Information Transmission as a Rationale for Voluntary Agreements

(draft)

Patrick González[1]

May 11, 2010

[1]Département d'économique, Université Laval.

Abstract

I consider an industry where production generates pollution, but the marginal damage differs among firms. If the Regulator has no information about the firms' types, he taxes clean and dirty firms alike with the same uniform pigovian tax. Firms may participate to a voluntary agreement where their type is revealed to the Regulator by the mean of a perfect costly audit. In equilibrium, participation entails a positive informational externality as the dispersion of types decreases within the pool of non participating firms. Participating firms benefit from the elimination of the distortion induced by uniform taxation. That distortion is endogenous since the uniform tax rate rises as the pool of non participating firms is reduced. I show that social welfare increases if the audit cost is not too high, in which case there is a rationale for the Regulator to advocate such agreement. A key result is that the efficiency of a voluntary agreement can't be inferred from the participation rate at the margin.

1 Introduction

The economic analysis of industrial regulation contrasts the public interest pursued by the State (hereafter, the "Regulator") with the firms' private interests. Industrial pollution is an important example where the firms' production entail a social cost that they neglect. Because he has coercive powers, the Regulator can maintain an efficient level of pollution by constraining the firms to reduce their polluting emissions. Firms comply to avoid costly penalties or downright closure of their plants.

Yet, the Regulator will often propose the firms to reduce their emissions even if non compliance involves no penalty. Such intervention takes the form of a "voluntary agreement" where the industry and the Regulator accept in principle to address a problem involving some external cost or benefit. Not all firms choose to participate so that, in the end, similar firms are subjected to different constraints. To the extent that the Regulator's intervention involves constraints imposed on firms, the phenomenon of voluntary agreements looks paradoxical. Why should participating firms accept constraints that reduce their profits if they can avoid them? And why the Regulator would renounce to use its coercive powers in the first place?

The industry may voluntary accept constraints today to avoid stringent constraints tomorrow. They may fear that current non compliance could lead the Regulator to impose penalties later (Segerson and Miceli 1998), or lead their concerned (green) customers to adopt a similar stance (Arora and Gangopadhyay 1995, Arora and Cason 1996). The Regulator may prefer the voluntary approach to traditional regulation if it economizes on transaction costs (Lévêque 1999, Segerson and Miceli 1999).

Transaction costs account for any factor likely to impact the efficiency of the regulatory process. For example, in Nyborg (2000), the Regulator control only some forms of pollution by firms: a voluntary agreement with the industry dominates traditional regulation if it allows a better tradeoff among the various sources of pollution. David (2005) considers a similar case where there is an arbitrage between a direct or an indirect reduction of polluting emissions (indirect reduction being the result of an inefficient reduction of production). Schmelzer (1999) analyzes a case where players in the industry can observe the emissions of firms at a lower cost than the Regulator. The Regulator can then economize on monitoring cost by delegating this function to the industry within a voluntary agreement.

In all these cases, the industry advocates a form of self-regulation which allows a lighter regulation and a better allocation of resources. But, it is not clear that the industry can regulate itself at a lower cost. Contrary to the Regulator, the industry must reconcile the conflictive incentives of its members. These may result in an under provision of the public good (the reduction of emission). Ashby and als (2004) illustrate a few cases where these incentives are so contradictory that self-regulation is impossible: each firm has an incentive to rely on the others to implement the voluntary agreement. Segerson and Dawson (2008) present a case where internal discipline within the industry is maintained only because each participating firm is pivotal and believes that if it reneges on the voluntary agreement, the Regulator will react boldly and that will affect the whole industry.

In this paper, I develop the idea that firms participate to a voluntary agreement not to ensure a lighter regulation for the industry but for themselves individually. My approach combines transaction costs and heterogeneity among firms under incomplete information. I show that in these conditions there may be a rationale for the Regulator to implement a voluntary agreement with firms in the industry.

Firms differ by their impact on the environment, but the Regulator can't initially perceive these differences. To reduce the emissions, the Regulator imposes a uniform tax upon all firms whether they pollute a lot or not. Hence, the cleanest firms support a disproportionate share of the tax burden. In this paper, participation to a voluntary agreement means paying for a perfect audit that informs the Regulator about one's type. To avoid the impact of an undiscriminating regulation, clean firms have an incentive to join a voluntary agreement and inform the Regulator about their type. Once informed, the Regulator is willing to discriminate the tax rate.

Participation to a voluntary agreement entails an informational externality as the statistical properties of the pool of non participating firms change. In equilibrium, the cleanest firms participate so that the dispersion of types within the pool of non participating firms decreases and the default uniform tax rate rises.

Arora and Gangopadhyay (1995) were the firsts to explain the participation of firms to a voluntary agreement by analyzing the effect of this participation on the regulation applied to the whole industry. Maxwell, Lyons and Hackett (2000) as well as Lyons and Maxwell (2003) have developed an extreme version of this argument where the firms manage to preempt any intervention of the Regulator by adopting slack voluntary measures. More recently, Denicolò (2008) has shown how the early adoption by a firm of a clean but costly technology could be a signal sent to the Regulator to motivate him to impose its adoption to the whole industry.

In Denicolò, participation to a voluntary agreement involves the transmission of *soft* information to the Regulator (of a deductive nature). A firm participates to improve its future position with respect to its competitor. The Regulator does not know initially whether a regulatory intervention is needed and learns it very indirectly looking at the choice of technology of some firms. In my model, by comparison, participation involves the transmission of *hard* information obtained through a perfect costly audit. Firms are perfectly competitive and have no strategic motive to participate besides obtaining a lesser tax rate.

The popularity of voluntary agreements has faltered over the years. From the start, people wondered why participating firms would voluntary impose themselves constraints that reduce their profits. In support of this view, some voluntary agreements did result in little gain with respect to the business as usual scenario if we take into account the human resources required to make them run. My analysis indeed shows that not all voluntary agreements enhance social welfare. This is because firms participate not only to avoid the uniform tax distortion effect but also its distributive effect (from the firm to the Regulator). The Regulator is insensitive to the latter effect so that there may be too much participation. Specifically, I show that, through participation, infra marginal firms have always a lesser impact on welfare than the marginal firm who is indifferent between participating or not. The participation of these infra marginal firms may decrease expected welfare below the level reached without a voluntary agreement. With a numerical example, I show that this happens when the audit cost is too high. In that sense, a voluntary agreement is a good regulatory instrument as long as it does not cost too much. The difficult part is to evaluate this endogenous cost.

