An inter-disciplinary assessment of the implementation of a european environmental policy: the municipal waste incineration directive (89/429/CEE)

Simone Schucht

Part C Econometric Evaluation of Compliance and Enforcement

The ultimate objective of this third part of the thesis is to econometrically analyse the enforcement of the 1989 municipal waste incineration Directive in France, this being the implementation case where enforcement was incomplete and an implementation gap occurred.

Two strands of literature are investigated with a view to this econometric study. Firstly, we draw on that part of the environmental economics literature that deals with the implementation issue: the monitoring and enforcement literature. Focusing on three models we will show in chapter 8 how an implementation gap can increase the efficiency of an inefficient policy and formulate a number of hypotheses necessary for our econometric study. Secondly, we review in chapter 9 econometric studies on compliance, monitoring and enforcement behaviour in environmental regulation. This is done with a view to assessing what kind of findings with respect to the effectiveness and efficiency of enforcement and compliance this literature has established and which questions it has so far left open. It also serves to identify some methodological issues relevant to our econometric study.

This econometric study is carried out in chapter 10. Compared to other available econometric studies, it shows three specific features: it studies a case where the polluter is a political agent, applies to the European context, and investigates initial compliance. The major objective of the analysis is to find quantitative evidence on whether enforcers, and the regulated incinerators, behaved cost-effectively when making enforcement and compliance decisions.

Chapter 8 A Discussion of Three Selected Monitoring and Enforcement Models

1 Introduction

The previous chapters argue that a gap in implementation might be justified if the initial policy is not optimal and that, from an economic point of view, selective compliance may increase the cost-effectiveness of implementation when rather uniform regulations are applied to heterogeneous pollution sources. In order to show how an implementation gap can increase the efficiency of an inefficient policy and to formulate a number of hypotheses necessary for the econometric study of the implementation of the 1989 municipal waste incineration Directive in France, we draw on that part of the environmental economics literature that deals with the implementation issue: the monitoring and enforcement literature.

This rapidly growing literature starts with the seminal article by Becker (1968) and generally deals with the optimality of monitoring and enforcement schemes, including a determination of the optimal monitoring probability and fines. It furthermore discusses a large number of more specific questions applying to varying contexts and characteristics of the regulated agent. This literature has been reviewed in a number of surveys (see for recent examples Cohen, 1999; Heyes, 1998 and 2000; and Polinsky and Shavell, 2000). Instead of adding a further survey to this list, we focus on three models central to the objective of this chapter.

Becker (1968) was the first to suggest that minimising social costs can imply imperfect enforcement and the existence of an implementation gap. Section 2 presents an application of this author's approach to the environment and to the question of how an implementation gap can improve the efficiency of an inefficient policy. Section 3 introduces a cost-effective selective enforcement scheme modelled by Harrington (1988). Particularly important for the econometric study of France in this author's approach is the fact that he models the enforcer-firm relationship as dynamic and interdependent. Finally, the econometric study of the French implementation path is confronted with a specific compliance issue: that of initial compliance, which refers to necessary capital abatement investment in contrast to those efforts necessary to ensure continuing compliance with emission limit standards once the necessary capital investment is made. What this distinction implies for the optimality of policy objectives and how far monitoring issues can be assumed as being specific in the case of initial compliance was first modelled by Harford (2000). This is the subject of section 4. Section 5 concludes.

2 An application of Becker's (1968) approach to the environment

When defining the optimal level of an activity entailing negative externalities, for example pollution caused as a side effect of production processes, traditional environmental economics did abstract from the fact that related pollution standards might require enforcement. It defined the overall damage D caused by pollution as being determined by the related harm to society and the private gain (i.e. avoided pollution reduction costs) to the polluter. The optimal pollution level, the level where overall costs to society as a whole (including the polluter) -equivalent to the social loss- are minimised was subsequently the level where the marginal external harm to society equals the marginal private gain of the polluter.[1] Under this rule it is considered as optimal to restrict a negative externalities causing activity up to the level where marginal net damages to society overall equal zero (e.g. Baumol and Oates, 1988).

[1] Assuming the following functional characteristics of the harm to society H and the gain to the polluter G in dependence of the level of pollution e: $H_e > 0$, $G_e > 0$ and $H_{ee} > 0$, $G_{ee} < 0$, that is an increasing function of the marginal harm of pollution to society and a decreasing function of the marginal gain of pollution to the offender. These assumptions are necessary to ensure that $D_{ee} = H_{ee} - G_{ee} > 0$.

The underlying hypothesis is that a certain pollution standard could be enforced without using any administrative resources for enforcement itself. This would either imply that regulated agents always obey to law or that they can be forced to do so at zero cost. In traditional environmental economics implementation is thus assumed to be perfect and enforcement to be costless. The first coherent analysis of the role of enforcement was Gary Becker's (1968) seminal article 'Crime and Punishment: An Economic Approach' which did not specifically deal with the environment but developed a more general economic model of crime. This author assumes that obedience to a law cannot be taken for granted and that, therefore, generally private and public resources are spent in order to prevent offences. Otherwise, violations of the law were to be expected.

Building on Becker's (1968) analysis and results, but applying them to the environment, a simplified framework defining the optimal strictness of a regulation is developed below. We will see that when the administrative costs of monitoring and enforcement are taken into consideration the optimal regulatory standard is less strict than when they are not. This implies that a policy which is defined without considering these costs is not optimal, and an implementation gap may be efficient.

2.1 The properties of the optimal policy when there exist positive monitoring and enforcement costs

In this section it will be shown that the optimal strictness of a regulation is lower when monitoring and enforcement costs are taken into consideration than when these costs are neglected. As a corollary of this an implementation gap is efficient if policy objectives with respect to aggregate abatement are set without considering monitoring and enforcement costs.

Consider a population of polluters which bear an aggregate abatement cost C(a), where a is the aggregate level of abatement. The cost function exhibits the classical properties $C_a > 0$ and $C_{aa} > 0$, implying that the marginal cost of abatement increases in abatement. As a further hypothesis the level of abatement a entails the damage D(a) with $D_a < 0$ and $D_{aa} < 0$. Abatement reduces damage but it does so at a decreasing rate.

To keep the analysis simple, we further assume that the aggregate abatement level of the population of polluters is a function of monitoring and enforcement activity:

$$a = a(p, F) \tag{1}$$

where p is the monitoring probability and F the enforcement incentive (for example a fine). Of course, it is assumed that

$$\frac{\partial a}{\partial F} > 0 \text{ and } \frac{\partial a}{\partial p} > 0, \tag{2}$$

meaning that a higher penalty and an increased monitoring probability induce the polluter to abate.

The regulator faces a monitoring and enforcement cost M which depends on the level both of the monitoring and the enforcement activity M = M (p, F). As to the properties of this cost function M we make the following hypothesis[2]:

$$\frac{\partial M}{\partial F} > 0; \frac{\partial M}{\partial p} > 0 \tag{3}$$

[2] Becker (1968) actually assumes that $\partial M/\partial F \geq 0$, implying that an increase in the penalty does not always result in higher monitoring and enforcement costs. When discussing the overall (social) costs arising from various forms of penalties, this author assumes that fines only require collection costs. Neglecting these collection costs results in $\partial M/\partial F = 0$ for monetary fines. The assumption of costless fines is contradicted for example by Polinsky and Shavell (1992). These authors assume that the imposition of fines entails costs of prosecution and penalising of polluters which can be viewed as the expense of collecting fines through a formal enforcement process but also as the expense of settlement negotiations.

In order to establish the optimal policy, the social cost function can be written as follows

$$S = S(a,p,F) = D(a(p,F)) + C(a(p,F)) + M(p,F) \tag{4}$$

of which minimisation by the regulator leads to the two first order conditions defining the interior solution which yields the optimal policy a**:

$$\begin{cases} -D_a(a) = C_a(a) + \dfrac{\partial M / \partial F}{\partial a / \partial F} & \text{(5a)} \\ -D_a(a) = C_a(a) + \dfrac{\partial M / \partial p}{\partial a / \partial p} & \text{(5b)} \end{cases}$$

Note that if the regulator were myopic, in the sense that he would not take into account the monitoring and enforcement cost, he would set a (sub optimal) policy a* implicitly defined by

$$-D_a(a) = C_a(a) \tag{6}$$

Comparing (6) with (5a and 5b) it is immediate that a**< a*.

