) \in \Gamma\left(\overline{X}\right)} \sum n^i c^i \left(x^i\right).$$

We have that our mechanism still implements truth telling, and this results in the complete information standards for this problem.

Theorem 4. *For any $\overline{X}$ the unique equilibrium of the direct revelation mechanism of Section 4, is truth telling.*

The proof is identical to that of Theorem 1, and is therefore omitted.

6 On the Novelty of Our Theorems

We believe that the main merit of our results is their applicability given the simplicity of the mechanism and of the proof, which makes it "likely" that players will understand their incentives.[12]

[12] We thank Matt Jackson for many of the references in this Section, and for his comments regarding the importance of the simplicity of the mechanism and the proof.

In particular, we do not use some of the standard techniques, like cross reporting, used in the literature on implementation with complete information. Nevertheless, in this section we argue that our results are new, and discuss the relationship with the literature on mechanism design.

First, our results do not follow from any of the existing theorems in the literature. That is, there is no theorem that ensures that the social choice correspondence defined by equation 1, or any selection from it, is fully implementable in Nash equilibrium. The results in Jackson, Palfrey and Srivastava (1994) do not apply either to our mechanism, or to the simpler version in which there is only one industry and two firms. Most importantly, their theorems are for implementation in undominated Nash, and our results are full Nash implementation (we get uniqueness without requiring that the strategies be undominated). Moreover, their Theorem 1 is for three or more firms, and their Theorem 3 requires the existence of a "worse outcome" that is not present in our setup.[13]

Second, although inspections and fines have been used in the past and it is "known" that they help in the implementation problem, our assumptions are weaker and different than the ones that have been used before. For example, the important works of Mookherjee and P'ng (1989) and Ortuño-Ortin and Roemer (1993) used costly but perfectly informative inspections and sizeable fines. Our inspections can be as uninformative as one wants, and the fines can be arbitrarily small. Arya and Glover (2005) use a public signal that may be only slightly correlated with the player's reports to implement truth telling (to the owner of a firm) by a manager and his auditor. In their model, however, fines for lying can be large.

Finally, our results are not subject to the criticisms to full implementation in complete information that have been raised by Chung and Ely (2003), since our setup is, in their terminology, one of "private values".

7 Summary

We have presented a mechanism that may help in solving the important problem of how to get polluters to tell the truth about their abatement costs. Our solution is simple, shares some features of how the actual regulatory process works in the US and other places, it implements truth telling by firms and the efficient level of pollution. Also, we have argued that one of the reasons why one does not observe in practice alternative mechanisms that have been proposed in the literature is because they were complicated and relied on taxes and subsidies, which may be too difficult to implement for regulators.

Our main assumption is that there are at least two firms in each industry. We have argued

[13] A worse outcome in that setting would be a standard of 0 and for each firm a lottery which yields the fine with probabilty ε. We do not need to include such an outcome in our space of allocations for our mechanism to work. Our mechanism inspects only one firm.

that this is a reasonable assumption, and we have shown how our mechanism can still be used even when that assumption is not satisfied.

8 Appendix

Proof of Theorem 1. Truth Telling is an Equilibrium. We first show that truth telling is an equilibrium. Without loss of generality, consider the situation of firm 1 when all other firms in all industries are reporting the true costs $\left(c^1, c^2, ..., c^m\right)$. Notice that declaring the true c^1 leads to the implementation of $x_2^1 = ... = x_{n^1}^1$ consistent with all the declarations of firms 2 through n^1. If firm 1 reports $\widehat{c}^1 \neq c^1$, three things could happen, depending on the profile of types announced by industries $2, ..., m$:

- $x_1^1 \geq x_j^1$ for all $j = 2, ..., n^j$. In this case the same standard is implemented in industry 1, and the firm could be fined.
- $x_1^1 < x_j^1$ for all $j = 2, ..., n^j$. In this case, a harsher standard is implemented for industry 1.

Since, no matter what is the profile of types announced in the other industries, firm 1 is worse off deviating, and hence, declaring the truth is better than declaring anything else, proving that truth telling is an equilibrium.