The mathematical model is presented in the next section. I set up an incomplete information environment where the firms' private types is the (constant) marginal external impact of their production. I show that second best taxation under incomplete information takes the form of a uniform rate if the firms' types are independent of their other observable attributes. I then show that the recourse to such taxation implies implicitly that the firms can't easily certify their types. The key step is then to assume that this certification can only be done within the specific setting of a voluntary agreement. The most interesting insights arise under partial participation. For that to happen, participation must be costly. This line of reasoning is pursued in section 3 with a dynamic version of the model.

In both following sections, participation to a voluntary agreement generates an informational externality as it allows the Regulator to update his beliefs about the distribution of types among non participating firms. In the first section, this externality has a far reaching effect since the information problem disappears in equilibrium. This is why I add an audit cost in the second section. Yet, in both sections, this externality is "marginal" in the sense that the participation of a small group of firms has a continuous effect on the updating process. This formalization is done for tractability, but I suspect that it does not represent well what is really going on in many real life instances of voluntary agreements. I discuss this issue in the conclusion.

I have tried to emphasize throughout the underlying logic the model. Specifically, the analysis is done using the common apparatus of profit functions. In an appendix though, one will find a running numerical example fashionable in the industrial organization literature. Yet, that numerical example is important because I rely on it to provide instances where a voluntary agreement increases social welfare and instances where it does not.

2 First and Second best Regulation

I model a competitive single-good market with a perfectly elastic demand and a population of firms with strictly positive, increasing and continuously differentiable supply functions for all positive prices. Let s denote a firm's supply. Its profit function π is strictly increasing and strictly convex and, by Hotelling's lemma, it has s for derivative.

Production entails an external damage cost proportional to supply by a factor d. This factor differs from firm to firm and is distributed in the industry over an interval $D = [d_0, d_1]$, with positive measure over any subinterval. To simplify the analysis, I assume that $p > d_1$ so that all firms have a non zero socially efficient level of production. Let $\hat{\delta}$ and $\check{\delta}$ denote respectively the expected value of d conditional on d being *no less* (in the first case) and *no greater* (in the second) than δ. Hence $\mathrm{E}(d) = \hat{d}_0 = \check{d}_1$. I assume that these functions are increasing and continuous[1]. There are positive measures of firms below and above the mean so that

$$d_0 < \hat{d}_0 < d_1$$

The factor d is distributed independently of any of the firm other attributes. As a consequence, if $s(p)$ now denotes the industry *average* supply, then $\hat{d}_0 s(p)$, denotes the average damage. When the Regulator imposes a unit tax t, the average supply decreases to $s(p-t)$. Average profit decreases to $\pi(p-t)$ but part of this reduction is the tax revenue $T = ts(p-t)$ that has no bearing on social surplus. The social value of production[2] thus amounts to

$$\pi(p-t) - (d-t)s(p-t)$$

The optimal tax t^* maximizes this expression. The first-order condition[3] yields

$$(d-t^*)s'(p-t^*) = 0$$

so that $t^* = d$ whenever $p > d$. Otherwise, one sets $t^* = p$ and the firm is basically shut down. The pigovian tax fully internalizes the damage into the firm's profit. Let

$$S^* = \mathrm{E}(\pi(p-d)) > 0$$

denotes the first best average social value of production.

The profit function π is common knowledge but d is private information to the firm. To implement t^*, d must be known. Without this information, the Regulator may set a

[1] In the conclusion, I show that these are actually very strong economic assumptions.

[2] To simplify the analysis, I discard consumer surplus. We may assume that all production is exported abroad.

[3] The second-order condition is $s'(p-t^*) > 0$ and is always satisfied.

uniform tax t that maximizes the expected social value of production

$$S(t) = \mathrm{E}(\pi(p-t) - (d-t)s(p-t))$$

The laissez-faire solution is the special case where $t = 0$ and welfare amounts to $S(0)$. The second-best tax t^{**} solves the first-order condition

$$(\hat{d}_0 - t^{**})s'(p - t^{**}) = 0$$

so that t^{**} is equal to the *average* value of d. With such tax, each firm pays the expected damages of its production

$$T^{**} = t^{**}s(p - t^{**}) = \mathrm{E}(ds(p - t^{**}))$$

and the expected social value of production reaches

$$S(t^{**}) = \pi(p - \hat{d}_0)$$

Since $t = t^{**}$ is socially preferred to $t = 0$, welfare increases: $S(t^{**}) > S(0)$. Jensen's inequality implies that

$$\mathrm{E}(\pi(p-d)) > \pi(p - \hat{d}_0)$$
$$S^* > S(t^{**})$$

To resume $S^* > S(t^{**}) > S(0)$. At the firm level, though,

$$d_0 < \hat{d}_0 < d_1$$

or equivalently

$$t^*(d_0) < t^{**} < t^*(d_1)$$

The second-best tax is higher (lower) than the first-best tax for clean (dirty) firms. As a consequence, clean (dirty) firms produce less (more) and make less (more) profit with t^{**} than they should

$$s(p - d_0) > s(p - t^{**}) > s(p - d_1)$$
$$\pi(p - d_0) > \pi(p - t^{**}) > \pi(p - d_1)$$

The second-best solution improves social welfare with comparison to the laissez-faire solution but it is less efficient than the first-best solution because it unduly constrains the production of relatively clean firms and it allows dirty firms to produce more than they should.

2.1 The First Best for Free

Although the Regulator lacks information about d, this information is available from the firms themselves, although not all firms have an incentive to provide it: those for which $d > t^{**}$ would pay a higher rate with an optimal tax. But there is an easy way out for the Regulator which is to privatize the public bad so that all relevant firms have every incentive to prove that it does not belong to them.