Unsurprisingly, the optimal aggregate regulatory standard is less strict when monitoring and enforcement costs are taken into consideration. This is so because monitoring and enforcement imply costs to society as a whole which reduce the social welfare of pollution reduction. This implies that, if policy makers fail to take into account enforcement and monitoring cost -leading to a sub-optimal standard a*- an implementation gap (a*-a**) is socially optimal.

2.2 Heterogeneous polluters and uniform standards

In this section, we extend the reasoning of Becker to the case where polluters are heterogeneous and standards are uniform. The reason for investigating this case is related to the municipal waste incineration Directive which relies on rather uniform standards whereas incinerators are very heterogeneous. Standards are thus clearly not cost-effective since they are not differentiated according to individual cost characteristics. In other words, the standards fail to equal marginal costs across polluters. In this context, can an implementation gap enhance the efficiency?

To investigate this question, we now assume that the population of polluters is a continuum of polluters of mass 1. For the sake of simplicity, each individual polluter indexed *i* makes a binary compliance decision which is either 1 or 0, with compliance costs equal to $C_i \neq 0$ or 0, respectively. The firms' compliance costs *C* are uniformly distributed over the interval $[0, \overline{C}]$. The respective distribution function and cumulative function are denoted g(C) and G(C).

We assume that polluters consider penalties similarly to any other cost related to their activity. When making compliance decisions they thus minimise the sum of the expected abatement and non-compliance costs. The non-compliance cost is the expected penalty the polluter will have to pay if the enforcer detects non-compliance. The expected value of penalties depends on the monitoring probability *p* and on the penalty *F*.

More specifically, the polluter *i* complies if and only if

$$C_i \le pF, \tag{7}$$

that is if its compliance cost is lower than the expected penalty pF.

Notice that given our assumptions, the compliance rate is equal to the aggregate level of abatement a. Using the uniform distribution function, the aggregate level of abatement is

$$a(p,F) = G(pF) = \frac{pF}{\overline{C}} \tag{8}$$

In fact, p and F constitute the available political action variables for the regulator. Furthermore, assuming risk-neutral firms, an increase in p compensated by an equal percentage reduction in F would not change the expected income or utility from non-compliance. It would therefore have no effect on compliance. Put differently, a given a can be reached by either increasing p or F.[3] As compliance incentives, p and F are perfect substitutes.[4]

But the most important remark is that a uniform standard which is uniformly enforced leads to a heterogeneous cost-effective response of the polluters, i.e. to a cost-effective distribution of compliance effort between firms. With all firms facing equal expected non-compliance costs pF, only those firms comply for whom the compliance costs do not exceed the expected non-compliance costs. In other words, the aggregate abatement cost

$$C(a) = \int_0^a f(c)c\,dc = \frac{a^2}{2\overline{C}} \tag{9}$$

is minimised.

With respect to the cost-benefit efficiency of the induced pattern of compliance it is, however, worth noting that -assuming social welfare to be the unweighted sum of compliance costs and environmental damage- the number of firms complying and, hence, the aggregate abatement level would be first-best only if the product pF happens to equal the marginal expected environmental damage caused by non-compliance. In the end pF is equivalent to a uniform tax.

If compliance decisions were not binary, and if firms instead would choose the exact level of abatement from a continuum, each firm would apply abatement effort up to the point where their marginal abatement cost equals the marginal expected penalty. This would result in an equalisation of abatement costs amongst the complying firms.

Alternatively, if pF were not uniform, the regulator could apply differentiated monitoring and enforcement effort across firms, i.e. stricter monitoring and enforcement to those firms for whom compliance costs exceed the expected penalty and lower monitoring and enforcement to those whose compliance costs are lower. This might in principle reduce monitoring and enforcement costs compared to the previous case, while still aiming at 100% compliance. But in this case the policy would be implemented in its inefficient form, i.e. abatement would be uniform despite the firms' non-uniform abatement costs. With this chances are that the overall abatement effort would be too high and that the overall social costs of policy implementation would outweigh the policy's benefits.

Note that our above analysis is simplistic. In particular, the expected penalty is lumpy, i.e. it does not vary with the size of the violation. Results will be changed with variable fines.

[3] In this context Becker (1968) has made a further policy suggestion. Assuming monetary fines as enforceable quasi without cost he considers monetary sanctions as preferable over non-monetary sanctions, as the latter are socially more costly. He further suggests that in the case of risk-neutral regulated agents the social loss from non-compliance would be minimised if p were lowered arbitrarily close to zero in order to save monitoring costs and F were raised sufficiently high so that pF induces the attempted level of compliance.

[4] This would no longer hold if actors had a preference for risk or were risk averse. While an increase in p compensated by a reduction in F would still not change the expected income from non-compliance, it could change the expected utility. In the case of a risk preference, an increase in p would reduce the expected utility and with that non-compliance (or: increase a) more than the equal percentage increase in F. In the case of risk aversion, F would have the greater effect.

But this is not a problem, since our analysis definitely applies to the initial compliance of waste incinerators which is a binary decision (as opposed to continuous compliance).

2.3 Summary of the results and discussion

The simple modelling exercise has identified two ways of improving the efficiency of a policy through an implementation gap.

First, when policy makers fail to take into account monitoring and enforcement costs - leading to a policy which is too strict- imperfect compliance is a way to restore *ex post* the efficiency of a policy by reducing the strictness of the standard. This result is particularly relevant when discussing EU Directives. EU policy makers are not at all involved in enforcement and are more inclined than others to forget monitoring and enforcement costs.

Second, the model establishes how the uniform enforcement of a uniform -inefficient- regulatory standard can lead to a cost-effectively differentiated implementation gap. This is also relevant for us as shown by the environmental federalism literature which establishes that higher jurisdictional levels are more likely to enact over-uniform standards.

3 Conditions for a cost-effective implementation gap - Harrington's (1988) targeting model

Our analysis based on Becker (1968) suggests that limited enforcement -and the related implementation gap- can be efficient. This obviously does not imply that all observed implementation gaps are efficient. Efficiency will depend on the enforcement rules according to which enforcers select plants that are forced to comply or not. For instance, we show in section 2.2 that a uniform lumpy penalty is cost-effective. One example of how rules can be determined that define a desirable monitoring and enforcement scheme was developed by Harrington (1988).

This author models a monitoring and enforcement scheme minimising administrative costs of monitoring and enforcement under a given a target compliance rate. With this, the model takes a different angle to cost-effectiveness of monitoring and enforcement than our analysis in section 2.2. Instead of targeting all firms with the same uniform penalty, this author derives a cost-effective result by applying differentiated enforcement to a specific set of firms. This brings the model closer to the French case where non-compliant incinerators were subject to differentiated enforcement measures. The effect of these different approaches is that overall implementation costs in Harrington's model are reduced through a minimisation of administrative costs, while in section 2.2 they were reduced through a cost-effective allocation of abatement effort between firms. Put differently, in our model the enforcement policy was externally defined, whereas it is the target compliance rate of firms which is defined in Harrington's model.

3.1 The modelling approach

The author derives rules for monitoring and enforcement behaviour which follow a more complex approach than presented above. He models a dynamic relationship between regulated agents and enforcers and determines how the monitoring and enforcement scheme should be structured, towards which firms monitoring and enforcement measures should be applied, and which firms should be left non-compliant in order to reach a target compliance rate at minimal administrative cost.

Compared to a welfare maximising scheme Harrington's (1988) scheme is second best: It abstracts from environmental damage and compliance costs in the enforcer's behavioural function and is restricted to modelling administrative cost-effectiveness. Harrington's (1988) model is also second best in that it introduces a restriction of the maximum applicable fine per period.[5] Harrington's approach however comes closer to reality. In practice, there are limits to fines and levying fines may be costly to enforcers as well. Furthermore, enforcement budgets are generally limited.

A further interesting and central aspect of this model, and in which it diverges from Becker's analysis, is the dynamic view it takes, introducing the idea that enforcers and regulated agents make decisions based on the anticipated behaviour of the other. By introducing a dynamic firm-enforcer relationship, firm and enforcer behaviour become endogenous functions. This assumption also brings the model closer to real world conditions and constitutes an issue important to an econometric estimation of empirically observed compliance and enforcement behaviour, which is discussed in more detail in the following two chapters.