There is no other equilibrium. Suppose there is a profile of strategies $s = \left(s_1^1, ..., s_{n^1}^1, ..., s_{n^1+...+n^m}^m\right)$ such that for some industry i and firm j, $s_j^i\left(c_{lie}^i\right) \neq c_{lie}^i$ for some $c_{lie}^i \in \mathcal{C}$ in the support of P^i (the probability distribution over industry i's types induced by P) and suppose it is an equilibrium. That is, suppose there is an equilibrium without truth telling. Without loss of generality, suppose $i = 1$. Then, for the state $c = \left(c_{lie}^1, ..., c^m\right)$ the profile of announcements $C = \left(C^1, C^2, ..., C^m\right)$ (see equation (2)) is such that not all firms in industry 1 are telling the truth. Notice that *all* announcement in C^1 are lies, since if one firm were telling the truth all firms would be strictly better off telling the truth, since (relative to lying) they would weakly increase the standard for the industry, and strictly reduce the chance of being fined (to 0). Since one firm is inspected the average chance of a firm being inspected is $1/n^1$. Take any firm that in state c_{lie}^1 has a chance (depending on other firms' declarations) of being inspected which is (weakly) larger than $1/n^1$. Suppose it is firm 1.

The standards claimed by firms $2, ..., n^1$ in industry 1, $\left\{x_2^1, x_3^1, ..., x_{n^1}^1\right\}$ depend on: the strategies s_j^i of the firms j in other industries i and on their types c^i (true cost functions). If, given s_j^i for firms j in other industries i, firm 1 in industry 1 can ensure that for "almost all" of the types of the other industries x_1^1 will be strictly (but slightly) smaller than $\min\left\{x_2^1, x_3^1, ..., x_{n^1}^1\right\}$, then the probability of inspection when c_{lie}^1 happens will be very close to $(1-\pi)/\left(n^1-1\right)$. That way, it reduces the standard only slightly, but is sampled with a probability close to

$$\frac{1-\pi}{n^1-1} < \frac{1}{n^1}$$

which corresponds to the probability of being inspected for a firm that claimed a standard different from the maximum. Since firm 1 was lying, it will be strictly better off with that deviation (by strictly reducing the chance of being inspected and fined). To establish the existence of such a deviation, notice that if the firms in industry 1 were declaring cost functions $\{\widehat{c}_1, \widehat{c}_2, ..., \widehat{c}_{n^1}\}$ firm 1 can always declare a $\widetilde{c}$ defined by

$$\widetilde{c}'(x) = \begin{cases} \max_i \widetilde{c}'_i(x) + \delta & \text{if } \max_i \widetilde{c}'_i(x) + \delta \leq 0 \\ \max_i \widetilde{c}'_i(x) & \text{otherwise} \end{cases}$$

for small δ. Such a $\widetilde{c}$ is in $\mathcal{C}$, so we now show that it yields a strictly higher utility for firm 1. For the fixed c^1_{lie}, let $T(\delta)$ be the set of types of the other industries, $T(\delta) \subseteq \mathcal{C}^{m-1}$ for which the (proposed) equilibrium declarations of all firms in all industries would yield a standard $\underline{x}^1$ such that $\max_i \widetilde{c}'_i\left(\underline{x}^1\right) + \delta \leq 0$. Since $c'_i \leq 0$ for all c_i, for small enough δ, the set $T(\delta)$ has probability close to 1. We will now show that for any such δ, for any type in $T(\delta)$ the profile of declarations of cost functions of firms in the other industries $C^{-1} = \left(C^2, ..., C^m\right)$ induces a smaller standard for industry 1, and that standard corresponds to the declaration $\widetilde{c}_1$ of firm 1. Hence for this type of the other industries the probability of inspection for firm 1 is $\pi / \left(n^1 - 1\right)$. Since the type was arbitrary in a set of measure close to 1 (for the small δ that we chose), the overall probability of inspection will be close to $(1-\pi) / \left(n^1 - 1\right)$.

The assumptions that D and all cs are differentiable, and $D'(x) + c'(0) < 0$ for all c ensure that the solution of the regulator's problem is interior, and hence the first order condition of the problem of the regulator is satisfied. Now fix any type c for which $c^1 = c^1_{lie}$, and given the strategies of firms in other industries, fix the selection $\left(c^2, ..., c^m\right)$ from $\left(C^2, ..., C^m\right)$ (one firm's declaration per industry) and the firm j in industry 1 such that $\underline{x}^1 = x^1_j$. When choosing x^1_j the regulator solved:

$$0 = D'\left(\sum_{i \neq j} n^i x^i + n^1 x^1_j\right) n^1 + n^1 \widetilde{c}'_j\left(x^1_j\right) < D'\left(\sum_{i \neq j} n^i x^i + n^1 x^1_j\right) n^1 + n^1 \widetilde{c}'\left(x^1_j\right).$$

Convexity of D and $\widetilde{c}$ imply that for this combination of declarations in the other industries, the standard set by firm 1 will be lower. Moreover, by the Theorem of the Maximum, and uniqueness of the optimal x (a consequence of the strict convexity of the cs) the new standard corresponding to $\widetilde{c}$ will be close to $\min\left\{x^1_2, x^1_3, ..., x^1_{n^1}\right\}$.