Suppose the Regulator announces the imposition of a second-best tax

$$t_1 = t^{**} = \hat{d}_0$$

All firms for which $d < \hat{d}_0$ then have an incentive to inform the Regulator that such tax would unduly destroy social welfare in their case. These private incentives generate an informational externality. If the left portion of the distribution of firms come forward, and ask for a discriminating optimal tax, the Regulator will learn that the distribution of d within the remaining population is bounded below by $\hat{d}_0$ and he will set a new uniform rate

$$t_2 = \hat{t}_1 = \hat{\hat{d}}_0$$

for those firms. Since the conditional expectation of d within this population is higher than $\hat{d}_0$, that uniform rate will be higher, $t_2 > t_1$, and all firms for which $t_1 \leq d < t_2$, that stayed silent in the first round will now come forward, show compelling evidence about their type and also ask for a discriminating rate. Again, the conditional expectation of d in the remaining population will shift to the right and the Regulator will revise its rate to

$$t_3 = \hat{t}_2 = \hat{\hat{\hat{d}}}_0$$

Again, $t_3 > t_2 > t_1$. This run to revelation stops when the tax rate reaches $t = d_1$, since $d_1 = \hat{d}_1$. At this point, all firms have willingly revealed their types.

Hence, there exists a Bayesian equilibrium where the Regulator manages to implement the first-best allocation under incomplete information. The two following assumptions are needed for this result.

Assumption A. *The Regulator can constrain the firms' production choices.*

Assumption B. *Firms can freely convey hard information about their type.*

Without assumption A, the Regulator is pretty much emasculated. Without assumption B, the Regulator would have to rely on soft (non verifiable) information of dubious quality since presumably all firms would state that they have the lowest possible d in order to get the lowest possible tax.

Both assumptions make sense in the long run since the Regulator can channel vast powers and information diffusion is an increasing return technology[4]. In that sense, the regulation problem is acute in the short run when either one or both of these assumptions do not hold.[5]

The literature on regulation with incomplete information[6] relaxes assumption B. In the short run, it is plausible that the regulator does not have all the information needed to implement a first-best allocation. This literature provides a rationale for incentives schemes that allow firms to produce inefficiently and to gather costly social funds. When incomplete information is taken into account, these seemingly inefficient schemes are shown to realize a delicate arbitrage between information acquisition by the regulator and production efficiency.

The literature on voluntary agreements relaxes assumption A. Again, it is plausible that the regulator can't constrain the firms' production choices in the short run. The Regulator is one among many institutional players. Constitutions usually endow the State with unrestricted powers in the short run but only in critical instances like war. In the normal course of business, fundamental institutions like private property, counterbalance many of the Regulator powers. Even if the Regulator could proceed with a quick nationalization, this is no guarantee that he will have a firm grip on all the interval levers of a big firm, like many initiators of hostile takeovers have experienced in the past. To regulate a firm in the short run, cooperation is required.

In the long run, both conditions may approximately hold but they reflect very different

[4]The more people know something, the easier for someone else to learn it.

[5]Regulation is not "free" in the long run to the extent that new information asymmetries continuously emerge but this involves the issue of moral hazard which I do not analyze here.

[6]See Laffont-Tirole.

features of the world and there is no reason to assume that both will begin to hold at the same time. In the next section, I shall assume that assumption A holds much sooner than assumption B. The Regulator has a direct interest in providing itself with the regulatory tools it needs to regulate an industry. The firms, on the other hand, have only an indirect interest in providing signals about themselves through, for instance, costly audits.

3 A Rationale For Voluntary Agreements

Consider thus a two periods horizon with incomplete information about the firms' types. Denote $0 < \beta < 1$ the discount factor that weights the first and second period profit: a firm maximizing profits at price p_1 in the first period and price p_2 in the second period earns

$$(1 - \beta)\pi(p_1) + \beta\pi(p_2)$$

This division represents the time needed by the Regulator to build its regulatory powers. Assumption A does not hold in the first period but does so in the second period. The discount factor β thus stands for the readiness of the Regulator to exercise its powers: if $\beta \to 0$, the Regulator has no power and if $\beta \to 1$, it can enforce any regulation even in the very short run.

In concordance with the incomplete information literature, assumption B does not hold over the whole horizon which is to say that signaling one's type involves costs. I shall assume that these costs are prohibitive in comparison to what will be possible within a voluntary agreement (see below).

Without a voluntary agreement, the firms disregard the external effects of their production in the first period and the Regulator imposes the second-best pigovian tax, as described in section 2, in the second period. Expected social surplus then amounts to

$$S_\beta^{**} = (1-\beta)S(0) + \beta S(t^{**})$$

which lies between $S(0)$ and $S(t^{**})$, so that the Regulator intervention improves social welfare but not as much as if it could be pursued right away.

To assess an institution like a "voluntary agreement" we must define a category of what it is and what it is for. A voluntary agreement is more than just a slogan. Parties who engage in a voluntary agreement devote human resources to such enterprise. The people involved are doing something. My main thesis is that a voluntary agreement is an institutional arrangement through which there is information transmission in the short-run. Civil servants learn about the participating firms technology.

I consider a very simple information transmission mechanism: a perfect costly audit. An audit costs $c > 0$ (a dead-weight loss) per unit produced and informs the Regulator about a firm's type. As above, a participating firm gets a discriminating rate $t^*(d)$ in the second period instead of a uniform rate t.

Participation

A firm participates if the long run benefit $\pi(p-d) - \pi(p-t)$ of getting a lower tax rate covers the short run loss $\pi(p) - \pi(p-c)$ due to the audit cost (properly discounted). The gain to participate is denoted

$$V(p,d,t) = \beta(\pi(p-d) - \pi(p-t)) - (1-\beta)(\pi(p) - \pi(p-c))$$

Notice that V is convex, decreasing in d and increasing in t. A firm that expects a second period uniform rate t participates if $V(p,d,t) \geq 0$. Given t, the Regulator infers that

$V(p, d, t) < 0$ for the pool of all non participating firms and it revise t accordingly to the average d within that pool. To maintain a simple information structure, I assume thereafter that all profit functions are identical up to an affine transformation[7].

Since V decreases with d, the cleanest firms participate first and the second best period tax rate increases as participation increases. We have an equilibrium tax rate when no such revision take place[8].