3.2 The model

Harrington's (1988) model focuses on the relationship between the firm's compliance costs and the average level of compliance that can be reached given the restrictions to enforcement outlined above.[6] The agency is assumed to minimise monitoring and enforcement costs, consistent with achieving a given target compliance rate by the firm. The author then studies how the enforcer can select (i.e. target) firms towards which to apply monitoring and enforcement in a cost-effective way. In the model, the agency classifies firms in two groups G_i (with i=1, 2), where one set of firms faces more severe monitoring and enforcement than the other. More specifically, be the monitoring probability p_i and the penalty for a violation F_i, where $p_1 < p_2$ and $F_1 < F_2$. Firms have two discrete compliance choices (compliance and non-compliance) and move between the groups G_i depending on their previous compliance status. Detected violations in G_1 are punished with exile to G_2, and discovered compliance in G_2 offers a chance of return to G_1 with a probability denoted u.

The parameters are selected by the agency beforehand and determine the transition probabilities and payoffs to the firms, and thus the compliance rate. As further assumptions, the regulator knows the firm's compliance costs and the firms know the agency's enforcement parameters as well as to which group they are assigned. To the firm this monitoring and enforcement scheme presents a Markov decision problem, where the transition probabilities and the pay-offs (costs) incurred by the firm actually depend not only on the parameters chosen by the enforcer, but also on the action (complying, not complying) taken by the firm during this period.

As a first important result, Harrington (1988) shows that in a dynamic model of compliance and enforcement, with firms subject to different monitoring probabilities and fines dependent on the firms' previous compliance behaviour, a given compliance target rate can be reached at lower monitoring and enforcement costs overall compared to those achievable under a static approach. In static models, previously prevailing in the monitoring and enforcement literature, the penalty depended on the firm's rate of emissions (or compliance). The firm was thus inspected with a certain probability, and the firm was better off when complying once its expected non-compliance costs outweighed its compliance costs, i.e. when $pF > C$. Contrary to this, dynamic models assume that both the enforcer and the regulated agent react to previous actions by the other. The enforcer can thus alter the expected penalty based on the firm's previous compliance behaviour, and this changes the optimal strategy for the firm. In fact, as a second important result of the model, firms may have an incentive to comply even if their

[5] This restriction in Harrington's model implies that Becker's suggestions allowing to minimise social costs - by reducing (for the case of risk neutral agents) *p* arbitrarily close to zero, in order to save on monitoring costs, and by increasing the monetary fine arbitrarily high, in order to ensure deterrence of violations- cannot be applied.

[6] The author adapts the state-dependent models developed by Landsberger and Meilijson (1982) of income tax enforcement.

compliance costs each period exceed the expected penalty, and even if they exceed the maximum penalty feasible in any period.

This result, according to the author, may give an explanation for the frequent empirical finding that firms are in compliance despite actual inspection rates and penalties implying an expected penalty lower than what is considered necessary to induce compliance in static theoretical models.

3.3 Which firms should be targeted?

There are four policies available to the firm: complying when in both states (G_1 or G_2), denoted as f_{00}, complying in G_1 but not in G_2 (f_{01}), complying in neither group (f_{11}), or complying only in G_2 (f_{10}). It is then possible to calculate, firstly, the transition probabilities which depend on the firm's compliance decision, on the original group the firm is placed in, on p_i and on u. Secondly, it is possible to calculate the expected cost, in present value[7] of each policy f_{ij} when initially in state m by solving the simultaneous equations for the expected cost depending on m. Minimising the expected costs requires solution of the 4 simultaneous equation systems and yields the optimum policy for the firm. Firms' optimal policies depend on their compliance costs and the enforcement parameters chosen by the enforcement agency.

As an important result the author shows that while the compliance policy f_{01} is never optimal to the firm[8], the optimality of the other three policies depends on the firm's compliance costs. Furthermore, both f_{ij} and the expected cost of the optimum policies are linear functions of the compliance cost C. Given that the probability of a firm being compliant, depending on its compliance costs, equals 1 for f_{00} (low cost firms) and 0 for f_{11} (high cost firms), the regulator can save monitoring and enforcement costs by not surveying these two groups of firms and by accepting that the group f_{11} will not comply. The group of firms of interest to the regulator is therefore the one whose optimum policy is f_{10} and which complies only part of the time.

Given furthermore that these firms will be in compliance in G_2 and out of compliance in G_1, Harrington (1988, 39) suggests that F_1 can be set equal to 0 and that such firms will still be in compliance part of the time if p_2F_2 is sufficiently large. Finally, he argues that the firms will even be in compliance part of the time if $p_2F_2 < C$, because compliance in G_2 allows the firms to eventually return to G_1, where the cost of non-compliance is lower.[9] Contrary to this, in a static model the absence of fines would imply that only those firms comply whose compliance costs are zero; and an expected penalty lower than the compliance costs would not be sufficient to induce a firm to comply. This two-group state-dependent enforcement approach therefore allows regulators to ensure a certain rate of compliance at lower monitoring and enforcement costs.

Note that while the assumption of abatement costs constituting a limit to the enforceability of a regulation -which in this model results in high cost firms (firm group f_{11}) never complying- is coherent with the model assumptions, in practice it will only apply to very specific situations. In practice this assumption may hold if the limited fines are the only penalty possible. It will not hold when the regulator can suspend operations and enforce the closure of a plant by removing its operation licence. Such a drastic decision, however, may be subject to concerns that are not fully of an economic but at least partly of a political nature, an issue discussed in the following chapter.

[7] The cost in the current period plus the expected present value discounted one period.

[8] This follows because f_{01} is dominated by f_{00} when $C_i < p_2F_2$ and by f_{11} when $C_i \geq p_2F_2$.

[9] In fact, in this model 'good' firms can afford to cheat because they are in the low enforcement regime, while 'bad' firms do not cheat because they wish to remain in compliance and to return to the 'good' category.

3.4 The optimal targeting scheme towards f_{10} firms

The agency aims at minimising its monitoring and enforcement costs consistent with the firm achieving a target long-run compliance rate *a* and does so by minimising the average inspection rate p^*. The firm's compliance costs C_i are given and known by the agency. The variables the agency can modify to reach its goal are p_1, p_2, F_1, F_2 and *u*, while the maximum fine $\overline{F}$ is externally given.

Harrington (1988: 40) then establishes under which conditions the enforcer can achieve a given compliance rate:

- If the target compliance rate is a = 1, i.e. the enforcer wants all firms to comply (no implementation gap), it is necessary that $p^* \overline{F} > C_i$ just as in the static model and also that $\overline{F} > C_i$.
- If the present value of all future fines is outweighed by the firm's compliance costs the agency could not induce any compliance. This is expressed by $C_i > \overline{F}/(1-\beta)$ which indicates that the cost of complying even once exceeds the present value of all future fines.
- Even if $C_i > \overline{F}$, some degree of compliance can be achieved as long as $C_i < \overline{F}/(1-\beta)$. If this holds and if the given target compliance rate is a < 1, the optimal penalties are $F_1=0$ and $F_2=\overline{F}$. This is so because in the model the frequency of being in compliance for the firm with the optimal policy f_{10} as well as the average inspection rate p^*_{10} are only dependent on p_1, p_2 and u, and not influenced by the penalties. The results for the optimal penalties imply that, at the optimum, no fines are ever collected because the firm cheats only in G1 where $F_1=0$.[10]
- If, however, $C_i > \overline{F}$, p_2 must be at least $aC_i/\overline{F}$. This implies that the maximum feasible compliance rate is $a = \overline{F}/C_i$ in which case $p_2 = 1$.

In summary, as a further result, state-independent enforcement schemes can give incentives for (partial) compliance only as long as C_i does not exceed $\overline{F}$, while state-dependent enforcement schemes can achieve partial compliance even if $C_i > \overline{F}$. Making monitoring and enforcement schemes state-dependent can therefore set additional compliance incentives and reduce monitoring and enforcement costs relative to state-independent schemes. Finally, state-dependent models can achieve a greater compliance rate than state-independent models and are thus more cost-effective than state independent models when $C_i < \overline{F}$. This cost-difference however decreases when *a* increases towards 1.

From what was said before one can conclude that the superiority of state-dependent over static models in terms of cost-effectiveness hinges on the acceptance of an implementation gap. Secondly, one can think of the gap consisting of two analytical components. Firstly, the high cost firms whose probability to be in compliance is zero and who are not targeted at all by monitoring and enforcement schemes and, secondly, the non-compliance rate amongst the f_{10} type firms determined by 1-a.