We have proved that if c^1_{lie} occurs, firm 1 is strictly better off deviating. Since strategies are continuous, the same analysis can be conducted for types c^1 close to c^1_{lie}. Since c^1_{lie} is in the support of the distribution of types for industry 1, there is a positive probability of types for which firm 1 is strictly better deviating from the proposed equilibrium, and that is a contradiction. ∎

References

[1] Aliprantis, C. and K. Border, (1999), *Infinite Dimensional Analysis*, Springer: Heidelberg.

[2] Arya, A. and J. Glover, (2005), "Maintaining Auditor Independence," mimeo.

[3] Böhm, P. and C. S. Russell, (1985), "Comparative analysis of alternative policy instruments," in *Handbook of Natural Resource and Energy Economics*, Vol. I, A.V. Kneese and J.L. Sweeney. (Eds.), Elsevier.

[4] Chung, K. and J. Ely, (2003), "Implementation with Near Complete Information," *Econometrica*, 71, 3, 857-871.

[5] Dasgupta, P., P. Hammond and E. Maskin, (1980), "On Imperfect Information and Optimal Pollution Control," *Review of Economic Studies*, **47** (5), 857-60.

[6] EPA (1992), "Model Pollution Prevention Plan for the Kraft Segment of the Pulp and Paper Industry," NTIS # PB93-145605.

[7] EPA (2002), "Development Document for Final Effluent Limitations Guidelines and Standards for the Iron and Steel Manufacturing Point Source Category," EPA-821-R-02-004, at http://www.epa.gov/ost/ironsteel/tdd.htm

[8] Field, B. C., (1997), *Environmental Economics, an Introduction,* Irwin/McGraw-Hill: Boston.

[9] Jackson, M. O. (1992): "Implementation in Undominated Strategies: A Look at Bounded Mechanisms," Review of Economic Studies, 59, 757–775.

[10] Jackson, M. O., T. R. Palfrey, and S. Srivastava (1994): "Undominated Nash Implementation in Bounded Mechanisms," Games and Economic Behavior, 6, 474–501.

[11] Jackson, M. O., and S. Srivastava (1996): "Characterizations of Game Theoretic Solution Concepts which Lead to Impossibility Theorems," Review of Economic Studies, 63(1), 23–38.

[12] Keohane, N. O., R. L. Revesz and R. N. Stavins, (1998), "The Choice of Regulatory Instruments in Environmental Policy," *Harvard Environmental Law Review*, 22 (2), 313-367.

[13] Kwerel, E. (1977), "To Tell the Truth: Imperfect Information and Optimal Pollution Control," *Review of Economic Studies*, **44** (3), 595-601.

[14] Lewis, T. (1996), "Protecting the Environment when Costs and Benefits are Privately Known," *RAND Journal of Economics*, **27**, 819-47.

[15] Mookherjee, D. and I. Png, (1989), "Optimal Auditing, Insurance, and Redistribution," *the Quarterly Journal of Economics,* Vol. 104, 2, pp. 399-415.

[16] Moore, J. (1992). Implementation, contracts, and renegotiation in environments with complete information. In *Advances in Economic Theory: Sixth World Congress*, vol. I, edited by J.-J. Laffont. Cambridge, UK: Cambridge University Press.

[17] Ortuño-Ortin, I. and J. Roemer (1993), "Politically realistic implementation with inspection: The equity-honesty-welfare trade-off," *Economics and Politics* 5, 255-270.

[18] Palfrey, T. R., and S. Srivastava (1991): "Nash Implementation Using Undominated Strategies," *Econometirca*, 59, 479–501.

[19] Russell, C. S. and P. T. Powell, (1996), "Choosing Environmental Policy Tools, Theoretical Cautions and Practical Considerations," IADB, Washington D.C., - No. ENV-102.