Let $D^{**} = [t^{**}, d_1]$. To have information transmission, the audit cost must not be too high so that the cleanest firm will wish to participate. Hence, I assume that $V(p, d_0, t^{**}) > 0$. Since V increases with t, this inequality holds as well for any $t \in D^{**}$. At the other end of D, we can be sure that the dirtiest firm will never want to pay for an audit since the uniform rate is bounded above by its type: it can't expect to pay a lesser tax by revealing its type. This implies $V(p, d_1, t) < 0$. To resume,

$$V(p, d_0, t) > 0 > V(p, d_1, t) \quad \text{for all } t \in D^{**} \tag{1}$$

By continuity, for all $t \in D^{**}$, there exists a marginal firm $\delta(t)$ for which $V(p, \delta(t), t) = 0$. Using the implicit function theorem on this identity, we can establish that δ increases

[7] In the previous section, we had $c = 0$ so that the participation condition resumed to $d \leq t$. Hence, participation depended on d but not on π. Consequently, participation did not affect the statistical independence property between d and π within the pools of participating and non participating firms. This is not so anymore so that, theoretically, participation could result in some correlation between observable profits (or supply) and d. In particular, it would be wrong to assume that a uniform tax, independent of π would still yield the second best welfare within the pool of non participating firms and there would be a scope for more sophisticated scheme than a single hard audit to identify the participating firms.

Now if all profit functions are identical to some function π up to an affine transformation $\alpha + \gamma\pi$, then they will share the same participation function V, up to a linear constant, and their decisions to participate will still differ only if their factors d differ.

[8] Alternately, we could suppose that the Regulator commits in the first period to apply a rate t_e in the second period. This would not change much the analysis as the equilibrium rate is one among many that the Regulator could choose but it would complicate the mathematics tremendously because of the informational externality. I thus implicitly assume that the Regulator does not have this possibility. In the conclusion, I argue that the present formalization makes more economic sense.

strictly with t. A firm with $d > \delta(t)$ does not participate. For the marginal firm

$$\beta\left[\pi(p-\delta(t))-\pi(p-t)\right]=(1-\beta)\left[\pi(p)-\pi(p-c)\right]>0$$

Increasing marginally d, we find a bunch of non participating firms for which $\pi(p-d) > \pi(p-t)$; that is, firms that would benefit in the long run by participating yet that choose to abstain to economize on the audit cost.

Existence of an equilibrium

I now establishes the existence of an equilibrium. In equilibrium, the tax rate t_e induces a participation $\delta(t_e)$ such that the second best uniform rate within the pool of non participating firms is t_e. In short, $t_e = \hat{\delta}(t_e)$. Since V decreases with d, the l.h.s. of condition (1) ensures that $\delta(t^{**}) > d_0$ so that $\hat{\delta}(t^{**}) > \hat{d}_0$. Consider the continuous function $\hat{\delta}(\max\{t, t^{**}\})$ which maps D into D. It has a fixed point t_e. If $t_e = t^{**}$, we get the contradiction $t^{**} = \hat{\delta}(t^{**}) = \hat{d}_0$. It follows that $\hat{\delta}(t_e) = t_e > t^{**}$. For further reference, we denotes $d_e = \delta(t_e)$ the equilibrium marginal firm or measure of participation.

The existence of an equilibrium is depicted in Figure 1. The space is D^2. The horizontal axis represents the type d and the vertical axis, the tax rate t. Two strictly increasing continuous functions are drawn. The function $\hat{d}$ maps each d into a revised tax rate. In particular, $\hat{d}_0 = t^{**}$ and $\hat{d}_1 = d_1$. The function δ maps t into a marginal firm $\delta(t)$ that bounds above the set of participating firms. It is undefined around d_0 since no firm would agree to pay the audit cost c to avoid such a low rate. It is everywhere below t because only firms with $d < t$ will eventually want to participate. Yet $\delta(t^{**})$ exists and is strictly positive since we have assumed that the audit cost c is low enough so that type d_0 has a strict interest to participate. This implies that δ crosses the t-axis below t^{**}. At the other end of D, $\delta(t) < t$ implies that $\delta(d_1) < d_1$. These two facts imply the existence of an equilibrium point, like in point z, where the two functions cross. There, each function yields the inverse of the other: $\hat{\delta}(t) = t$.

Multiple Equilibria and Stability

As in most equilibrium models, instances of multiple equilibria are possible. Recall that many firms participate because the uniform tax rate rises. We could have a low participation equilibrium where the tax rate remains relatively low given the large dispersion of types within the pool of non participating firms; and a high participation equilibrium where the opposite applies. Such multiple equilibria generically come in odd number as it is the case here. There are three equilibria in x, y and z where x and z are respectively the low and high equilibria just described. Both are stable in the sense that if some firms expect a marginal discrepancy between the actual and equilibrium tax rates, their reaction will lead the Regulator to revise the rate by an amount less than this discrepancy. For instance, if the firms expect t to be $t_e - \Delta_1$, participation will decrease to $\delta(t_e - \Delta_1)$. But this should lead the Regulator to lower t by $\Delta_2 = t_e - \hat{\delta}(t_e - \Delta_1)$ which is less than Δ_1. Iterating on this reasoning, we see that $\Delta_n \to 0$. The equilibrium y is unstable: any discrepancy between the firms' expectations about the tax rate and its equilibrium value will lead to an ever increasing or decreasing revision process until the expectations converge again to a new stable equilibrium.

Welfare Analysis

Does the possibility for firms to transmit their information increases social welfare above S_{β}^{**}? This is so for the non-participating firms since they pay now a higher tax $t_e > t^{**}$ in the long run, closer to the marginal damage of their production. Things are more fuzzy for participating firms.

Notwithstanding its effect on the information structure, participation entails two external effects that matter for the Regulator. A participating firm reduces its production from $s(p)$ to $s(p-c)$ in the short run and that brings an additional social benefit as pollution is proportionally reduced. In the long run, the firm is taxed efficiently and thus one avoids the distortion $(d-t^{**})s(p-t^{**})$ induced by the second best uniform tax rate. Let

$$W(p,d) = \left[\beta s(p-d) - \frac{\partial V}{\partial p}(p,d,t^{**})\right] d - \beta T^{**} \tag{2}$$

resume these external effects[9]. Notice that the bracketed term in (2) is positive and independent of d (see footnote 9) so that W is a positive affine function of d.

The participation of a firm increases social welfare if

$$V(p,d,t) + W(p,d) \geq \beta(\pi(p-t^{**}) - \pi(p-t)) \tag{3}$$

The term on the r.h.s. is an adjustment to take into account that the Regulator considers the benefit of participation with respect to the base scenario, with no voluntary agreement, while each individual firm considers the benefit of participating given that such opportunity exists and that the uniform tax rate is expected to raise up to t.