While defining a cost-effective targeting strategy in terms of administrative costs, this model approach where the rate of compliance for an f_{10} firm depended only on the transition probabilities but not on compliance costs also implies the possibility of an inefficient outcome. Harrington (1988) himself points to the fact that if the regulator applied similar inspection probabilities and similar parameters *u* to all f_{10} type firms, their compliance rates would be the same even though their compliance costs may be different.

[10] An increase in the penalties does not affect the compliance rate unless it changes the class a firm belongs to, i.e. the policy optimal to a firm.

Moreover, his model implicitly assumes that compliance history is more decisive in shaping enforcement policy than for example environmental damages related to non-compliance. Finally, giving up the assumptions that the firm knows the agency's enforcement parameters as well as to which group it is assigned and that the regulator knows the firm's compliance costs might have impacts on the results.[11]

3.5 Implications of Harrington's targeting model for our empirical study of monitoring and enforcement

Harrington's model and its successors show two important characteristics that are taken into account in the majority of empirical econometric studies of monitoring, enforcement and compliance and/or are retained for the econometric study on the implementation of the 1989 municipal waste incineration Directive in France.

The first characteristic is the modelling of the enforcer-firm relationship as dynamic and interdependent, where both enforcers and regulated agents react on each other's behaviour. As will be shown in the following chapters, empirical econometric studies, in general, yield more significant results for the impact of monitoring and enforcement on compliance and for the impact of the compliance state of the firm on the enforcer's behaviour if this interdependency is taken into consideration in the analyses. But while Harrington assumes that both actors react on the other's previous behaviour, the econometric study in chapter 10 will model the actors' behaviour as dependent on the anticipated behaviour of the other.

The second characteristic is the divergence of the enforcer's behavioural function from pure welfare maximisation. With its positive approach Harrington's model introduced a new enforcer type, the enforcer characterised as 'bureaucratic' (cf. chapter 9), while its successors modelled an enforcer coming closer to welfare objectives. Whereas an enforcer motivated by 'bureaucratic' objectives would pay no attention to compliance costs and damage when making his enforcement targeting decision, a welfare maximising regulator would.[12] Indeed, the empirical econometric studies reviewed in the following chapter show that enforcers in practice seem to follow very different objectives when establishing their monitoring and enforcement policy.[13] Quite frequently, political considerations as well play a role for the enforcers' behaviour. Attempts are also made in chapter 10 to econometrically test the relevance of different enforcer types for the French implementation case, which represents an example of selective enforcement.

4 The first attempt to treat initial compliance – Harford's (2000) model

Harrington's (1988) model is not directly applicable to our econometric study of the French compliance and enforcement path in that it treats continuous compliance (e.g. the continuous meeting of a certain emissions standard) rather than initial compliance (investment in abatement technology), which is the issue of the econometric study for France. This is immediately obvious when considering Harrington's (1988) assumption of firms being able to move in and out of compliance in subsequent periods. This assumption makes more sense with respect to the necessity of continuously meeting pollution standards dependent on variable inputs than with respect to investing in abatement capital.

[11] This was partly studied in subsequent articles (cf. Harford and Harrington, 1991; Harford, 1991; Harford, 1993).

[12] Note that the use of the term 'bureaucratic' should not be confused with the designation in Niskanen's (1975) bureaucratic behaviour theory.

[13] Note that Harrinton's (1988) model is specifically tested by one econometric study (Helland, 1998b; cf. chapter 9). But opposite to Harrington this author doubts that regulated firms, in practice, can easily change their compliance status. This author therefore replaces Harrington's assumption of firms changing the groups by changing their compliance status by the assumption that they self-report in order to signal co-operation to enforcers, thus changing the enforcement groups. In this context it should also be mentioned that when continuing compliance is the issue (and not initial compliance; cf. section 4 below) it is more easily imaginable that plants frequently change their compliance status.

The fact that it applied rather to continuing than to initial compliance is a characteristic that Harrington's (1988) model has in common with the majority of models discussed by the normative theory on monitoring and enforcement. Although several authors pointed out earlier that monitoring and enforcement might differ between initial compliance, on the one hand, and continuing compliance, on the other hand (e.g. Harrington, 1988: 32; Amacher and Malik, 1996: 235), models generally applied to continuing compliance only. Harford (2000) provides the first study that explicitly models optimal monitoring and enforcement schemes for both initial and continuing compliance and discusses optimal standards for both types of compliance.

4.1 The general intuition and results

This model is constructed around a straightforward intuition. It assumes that pollution abatement will generally involve a mix of capital and labour input. The necessary capital input refers to initial investment necessary to comply (for example investment in acid removal scrubbing systems or ESP for dust control by municipal waste incinerators). The labour input ensures continuing compliance with a pollution standard (e.g. an emission standard). It is furthermore assumed that costs for monitoring initial compliance are lower than costs for monitoring continuing compliance. This is so because abatement technology for industrial polluters is quite visible whereas verification of continuing compliance requires (permanent) measuring of emissions.

4.2 The model

The regulator is assumed to maximise welfare by minimising the sum of pollution damages, enforcement costs and pollution control costs. In this the model comes closer to Becker's (1968) approach than did the one previously presented. Regulators can pre-commit either to a technology (or: capital) standard K_s or to an emission standard s, or both. Initial compliance is reached when the firm complies with the technology standard, while for continuing compliance the firm needs to comply with the emission standard. The firm's choice variables to determine its abatement level a are inputs of capital K at price r and a short-run variable input of labour L at price w. It is assumed that the firm can vary L in the short run. Furthermore, capital, once invested, lasts forever and is a fixed cost. One can therefore write the firm's long-run costs of abatement as

$$C = rK + wL \tag{11}$$

and its short-run abatement costs as

$$C = wL \tag{12}$$

Inputs of K and L and the firm's gross pollution amount z determine its net pollution x and with this the damages D(x).

$$x = z - a(K, L, z), \tag{13}$$

where a is assumed to be concave in K and L.

The (maximum) fine is a lump-sum total fine F per period and can be imposed with no administrative cost. A further decisive assumption is that monitoring abatement capital is costless and perfectly accurate so that the monitoring probability for initial compliance is $p_k = 1$. The enforcer furthermore monitors continuous compliance with the probability p_s at cost m. Finally, the firm is assumed to be risk-neutral.

4.3 Optimal abatement should have a capital bias

In this benchmark case it is assumed that the regulator can commit to a capital standard K_s, to an emission standard s, to the monitoring probability p_s for monitoring continuous compliance, and to the monitoring probability $p_k = 1$ for monitoring initial compliance. The firm thus knows that if failing to provide initial compliance it will be fined F per period. It also knows that once initial compliance is reached, the regulator will monitor its pollution to enforce continuing compliance and the firm hence determines its level of pollution (or: continuous compliance). Continuous compliance is then monitored with probability p_s at cost m and, if found non-compliant, the firm is fined F.

In a first step the level of F required for initial compliance can now be determined. Knowing that the regulator will monitor pollution to enforce continuing compliance once initial compliance is reached, the firm will only invest in the required technology standard if the fine at least equals both abatement capital and operating (labour) costs:

$$F \geq rK_s + wL_s \tag{14}$$

In a second step the optimal standards for technology K_s and for pollution s can be determined. For this it makes sense to introduce the social cost function N which, given the fine constraint (14), can be written as:

$$N(s) = D(s) + mp_s + rK_s + wL_s + \lambda(rK_s + wL_s - F) \tag{15}$$

Substituting out p_s by $p_s = wL_s/F$, several transformations of the equation, and minimising the social cost function with respect to K_s and L_s yields

$$(1+\lambda)\frac{r}{a_K} = D_x = \frac{w}{a_L}\left(1 + \frac{m}{F} + \lambda\right) \tag{16}$$

where (r/a_K) and (w/a_L) are the firm's marginal cost of abatement through abatement investment and through the use of labour respectively.

As was to be shown, and as a central result, equation (16) implies that overall abatement should have a capital-using bias. This can be best demonstrated for the case where the fine constraint does not hold ($rK + wL < F$), so that $\lambda = 0$. For this case equation (16) implies that the firm's marginal cost of abatement through abatement investment should equal the marginal damage of pollution, while the marginal cost of abatement through the use of labour should be lower.[14] This is so because capital is less costly to monitor than labour inputs and pollution (continuous compliance).[15]

This result implies that a policy including only an emission standard is not efficient where monitoring abatement capital is less costly than monitoring continuing compliance.[16] What is also important to note here is that under the optimal solution the regulator forces the firm to abate at a capital-to-labour ratio where $a_L/a_K < w/r$ and with this at a point where the firms' costs of meeting the standard are not minimised. Firms' cost minimisation would require $a_L/a_K = w/r$, that is a lower capital standard. Overall cost minimisation, therefore, implies that firms do not minimise costs. Let us now turn to the case where the regulator cannot commit to a technology standard to determine what happens when the technology choice is left to the firm.

---

[14] In fact it is lower than the marginal damage by $(w/a_L)(m/F)$, which represents the marginal cost of enforcing one more unit of abatement through labour.

[15] If the fine constraint were binding, so that $\lambda > 0$, pollution abatement would be even lower.

[16] The finding that pure compliance standards are not efficient when monitoring costs differ between initial and continuing compliance also has consequences for the general economic analysis of environmental policy instruments. Economists have traditionally advocated for economic instruments (tradable permits or emission taxes) over command-and-control type instruments (regulatory standards) because the former show better cost efficiency properties, allowing an efficient allocation of abatement effort between pollution sources. A question logically arising from the traditional analysis of policy instruments for the monitoring and enforcement literature was whether market based incentives -which are assumed to yield lower compliance costs for firms- might result in higher compliance as compared to regulatory standards and thus save costly monitoring and enforcement expenditure. The literature, consequently, turned its interest to policy instruments, studying the impact of monitoring and enforcement costs on the instruments' cost efficiency properties (see for example Malik, 1990 and 1992). The only general result this literature has yielded so far is that economic instruments are no longer systematically preferable when these costs are taken into consideration. In this line is also Harford's (2000) analysis which implies that the efficiency gains expected from applying tradable permits instead of technology standards may be smaller than generally expected when monitoring and enforcement costs are taken into consideration.

4.4 Bargaining for more capital input is optimal when the regulator cannot use a technology standard

A second case studied by Harford (2000) assumes that regulators cannot pre-commit to a technology standard, but only to a pollution standard s=σ. In contrast to the first case the firm here is free to seek the least-cost way of meeting the standard. This case matches with the municipal waste incineration Directive's policy approach, which indeed set pollution standards in the form of emission limit values.

If the firm is free to choose the least-cost way of meeting the standard, the regulator's constraint in the social cost function becomes $a_L/a_K = w/r$. This is inconsistent with the optimality conditions derived under the first case since it implies a lower capital-to-labour ratio. The difference is explained by the fact that in the previous case the firm was forced to comply with a certain technology standard, hence make the respective capital investment, and only then chose the cost-minimising level of pollution abatement through labour. In the present case the firm does immediately seek the cost-minimising abatement way and is free to adjust both capital and labour inputs to meet the pollution standard at least cost.

The social cost function becomes

$$N(\sigma) = D(\sigma) + rK_\sigma + wL_\sigma + mw\frac{L_\sigma}{F} + \mu(rK_\sigma - wL_\sigma) \tag{17}$$

where the monitoring probability was substituted out by the relationship $p_\sigma = w(L_\sigma/F)$.[17]

Several transformations and derivation of (17) lead to the marginal conditions for a social cost minimum as given in the following expression:

$$\frac{w}{a_L}\left[1+\mu-a_L\left(\frac{dL_\sigma}{d\sigma}\right)\left(\frac{m}{F}\right)\right] = D_x = \frac{r}{a_K}\left[1+\mu-a_L\left(\frac{dL_\sigma}{d\sigma}\right)\left(\frac{m}{F}\right)\right] \tag{18}$$

For the same amount of pollution eliminated in both cases, equation (18) implies that compared to the case described by equation (16), relatively more labour and less capital is used to meet the pollution standard. Furthermore, under ordinary circumstances, social costs are higher than in the first case. This is so because a) the regulator has one less instrument, and b) with a higher labour-to-capital ratio monitoring pollution becomes more important, and as monitoring pollution is more expensive than monitoring capital, overall monitoring and enforcement costs increase. Altogether, the sum of abatement cost, control cost and damage costs is higher. From this it follows that the optimal abatement level in the second case would be lower than in the first case, and so would be the abatement cost to the firm. Equation (18) also implies that the marginal cost for both capital and labour inputs would be lower than the marginal damage.

Given these results, Harford (2000) suggests that social costs could be reduced if the regulator did not fully enforce the pollution standard but instead negotiated with the firm a higher capital-to-labour ratio that did not increase the firm's total abatement costs but lowered the regulator's monitoring costs sufficiently to compensate for the costs resulting from increased damage.[18]

Summing up, under the model assumptions, compared to the situation of pure pollution standards, efficiency can be increased by either establishing a complementary capital standard or by negotiating the capital-to-labour ratio with the firm. In the case of the

[17] Furthermore, L_σ and K_σ are functions of σ and defined by $a_L/a_K = w/r$ and $z–\sigma = a(L_\sigma, K_\sigma, z)$.

[18] Investigating the relationships between the parameters of this problem the author shows that the reduction in social cost possible from substituting capital for labour while holding the firm's costs constant is negatively related to the curvature of the isoquant (Harford, 2000: 158).

complementary capital standard enterprises are worse off compared to a pure pollution standard because they are forced to apply a higher capital ratio than would be cost-minimising. However, overall social costs are lower as monitoring costs decrease. If the regulator could only set a pollution standard, overall costs could be reduced if the regulator negotiated a higher capital-to-labour ratio than the one firms would otherwise choose in order to reduce his monitoring costs. To give the firms an incentive to enter into such bargaining the regulator would have to ensure that the outcome would not increase the firms' overall abatement costs, which would imply lower abatement overall.[19]

4.5 Synthesis

Harford's (2000) major results are that, firstly, the optimal mix of capital and labour inputs for abatement will involve a capital bias. Secondly, pure performance standards are not efficient where monitoring abatement technology is less costly than monitoring continuing compliance. Instead, a policy mix is necessary to ensure that firms meet the optimality condition of a capital bias in abatement, because firms would otherwise choose an insufficient capital input. This holds no matter whether the regulator can pre-commit to both a technology and an emission standard or only to an emission standard. The second case is of specific interest here because it comes close to the 1989 Directive's policy approach consisting in the definition of emission limit values.

4.6 Assumptions retained for an empirical econometric study of enforcement and compliance in France

Two elements of Harford's (2000) model are further used and introduced into the econometric analysis of France's compliance path with respect to the 1989 municipal waste incineration Directive in chapter 10. This is, firstly, the study of the relationship between monitoring and enforcement, on the one hand, and initial compliance, defined as having the technology necessary to comply with emission standards in place, on the other hand. In fact the empirical econometric study focuses solely on initial compliance, driven by a lack of data on continuous compliance.

Secondly, Harford (2000) assumed that monitoring initial compliance is costless. Even though monitoring costs for measuring initial compliance, in practice, may not be equal to zero, it seems justified to assume that they are lower than those for measuring continuing compliance.[20] Given that there was a limited number of technologies able to meet the MWI Directive's standards, a failure by the firm to apply either of them made detection of non-compliance with initial compliance (and in that case also continuing compliance) quite straightforward to the regulator. The econometric analysis of the French case hence uses the assumption that monitoring initial compliance is possible at low costs. It also makes the assumption that initial compliance, although frequently assumed to be more easily achievable than continuing compliance, nevertheless requires enforcement. Given the relative ease of monitoring, the focus in this econometric study is therefore on the penalty applied towards non-compliant plants.

While there are some similarities between Harford's (2000) assumptions and the econometric study of the implementation of the 1989 EU Directive, there are also aspects that do not apply to the empirical case or that simply cannot be tested. As one aspect, unlike Harford, the analysis in chapter 10 does not assume enforcement (penalties) to be free of cost. Another important aspect is that while the alternative technologies capable of

[19] For further literature investigating bargaining incentives both for the regulator/enforcer and the regulated agent see for example Amacher and Malik (1996 and 1998); Heyes (1998) and Lehmann (1997).

[20] The obviousness of initial compliance may however rather hold for the local than for a central regulator. The French case showed that the central regulator (the Environment Ministry) in the early phase of the municipal waste Directive's implementation had a rather vague picture about the exact state of the country's incinerators, especially that of the small plants. This however refers rather to a communication issue within the administration than to a visibility issue in general.

meeting the 1989 Directive's standards entailed indeed quite heterogeneous investment and operating costs to the firms, the overall abatement costs were relatively similar across these technologies.[21] And although the abatement effectiveness with respect to different pollutants varied across the abatement technology alternatives, for the whole of the pollutant-package regulated by the 1989 Directive and the ministerial order transposing it into French legislation none of the alternative technologies was considered as more effective than the other (Milhau and Pernin, 1994). If Harford's assumptions hold in practice, this would imply that compared to this author's case even more administrative cost gains would have been possible if the Directive had defined technology standards. By increasing the technology bias, i.e. pushing towards technologies where the capital costs were more important than the labour costs, social costs could have been reduced without offsetting environmental effectiveness since the overall abatement costs for the firm remained rather similar.