As a function of d, $V+W$ is the sum of a decreasing convex function and an increasing

[9] We have

$$\begin{aligned} W(p,d) &= (1-\beta)d(s(p)-s(p-c)) + \beta(d-t^{**})s(p-t^{**}) \\ &= \left[\beta s(p-t^{**}) + (1-\beta)(s(p)-s(p-c))\right] d - \beta t^{**} s(p-t^{**}) \end{aligned}$$

where the bracketed term is positive and independent of d. Yet, we can rewrite it as

$$\begin{aligned} &= \left[\beta s(p-d) - \{\beta(s(p-d)-s(p-t^{**})) - (1-\beta)(s(p)-s(p-c))\}\right] d - \beta T^{**} \\ &= \left[\beta s(p-d) - \frac{\partial V}{\partial p}(p,d,t^{**})\right] d - \beta T^{**} \end{aligned}$$

affine one; it is thus convex. It's derivative is

$$\frac{\partial V}{\partial d}(p,d,t) + \frac{\partial W}{\partial d}(p,d,t) = -\frac{\partial V}{\partial p}(p,d,t^{**})$$

Since $\frac{\partial V}{\partial p}(p,d,t^{**}) < 0$, for all $d \in D^{**}$, $V + W$ is increasing at the point $d = d_e$. It follows that even if (3) holds for the marginal firm d_e, it does not necessarily hold for all participating firms $d < d_e$. There are always firms whose participation add less to social welfare than the marginal firm.

In a voluntary agreement, the cleanest firms participate because they have a lot to gain, notably by avoiding the long run distortion $(t_e - d)s(p - t_e)$, and the Regulator recognizes these gains. But the infra marginal firms get in only because the ex post uniform tax rate increases. At the margin, they are indifferent between participating or not, that is paying a unit cost c today or enjoying a tax rebate $t - d_e$ tomorrow. The Regulator does not necessarily share that view since participation involves an audit deadweight loss $cs(p - c)$ altogether different than the long run distortion. The more participation there is, the higher t_e and the closer d_e is to to t_e so that the ex post distortion becomes small with respect to the audit deadweight loss.

Even if the marginal firm's participation improves social welfare, that does not imply that the participation of infra marginal firms is also desirable. Hence the social desirability of a voluntary agreement can't be judged at the margin but globally by taking the expected social return over the contributions of all firms with a voluntary agreement in place. Since I haven't made any specific assumptions about the marginal distribution of d, I can't decide of this issue once and for all. I will need to *assume* that a voluntary agreement is socially desirable.

Let $SV(z)$ denotes expected welfare with a voluntary agreement when the participating firms pay z in audit costs per unit produced. If $z = 0$, all firms will wish to participate as it was shown in section 2 so that

$$SV(0) = (1-\beta)S(0) + \beta S^*$$

The Regulator may wish to subsidize the audit to motivate the firms to participate. If the audit costs is c, he could subsidize $c - z \geq 0$. Social welfare then amount to

$$SR(z, c) = SV(z) - (1-\beta)(c-z)s(p-z)F(d_e)$$

where the second term is the expected subsidy ($F(d_e)$ being the measure of participating firms). Full subsidization ($z = 0$) yields

$$SR(0, c) = SV(0) - (1-\beta)cs(p)$$

Hence, maximum welfare under a voluntary agreement (with subsidies) is

$$SR^*(c) = \max_{z \in [0,c]} SR(z, c)$$

Now recall that the Regulator can achieve S^{**}_β by not promoting a voluntary agreement. To provide a rationale for the phenomenon of voluntary agreements, I need the following participating condition for the Regulator.

Assumption C. *Given the audit cost c, $SR^*(c) > S^{**}_\beta$.*

In the numerical example presented in the appendix, $SR^*(c) = SR(c, c) = SV(c)$ for all c, so that it is always optimal for the Regulator not to subsidize the audit. Whether this ancillary result has some generality beyond this example I cannot say. More to the point, I show in this example that assumption C holds when the audit cost is not too high.

4 Conclusion

There is a rationale for voluntary agreements when assumptions A and B do not hold, but assumption C holds. Clean firms participate because they get a tax rebate; marginal firms participate because the participation of the clean firms makes the uniform rate rise; and, tautologically, the Regulator participates because the whole enterprise increases expected social welfare. I emphasize that assumption C is not trivial because it may not hold. The fact that firms would participate in a voluntary agreement, as defined in this paper, does not ensure that the Regulator should initiate such agreement in the first place.

An implicit assumption of this model is that the audit can take place only in the first period within the very specific setting of a voluntary agreement. The audit *is* the voluntary agreement. The audit stands as a reduced form of the sophisticated information transmission mechanism that goes on within a voluntary agreement. Meetings, exchange of documents, on site controls, etc, by civil servants, and participating firms' employees result in the Regulator being informed, at a cost, about the firms types. Obviously, information about firms can come from other sources and at other points in time but the point is that proceeding early on a voluntary basis makes sense if the transaction costs are low.

I have argued above (see footnote 8) that it is better to assume that the long run uniform tax rate maximizes ex post expected welfare conditionally on the participation rate. To assume this is equivalent to assume that the Regulator can't commit ex ante to a given tax rate. Here, the Regulator can compute (as we do) ex ante the second best rate but that is a misleading feature of the model. In reality, the information structure does not always easily permit such inference and the informational externality may turn out to be much stronger than it is modelled here. Remember that I have assumed that the expectation of d among the pool of non participating firms, conditional on the participation of the cleanest firms, was non decreasing and continuous. These are strong assumptions as the following simple example will demonstrate.

Suppose that d is distributed uniformly over $[0, 1]$ so that $t^{**} = \frac{1}{2}$. If the marginal firms within $[0, \Delta]$ choose to participate, the ex post rate raises marginally by $\frac{\Delta}{2}$. Suppose now that it is common knowledge that the distribution of firms is either $F(d) = d(2 - d)$ or $G(d) = d^2$ with equal probabilities. From an ex ante point of view, firms are distributed according to $\frac{1}{2}F(d) + \frac{1}{2}G(d) = d$ which is the uniform distribution. Since the ex ante distribution is the same, t^{**} will still equal $\frac{1}{2}$, but if the firms within $[0, \Delta]$ participate, the Regulator will immediately learn which of F or G is the true distribution. If it's F, the ex post tax will plunge to $\frac{1}{3} + \frac{2}{3}\Delta$ and if it is G, it will jump to $\frac{2}{3} + \frac{2}{3}\frac{\Delta^2}{\Delta+1}$. Here, the conditional expectation is not continuous as participation increases and it is not even "increasing" or "decreasing" in a meaningful way. The informational externality here has a non marginal (non continuous) effect on the Regulator's beliefs. The point is that it would make little sense for the Regulator to commit ex ante to a second period tax rate. I do not dispute the possibility of committing to a menu of tax rates in some simple cases, but I think that generally the scope of uncertainty is simply too great to proceed that way.