But how to explain the assumption of monitoring costs differing in dependence on technologies of different capital content? Harford (2000) does not specify this issue. The most convincing argument seems to relate to the firms' incentives to cheat.[22] Given Harford's assumptions, undetected non-compliance would only be possible with respect to continuing compliance, once technology investment has been undertaken. The monitoring and enforcement literature in general establishes that the higher the incentives for non-compliance, the higher the necessary expected costs of non-compliance to enforce compliance, and hence the higher are overall monitoring and enforcement costs. Coming back to the model, once investment in abatement technology has taken place, these costs are sunk. The remaining abatement necessary is less costly than it would have been otherwise. If with rising capital content the variable cost part is lower, a higher share of overall abatement costs is sunk, and firms have less reason to cheat (i.e. not to comply). Expressed differently, with lower remaining abatement costs the expected penalty necessary to enforce compliance decreases, and hence, -given Harford's (2000) assumption of costless fines- monitoring costs can be reduced.

Finally, a central issue in Harford's (2000) model is the study of compliance in dependence of the capital content of abatement. It is impossible to test the relationship of capital input, compliance and enforcement empirically for compliance with the 1989 Directive. The reasons are that not only data on continuing compliance are lacking but more in general also information on which part of emission reductions (and therefore compliance) are due to the capital part and which are due to the variable abatement input part for different abatement technologies applied in France.

5 Conclusions

As one important result our simple modelling exercise has provided support for one of the hypotheses underlying this thesis: that an implementation gap can improve the efficiency of an inefficient policy. It was shown that where the policy-maker fails to take into account monitoring and enforcement costs, policy standards are likely to be too strict. In this situation, imperfect compliance is a way to restore the efficiency of the policy by reducing the strictness of the standard *ex post*. It was also shown that imperfect enforcement can lead to a cost-effectively differentiated implementation gap where a policy sets regulatory standards that are too homogenous compared to the heterogeneity of the pollution sources to which they apply. Given the inefficiencies in the EU Directive

[21] These cost characteristics are presented in more detail in chapter 10.

[22] Harford's (2000) paper builds on Amacher and Malik (1996 and 1998), but these authors applied a different logic when arguing for varying enforcement costs with different technology choices. Their benchmark cases were either the adoption of an end-of-pipe technology by the firm or the replacement of its production process by a new, cleaner one. As the second option implies lower marginal abatement costs, a smaller expected fine is necessary to ensure compliance, and enforcement costs are thus lower (Amacher and Malik, 1996). As a major difference to Harford (2000), Amacher and Malik (1996) therefore study solely issues of continuing compliance and put the emphasis on enforcement costs in general rather than on monitoring costs.

established in previous chapters we will econometrically test whether enforcers paid attention to monitoring and enforcement costs and to firms' compliance costs when enforcing this policy in France.

Furthermore, the analysis identified two central hypotheses that are retained for the econometric study of the French case. Firstly, where initial compliance is at stake, monitoring was shown to be not so much the issue of the enforcer. This is so because non-compliance with respect to the installation of abatement equipment necessary to reach initial compliance is highly visible and the related monitoring costs low. Where this is the case, deterrence of non-compliance will basically depend on the penalties applied. In the econometric study we, therefore, focus on the occurrence of sanctions towards non-compliant municipal waste incinerators. Secondly, compliance decisions of regulated pollution sources and enforcement decision of the regulator are likely to be interdependent. This means that, on the one hand, one can expect enforcement agencies to target specific municipal waste incinerators, for example those that can be expected to be non-compliant, and, on the other hand, one can expect incinerators to make their compliance decisions depending on anticipated enforcement. We econometrically investigate this assumption in the French case.

Chapter 9 A Review of Econometric Studies on Compliance, Monitoring and Enforcement Behaviour in Environmental Regulation

1 Introduction

Having presented some theoretical and primarily normative models on monitoring, enforcement and compliance, it is in order to now investigate empirical evidence provided by the quantitative econometric literature on the behaviour of regulated actors and enforcers. Indeed, there is an increasing body of literature that empirically studies, with help of econometric methods, the variables determining compliance and monitoring and enforcement behaviour with respect to environmental regulation. Contrary to the literature presented in the previous chapter, these studies take a positive view, in that they do not ask 'what should be', but rather 'what is happening in reality'. The empirical focus implies that, compared to the normative studies, the econometric analyses take into account a larger number and a larger variety of variables thought to influence the behaviour of the regulated agents and the enforcers, and in particular include political variables. Furthermore, they open the 'black box' of the regulator.

The present chapter reviews those econometric studies, which are likely to give evidence of the effectiveness and efficiency of monitoring and enforcement directed at industrial plant. Specific questions to be investigated are:

- The normative theory suggests that monitoring and enforcement are the central political action variables to influence the behaviour of regulated actors in the desired way. Do the econometric studies give evidence for the effectiveness of monitoring and enforcement, i.e. does such activity create sufficient signals so that plants change their behaviour? And do the studies say anything about the relative effectiveness of monitoring and enforcement activity?
- Normative theory furthermore assumes that enforcers behave as welfare maximisers. Does the empirical evidence support the assumption that enforcers aim at efficient solutions? Or do they maybe follow personal or political objectives?
- Given that enforcement budgets are limited, are there alternative enforcement tools that might alleviate the enforcer's limited resources? How far are market forces and community pressure able to send out cost signals that induce firms to improve compliance, in addition to, or as a substitute of, traditional public enforcement? Does community pressure directed at enforcers provide information about the plants that cause pollution concern and thus help save scarce public monitoring resources?

Finally, this chapter also acts as a link between the mainly qualitative empirical investigation of the factors that drove compliance and enforcement behaviour in the implementation of the municipal waste incineration Directive in four European Member States and the econometric, quantitative investigation of the French compliance and enforcement path in the chapter that follows. Have comparable cases been investigated in the econometric studies reviewed here and what methodological issues are relevant to the econometric study of France?

The rest of this chapter is structured as follows. Section 2 discusses a basic methodological issue relevant to an estimation of compliance and enforcement behaviour and which is crucial to correctly determining the effectiveness of enforcement activity: the issue of compliance and enforcement decisions potentially being endogenous. Section 3 investigates whether empirical studies give indications of the effectiveness of monitoring and enforcement activity with respect to compliance and self-monitoring of regulated firms. It also asks whether monitoring and enforcement can be considered as substitutes, and goes on to discuss alternative enforcement tools. Section 4 presents a set of stylised objectives regulators may follow when enforcing environmental regulation and discusses their impacts on the occurrence and the efficiency of a possible implementation gap. This section also investigates whether there are any indications that the rules according to which enforcers apply monitoring and enforcement vary across industrial sectors or geographic zones. Section 5 concludes.

2 A methodological issue first: endogeneity of compliance and monitoring and enforcement decisions

Before analysing the econometric studies on compliance, monitoring and enforcement with respect to their empirical findings about what explains the behaviour of regulated actors and enforcers, it is worth outlining one methodological issue which has an impact on model results, and which is relevant for the econometric study in France. This has to do with the fact that the decisions of regulated agents and enforcers are not necessarily independent of each other.

2.1 An early study: compliance and monitoring decisions modelled as independent of each other

One of the first studies dealing with enforcement and compliance of environmental policy was published in 1990 by Magat and Viscusi. These authors investigate how inspections influence compliance, absolute emission levels and self-reporting of plants, hence testing the assumption underlying the monitoring and enforcement theory that inspections and enforcement raise the expected cost of non-compliance and by this provide additional incentives for pollution reduction and compliance. The dependent variables of interest are the absolute amount of emissions per plant and a dummy indicating whether or not the plant is in compliance. In order to measure the effect of the explanatory variable inspections on pollution and compliance, these dependent variables are regressed - amongst a series of other variables- against a dummy variable pertaining to whether the firm was inspected in the past or not. Given that it takes some time for a plant to bring itself into compliance after an inspection, it is often expected that it is not the current but rather past inspections that show an impact on compliance. Taking note of this, the authors use lagged inspection variables.