References

Arora, S. and N. Cason, Timothy (1996). Why do firms volunteer to exceed environmental regulations? Understanding participation in EPA's 33/50 program. *Land Economics 72*(4), 413–432.

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

Ashby, S., S.-H. Chuah, and R. Hoffmann (2004). Industry self-regulation: A game-theoretic typology of strategic voluntary compliance. *International Journal of the Economics of Business 11*(1), 91–106.

David, M. (2005, Jan). Regulating a polluting oligopoly: Emission tax or voluntary agreement? *Review of Development Economics 9*(4), 514–529.

Dawson, N. L. and K. Segerson (2008). Voluntary agreements with industries: Participation incentives with industry-wide targets. *Land Economics 84*, 97–114.

Denicolò, V. (2008). A signaling model of environmental overcompliance. *Journal of Economic Behavior & Organization 68*(1), 293–303.

Lévêque, F. (1999). Externalities, collective goods and the requirement of a state's intervention in pollution abatement. In C. Carraro and F. Lévêque (Eds.), *Voluntary Approaches in Environmental Policy*, Economics Energy Environment, pp. 17–26. Kluwer Academic Publishers.

Lyon, T. P. and J. W. Maxwell (2003). Self-regulation, taxation and public voluntary environmental agreements. *Journal of Public Economics 87*, 1453–1486.

Maxwell, J. W., T. P. Lyon, and S. C. Hackett (2000). Self-regulation and social welfare: The political economy of corporate environmentalism. *The Journal of Law and Economics XLIII*, 583–617.

Nyborg, K. (2000). Voluntary agreements and non-verifiable emissions. *Environmental and Resource Economics 17*, 125–144.

Schmelzer, D. (1999). Voluntary agreements in environmental policy: Negotiating emission reductions. In C. Carraro and F. Lévêque (Eds.), *Voluntary Approaches in Environmental Policy*, Economics Energy Environment, pp. 55–74. Kluwer Academic Publishers.

Segerson, K. and T. J. Miceli (1998). Voluntary environmental agreements: Good or bad news for environmental protection? *Journal of Environmental Economics and Management 36*, 109–130.

Segerson, K. and T. J. Miceli (1999). Voluntary approaches to environmental protection: the role of legislative threats. In C. Carraro and F. Lévêque (Eds.), *Voluntary Approaches in Environmental Policy*, Economics Energy Environment, pp. 105–120. Kluwer Academic Publishers.

A Numerical Example

The purpose of this appendix is two folds: i) it may help some readers understand the intricacies of the model (at least it helps me!) ii) it provides instances of cases where assumption C holds and instances where it doesn't hold.

Let all firms have the same profit function $\pi(p) = p^2/2$ so that $s(p) = p$ with $p = 4$. Assume that d is uniformly distributed over $[0, 4]$ so that $t^{**} = 2$. An unregulated firm's contribution to social welfare is

$$\pi(p) - ds(p) = 2(4 - 2d)$$

If taxed with t, its social contribution becomes

$$\pi(p-t)-(d-t)s(p-t)=\frac{4-t}{2}(4+t-2d) \tag{4}$$

When $t=d$, the firm produces efficiently $s(p-d)=4-d$ and realizes $\pi(p-d)=(4-d)^2/2$. First best social welfare amounts to[10]

$$S^*=\frac{1}{6}\int_0^4 3(4-d)^2 d\delta=-\frac{1}{6}\left[(4-\delta)^3\right]_0^4=\frac{32}{3}=10.\overline{6}$$

S^* is the sum of the contributions of each quartile of the population:

$$\begin{aligned} S^* &= \frac{37}{6}+\frac{19}{6}+\frac{7}{6}+\frac{1}{6}=\frac{32}{3} \\ &= 6.1\overline{6}+3.1\overline{6}+1.1\overline{6}+0.1\overline{6}=10.\overline{6} \end{aligned}$$

With laissez-faire, social welfare amounts to

$$\begin{aligned} S(0) &= 2\int_0^4 2\,(2-\delta)\,d\delta=-2\left[(2-\delta)^2\right]_0^4=0 \\ &= 6+2-2-6=0 \end{aligned}$$

With second best taxation $t=t^{**}=2$, equation (4) becomes $2(3-d)$. Second best welfare amounts to

$$\begin{aligned} S(t^{**}) &= \int_0^4 2\,(3-\delta)\,d\delta=-\left[(3-\delta)^2\right]_0^4=8 \\ &= 5+3+1-1=8 \end{aligned}$$

[10]Since the density is constant (equal to $\frac{1}{4}$), I discard it in all computations. All expected payoff are thus reported multiplied by 4.

Let $\beta = \frac{1}{2}$. The private benefit to participate is

$$V(p, d, t) = \frac{1}{4}\left((p-d)^2 - (p-t)^2 + (p-c)^2 - p^2\right)$$

We obtain the type of the marginal firm $\delta(t)$ by setting $V(4, \delta(t), t) = 0$.

$$4 - \delta(t) = \sqrt{(4-t)^2 - (4-c)^2 + 16} \tag{5}$$

With participation $\delta(t)$, the second period rate is revised to $\hat{\delta}(t) = \int_{\delta(t)}^{4} \frac{\delta}{4-\delta(t)} d\delta = \frac{\delta(t)+4}{2}$. The equilibrium rate is not revised; we use this fix point property to compute it:

$$\begin{aligned}
\hat{\delta}(t_e) &= t_e \\
\delta(t_e) + 4 &= 2t_e \\
2(4 - t_e) &= 4 - \delta(t_e) \\
4(4-t_e)^2 &= (4-t_e)^2 - (4-c)^2 + 16 \\
4 - t_e &= \sqrt{\frac{16-(4-c)^2}{3}}
\end{aligned} \tag{6}$$