Important from a methodological point of view is that Magat and Viscusi (1990) model compliance as being dependent only on past inspections, but not on the plants' expectations about current or future inspections. This does not seem a very realistic assumption. Monitoring and enforcement are costly and agencies' budgets are generally assumed as being limited. Agencies are, therefore, not likely to inspect all plants in every period. Instead they will choose specific plants to be inspected, for example by focusing their resources on those plants they expect not to comply. If this were the case, an idea suggesting itself is that firms might develop expectations about future inspections, and if they did so, they would base their compliance decisions on the anticipated enforcement actions. Compliance would hence be endogenous. Given that one can assume both enforcers and plants to make their decisions based on the anticipated behaviour of the other, compliance as well as monitoring and enforcement decisions may both be interdependent and thus endogenous decisions. As endogeneity was not accounted for by Magat and Viscusi (1990), one can characterise their study as a naive compliance model.

2.2 Taking the endogenous character of compliance and enforcement decisions into consideration

From an economic point of view endogeneity arises when the decisions of different actors are determined in a non-cooperative game as suggested by Nash. In the context of enforcement and compliance decisions, enforcers and regulated agents then make decisions strategically by anticipating each other's behaviour. In a simple way, the problem can be formalised as in the following hypothetical example.

$$\text{COMPLIANCE} = f(X, \text{ENFORCEMENT}), \quad (1)$$

where X is a vector of independent variables.

Equation (1) indicates that a regulated agent makes his compliance decision based on a number of variables, including the expected action from enforcers. Similarly, enforcers determine their enforcement action, amongst other things, based on the compliance behaviour of the regulated agent, leading to the following hypothetical equation:

$$\text{ENFORCEMENT} = g(Y, \text{COMPLIANCE}), \qquad (2)$$

where Y is a vector of independent variables.

It is immediately obvious that equation (2) can be substituted into equation (1) and *vice versa*, leading to equations (3) and (4), illuminating the issue of the endogeneity of decisions.

$$\text{COMPLIANCE} = f(X, g(Y, \text{COMPLIANCE})) \qquad (3)$$

$$\text{ENFORCEMENT} = g(Y, f(X, \text{ENFORCEMENT})) \qquad (4)$$

(3) and (4) indicate that expected compliance (enforcement) behaviour determines, other things being equal, enforcement activity (compliance), which determines compliance (enforcement).

Where decisions are interdependent this has impacts for an analysis of the variables. One may easily find a negative correlation between enforcement and compliance when examining the correlation between these two variables without taking account of their endogeneity. This is, for example, a likely result when enforcers target non-compliant plants. Not taking account of the interdependency between variables by not modelling these as endogenous can thus bias econometric estimation results and make it difficult to empirically observe the ability of enforcement to increase compliance.

In a study taking an approach similar to Magat and Viscusi (1990) but increasing the level of sophistication, Laplante and Rilstone (1996) address the possibility that inspections are endogenous. Like Magat and Viscusi, the authors are primarily interested in the impact of monitoring and of a threat of inspections on self-reporting and self-reported emission levels. Tested is the impact of inspections on two sets of variables: absolute discharges and the level of discharges relative to their respective standards.

Interesting in this study is that it allows comparing results of a model calculating least-squares estimates without and with controlling for the possibility of inspections being endogenous (cf. Table 9.1).[23] The authors account for the possibility that decisions about inspections may be endogenous by estimating inspections as a function of variables used in the basic pollution equation and of a variable indicating the number of prior inspections, which have been conducted at the plant. The approach is based on the assumption that the probability of an inspection may be inversely related to the number of previous visits and that the regulator's current decision to inspect might also be affected by the plant's emission level. Put in other words, the authors assume that both inspections and the probability of an inspection may have an effect on emissions.

[23] The general econometric approach to a test for endogeneity of decisions uses two-stage estimation. In the first stage decisions are predicted by regressing the actual experience on a set of instrument variables. Instrument variables are exogenous variables that do not enter the second stage model. The predicted values are then used in the structural specifications in the second stage.

Table 9. 1: Impacts of inspections on BOD discharges in models of differing sophistication

Model and independent variables	Coefficients	Impact
	Basic model	
Current inspections impact of (at least 4-months lagged) past inspections	-4.7 -1.3	 - 7%
	Model controlling for endogenous decisions	
Current expected inspections Lagged/past inspections	-193.4 -5.4	 -28%

Source: Laplante and Rilstone, 1996

The table shows that control for endogeneity of decisions increases the coefficient estimates on both current and lagged inspections: the basic model without control for interdependence of decisions suggests that inspections, lagged by 4 moths or more, reduce absolute discharges of biochemical oxygen demand (BOD) by approximately 7%, while in the model controlling for interdependence, past inspections reduced BOD discharges by approximately 28%. Differences in size are also reflected in the coefficients for the effect of current and past inspections on absolute discharges: They are -4.7 and -1.3 respectively in the basic model, while they amount to -193.4 for current and -5.4 in the model controlling for endogenous variables. In the more sophisticated second model the authors find a significant negative and persistent effect on emissions, which is hence stronger than in the previous model. Note that here the variable 'current inspections' has been replaced by the conditional inspection probability. Since inspections (in the basic model) were replaced by expected inspections (in the model controlling for endogeneity), an alternative interpretation of these results is that the threat of an inspection may have the strongest effect on emissions. In any case the results indicate that polluting behaviour is also a function of the probability of inspection, not only of actual inspections (Laplante and Rilstone, 1996: 30-31). This finding is also in support of what was suggested by the normative theory on monitoring, enforcement and compliance: that it is p, the probability of inspection, together with F, the stringency of the penalty, which plays a role for compliance. The reason is that the probability of monitoring directly influences the expected penalty, and with this the expected costs of non-compliance.

2.3 Specific compliance solutions and enforcement may also be endogenous

As was shown in chapter 6, compliance with environmental regulation can not only be reached through investment in abatement equipment (or process technology or input changes) but also by shutting down operations of a plant. Plant-closure proved not to be a marginal compliance option both in France and the UK. Therefore, a further study of specific interest to the econometric analysis of the compliance and enforcement path France followed in the implementation of the MWI Directive was provided by Gray and Deily (1996). These authors take account of interdependencies which can arise with respect to plant closure decisions. They study the relationship between the regulator's enforcement decision and firms' compliance and closure decisions and with this go one step further than the previously presented study. With their empirical case relating to air pollution regulation towards integrated US steel plants, and thus to a declining industry, Gray and Deily (1996) assume that compliance decisions in reaction to enforcement may well consist in plant closure. Consequently, they assume that compliance and enforcement decisions may be also intertwined with plant-closure decisions, and model the three variables, compliance, closure and enforcement, as endogenous decisions in three equations in two stages. In the first stage, predicted versions of each decision are generated by regressing actual experience on a set of instruments. The compliance variable captures the aggregated compliance status for a year by taking into account the status in every quarter of a year. The enforcement variables cover inspections, on the one hand, and all enforcement measures (inspections and fines), on the other hand.

As shown in Table 9.2, the authors do find the expected interactions: as a first result, greater lagged enforcement (whether measured as inspections or total actions) leads to greater compliance at a plant level, while greater (predicted) compliance leads to less enforcement. One specification is however necessary: While lagged actual enforcement shows significant effects on compliance, predicted current enforcement does not. While no explanation was offered for these results, one may imagine that effects of predicted current enforcement are not immediately obvious because bringing a plant into compliance takes time. As a second result, predicting values also for compliance and plant closure allows to establish an interdependence of the three decisions. The predicted probability that a plant will close during industry contraction shows a negative impact on compliance which indicates that firms were less likely to bring marginal plants into compliance before shutting them down. Furthermore, the negative impact of predicted plant closure on enforcement suggests that enforcers consider the fact that a plant may be forced to shut down as a result of enforcement action and adjust their activity accordingly (Gray and Deily, 1996).