We obtain the equilibrium participation by substituting (6) into the r.h.s. of (5).

$$4 - \delta(t_e) = 2\sqrt{\frac{16-(4-c)^2}{3}} \tag{7}$$

In the first-period, each participating firm contributes

$$\pi(p-c) - ds(p-c) = \frac{4-c}{2}(4-c-2d)$$

In the second period, they efficiently contribute $(4-d)^2/2$. The total welfare from participating firms amounts to

$$\frac{1}{4}\int_0^{d_e}\left[(4-c)^2-(4-c)2d+(4-d)^2\right]d\delta$$
$$=\frac{1}{4}\left[(4-c)^2d_e-(4-c){d_e}^2-\frac{(4-d_e)^3-64}{3}\right] \quad (8)$$

Non participating firms generate $2(4-2d)$ in the first period and

$$\frac{4-t_e}{2}(4+t_e-2d)=\frac{1}{2}((16-t_e^2)-(4-t_e)2d)$$

in the second period. The total welfare from non participating firms is

$$\frac{1}{4}\int_{d_e}^{4}\left[(32-t_e^2)-(8-t_e)2\delta\right]d\delta=\frac{1}{4}\left[(32-t_e^2)\delta-(8-t_e)\delta^2\right]_{d_e}^{4}$$
$$=\frac{1}{4}\left[4t_e(4-t_e)-(32-t_e^2)d_e+(8-t_e)d_e^2\right]$$
$$\text{using } t_e=\frac{d_e+4}{2}, \text{ we get } \quad =\frac{1}{4}\left[-\frac{d_e^3}{4}+7d_e^2-28d_e+16\right] \quad (9)$$

Total welfare with a voluntary agreement is the sum of (8) and (9) :

$$SV(c)=4+\frac{d_e}{24}\left[6(c+d_e/2-4)^2-(d_e^2-18d_e+72)\right] \quad (10)$$

To simplify the exposition, re-parameterize the audit cost with a (decreasing) monotonous transformation (drawn in Figure 2 below):

$$c(\delta)=4-\sqrt{16-\frac{3(4-\delta)^2}{4}} \quad (11)$$

Using (7), we then have $d_e = \delta$. Substituting (11) into (10)

$$SV(\delta) = 4 + \frac{\delta^2}{4}\,(c(\delta) - m(\delta)) \tag{12}$$

where $m(\delta) = \frac{2\delta}{3} - 5 + \frac{8}{\delta}$. The conditions $V(p, d_0, t^{**}) > 0$ and $\frac{\partial V}{\partial p}(p, d_0, t^{**}) > 0$ yield $c(\delta) < 2$ or equivalently $\delta > 0$.

I now establish that the Regulator would not gain by subsidizing the audit. Suppose that $c(\delta)$ is only the portion of the audit cost paid by the firms and that the Regulator pays the rest $\overline{c} - c$ where $\overline{c}$ is the audit cost. Social surplus then amounts to

$$SV(\delta) - (1-\beta)(\overline{c} - c(\delta))s(p - c(\delta))\delta \tag{13}$$

where δ is the measure of the population that get audited in equilibrium. The Regulator may then choose δ, or $\overline{c} - c(\delta)$ (the value of the subsidy) under the constraint that $\overline{c} - c(\delta) \geq 0$. Taking the derivative of (13) with respect to δ, we get

$$4g(\delta)\overline{c} + h(\delta)$$

where $g(\delta) = 3(3-\delta)^2 - 35$ and $h(\delta)$ are polynomial forms. Notice that g is a parabola whose roots span an interval $[3 - \sqrt{35/3}, 3 + \sqrt{35/3}] \simeq [-0.42, 6.42]$ that includes D. It follows that g is negative on D. Since $\overline{c} \geq c(\delta) \geq 0$, it follows that

$$4g(\delta)c(\delta) + h(\delta) \geq 4g(\delta)\overline{c} + h(\delta)$$

on D. Evaluating the l.h.s. on D, we get a negative expression, so that the r.h.s. is negative as well. Hence, whatever the value of δ on D, social welfare decreases with δ. The Regulator should thus set δ as low as possible — or equivalently set $c(\delta)$ as high as possible;

which means $c(\delta) = \bar{c}$ — by not subsidizing the audit. Hence, in this numerical example, $SV(\delta)$ does measure the maximum surplus that the Regulator can achieve with a voluntary agreement.

With no voluntary agreement, the Regulator achieves

$$S^{\beta} = \frac{1}{2}S(0) + \frac{1}{2}S(t^{**}) = 4$$

Comparing this expression with (12), we see that assumption C holds whenever $c(\delta) \geq m(\delta)$. These two functions are drawn in Figure 2. Assumption C holds when δ is high, or equivalently, when the audit cost is low.

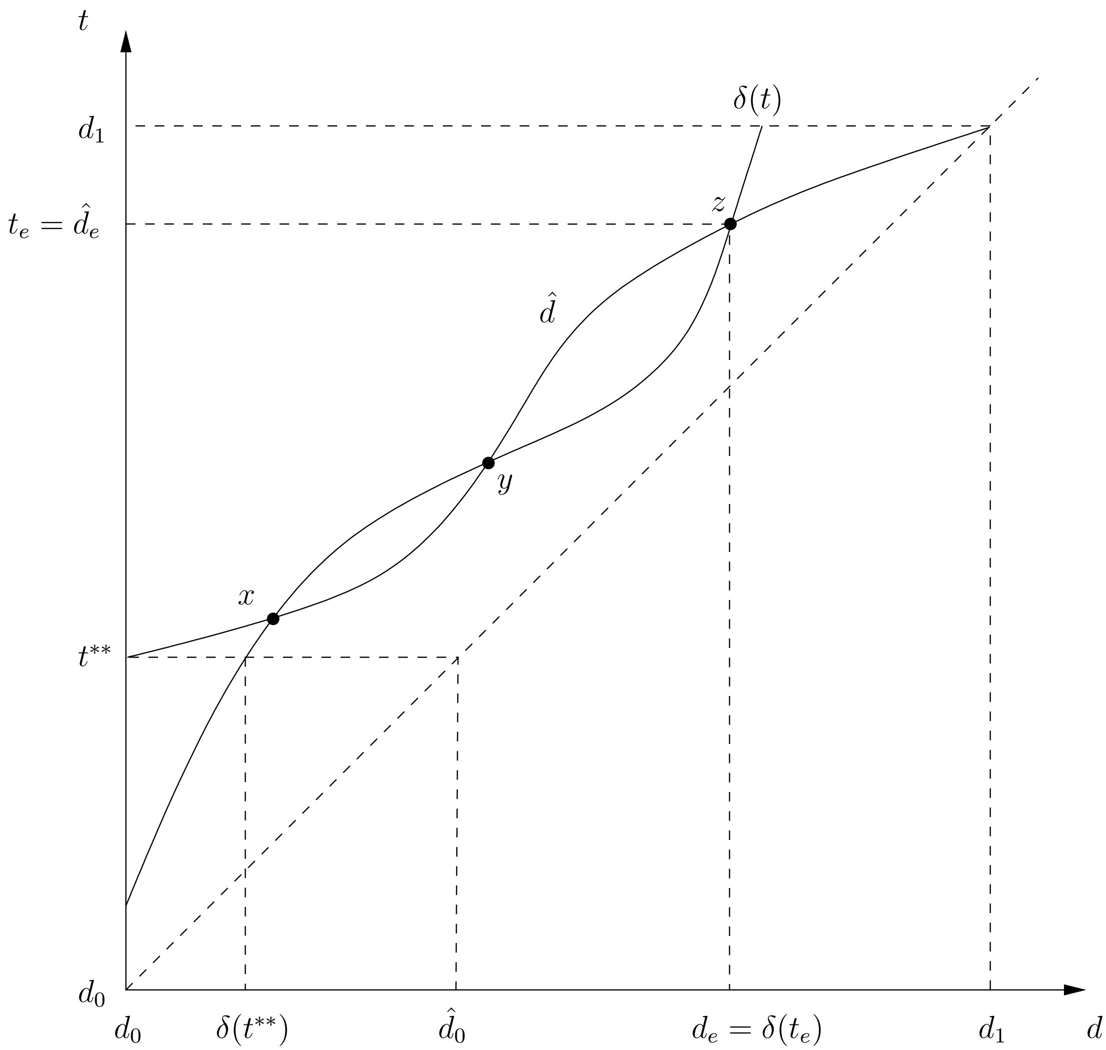

Figure 1: Equilibria.

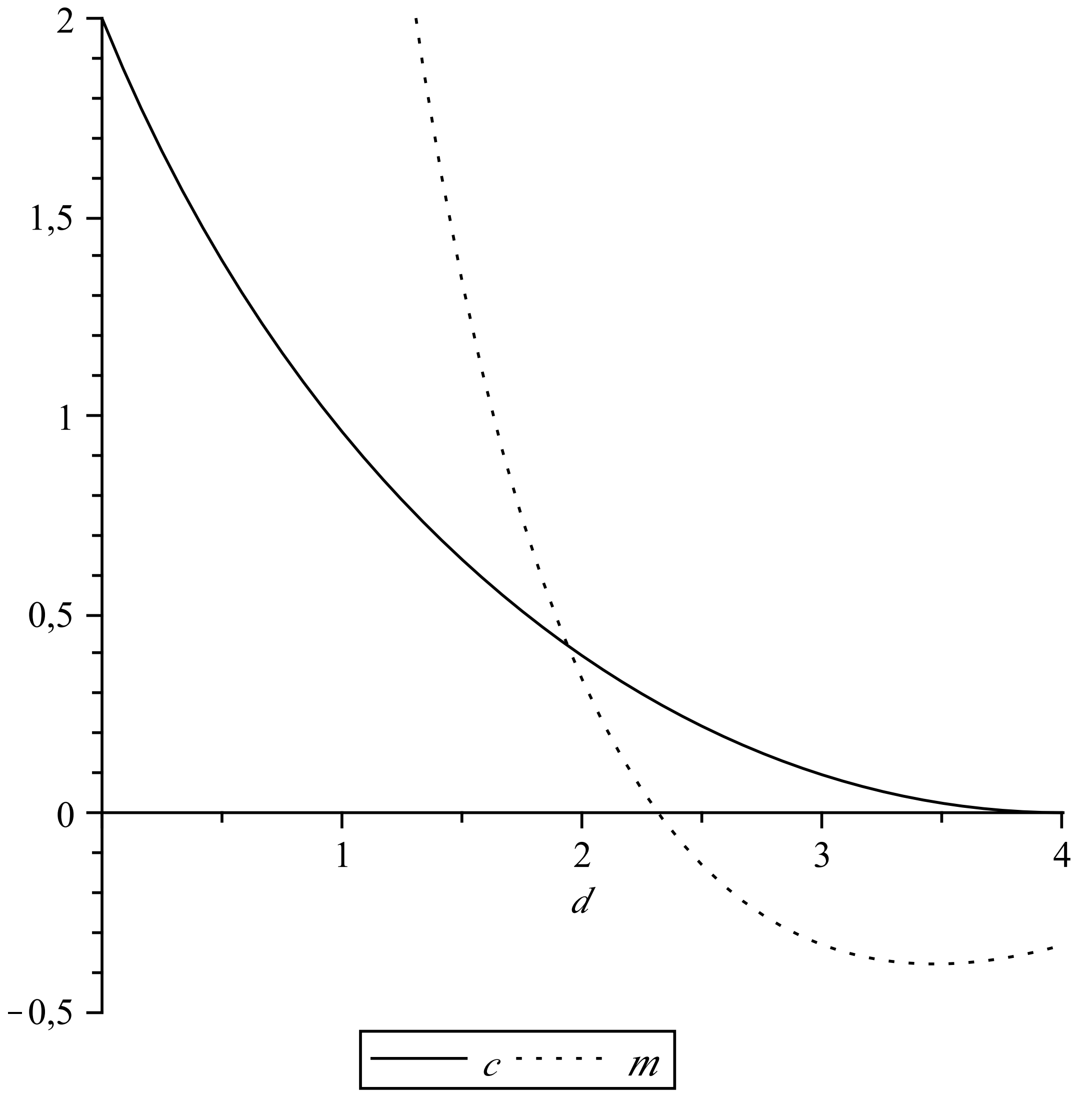

Figure 2: A voluntary agreement increases welfare when $c(d) > m(d)$. This happens when participation d is high, or equivalently, when the audit cost is low.