Table 9. 2: Estimation results of second-stage estimations

Independent variable	Dependent variable
	Compliance
Predicted total enforcement actions	Positive coefficient not significant
Lagged total enforcement actions	Positive impact Significant
Predicted number of inspections	Positive coefficient Not significant
lagged inspections	Positive impact Significant
Predicted probability that plant will close	Negative impact Significant
	Enforcement
Log of number of inspections	Positive impact Significant
Log of number of total enforcement actions	Positive impact Significant
Predicted probability that plant will close	Negative impact Mostly significant
Predicted compliance	Negative impact Significant for inspections and total enforcement actions

Source: Gray and Deily, 1996: 107

3 What do the empirical studies say about the effectiveness of monitoring and enforcement?

Let's now turn to the empirical findings of the econometric studies. Do they provide evidence that monitoring and enforcement help to bring plants into compliance with environmental standards? Do their results improve when the expected behaviour of actors is taken into consideration? Is there any evidence that monitoring and enforcement are substitutes as suggested by Becker (1968)? What is the effect of inspections on plant's self-reporting behaviour? And do the studies distinguish alternative enforcement tools?

3.1 Do monitoring and enforcement induce compliance with environmental standards?

For an investigation of which factors may explain an implementation gap it is interesting to know whether monitoring (i.e. inspections) or enforcement (for example penalties) actually help to bring plants into compliance. Several empirical studies do assess enforcer activity relative to compliance results.[24] The industrial sectors, countries, regulatory issues and independent variables studied are summarised in Table 9.3. While some authors focus on either monitoring or enforcement and their effect on compliance, others study both political action variables. The majority of studies analyse the pulp and paper industry. Striking is also the focus on North America, and primarily the USA.

[24] A number of further studies investigate the effect of monitoring and enforcement on the environmental performance of plants, but not on actual compliance (e.g. Laplante and Rilstone, 1996; Pargal et al., 1997).

Table 9. 3: Industrial sectors, countries and variables whose effect on compliance is studied

Author	Industrial sector	Country	Regulatory issue	Independent variable
Magat and Viscusi (1990)	Pulp and paper industry	USA	Effluent standards	Inspections
Helland (1998a)	Pulp and paper industry	USA	US Clean Water Act of 1972	Inspections
Foulon et al. (2000)	Pulp and paper industry	British Columbia/Canada	Water regulation	enforcement actions
Gray and Shadbegian (2000)	Pulp and paper industry	USA	Air pollution regulation	Inspections and enforcement
Nadeau (1997)	Pulp and paper industry	USA	Water and atmospheric regulation	Inspections and enforcement
Gray and Deily (1996)	Steel industry	USA	Air pollution regulation	Inspections and enforcement
Shimshack and Ward (2004)	Pulp and paper industry	USA	Water regulation	Inspections and fines
Shimshack and Ward (no date)	Pulp and paper industry	USA	Water regulation	Inspections and fines

The overall picture is quite homogenous: all studies presented find that enforcement action is more or less effective in triggering compliance behaviour with environmental standards.[25] The focus here will therefore be on the 'more or less', i.e. the question what the found results depend on in detail.

Two studies come to unambiguous results about the effectiveness of past inspections and of the probability of inspections in increasing plant level compliance. These are, firstly, Magat and Viscusi (1990) who find -in their naive compliance model shortly discussed in sub-section 2.1- that past inspections increase plant level compliance with a one quarter lag. Secondly, in Helland's (1998a) case, inspections three quarters ago, inspections one year ago, as well as the probability of inspections decrease violations. The latter two variables actually always reach significance levels of 1%. The results of the remaining studies require a more detailed investigation.

Gray and Shadbegian (2000) assess the impact of actual and predicted (two-year lagged) enforcement on the compliance status of plants. These authors capture enforcement activity in three variables, one covering the total number of actions directed towards a plant during a year, one covering only inspection actions (inspections, emission monitoring, stack tests, etc.), and a final one covering enforcement activity (notices of violation, penalties, phone calls, etc.). The results support what was discussed under section 2 above: not controlling for the fact that variables may be endogenous may on first sight yield counter-intuitive results indicating that those plants facing greater actual enforcement activity tend to have a higher probability of being non-compliant. This still holds when the authors use lagged actual enforcement faced by a firm. That these results rather reflect a targeting strategy of enforcers than perverse effects of enforcement, or counterproductive enforcement, is suggested by the results of their more sophisticated model applying an instrumental variables method to dealing with the problem of endogenous decisions. This method generates a predicted value of enforcement (which was also lagged) and which was used in a second stage estimation. As a result, the authors now find a positive and significant impact of 2-year lagged predicted enforcement, as well as of the probability of enforcement, on compliance rates. Raising the probability of inspection or the probability of enforcement action by one is found to increase the probability of a plant being in compliance by approximately 10%. Gray and Deily (1996) come to an inverse result. Also estimating the impact of either inspections

[25] One exception is Helland (1998b), not presented in the following, because he investigated the effect only of those past inspections on violations that did not detect a violation. The result that these inspection variables did not yield significant results is therefore to be taken with caution.

or total monitoring and enforcement activity (inspections and fines) on compliance, they find that lagged enforcement shows positive and significant coefficients, where the effects of total monitoring and enforcement actions are stronger than those of inspections only, whereas predicted enforcement does no longer yield significant results (Gray and Deily, 1996). Unfortunately, the authors do not offer any explanation for why predicted variables do not yield significant results.

It might be that Nadeau's (1997) findings could shed some light on this issue. This author shows how important it is to carefully choose the sample from which to predict monitoring and enforcement behaviour. He focalises on a specific aspect of the effectiveness of monitoring and enforcement by studying the enforcer's ability to reduce the duration plants spend in non-compliance with air pollution regulation and does so by using parametric survival models in the US pulp and paper industry.[26] In his case an agency is assumed as being effective at reducing the duration of non-compliance if monitoring and enforcement significantly increase the probability of plants exiting the non-compliance status. Nadeau (1997) studies two types of enforcer activity: enforcement actions, such as administrative, judicial, civil and penalty actions, and monitoring activity, consisting of inspections and tests. Variables for both activities use predicted values.[27] Monitoring and enforcement are predicted, alternatively from the full sample of plants and from a sample including only non-compliant plants. In a third model, enforcement activity is predicted from the non-compliant sample and monitoring from the full sample. The rationale behind this is that enforcement actions are generally taken in response to a violation, while monitoring might rather be done at all plants. When predicted from the full sample, monitoring and enforcement were insignificant in the second stage, while they were significant when predicted from the sample of non-compliant plants.[28] This indicates that the enforcer does indeed follow separate strategies based on the previous compliance status of plants, and that compliance, monitoring and enforcement were endogenous. For the sample of non-compliant plants the study finds that the enforcer was effective in reducing the time a plant spends in non-compliance: a 10% increase in monitoring activity led to a 4.2% reduction in the time plants spend in non-compliance, and a 10% increase in enforcement activity implied a 4% - 4.7% reduction in the length of non-compliance.

An interesting aspect of enforcement was studied by Shimshack and Ward (2004 and no date): the effectiveness of fines to not only reduce violations by the fined firm but also by other firms in the same regulatory jurisdiction. Their results do give evidence of the effectiveness of fines both to increase compliance of the fined firm and also to reduce the non-compliance rate of other firms in the same jurisdiction. Furthermore, these authors find that additional lagged inspections also somewhat deter non-compliance but that their deterrence effect is smaller and also decays more quickly than that of fines.

Finally, Foulon et al. (2000) investigate the effectiveness of enforcement activity (but not of inspections) with respect to plant compliance. Enforcement activity faced by a plant in a given year is captured by two variables, one indicating the number of prosecutions, the other the absolute amount of fines. The study does not yield any significant results for the impact of prosecutions on compliance rates. However, actually imposed lagged fines partly show a negative impact on compliance.

---

[26] Such models are used when the issue is to make interference about the length of time plants spend out of compliance. They estimate the probability of remaining in a state of existence for t periods, and thus allow studying the probability that a firm becomes compliant in t, given that it has been non-compliant for t periods. The method allows making interference about how monitoring and enforcement activity affects the probability of becoming compliant (Nadeau, 1997: 64).

[27] These are estimated in two-stage models accounting for simultaneity of compliance and enforcement decisions. First-stage estimations of monitoring and enforcement activity generate instruments used in the second-stage survival models.

[28] In the third model, where enforcement activity was predicted from the non-compliant sample and monitoring from the full sample, only predicted enforcement yielded significant, negative results.

3.2 Are monitoring and enforcement substitutes?

The findings presented above give indications of the effectiveness both of monitoring and of enforcement to increase compliance with environmental regulation. Becker (1968) had suggested that, as far as risk-neutral regulated agents are concerned, monitoring and enforcement were perfect substitutes with respect to their deterrence effect. In order to save overall social costs of monitoring and enforcement, this author recommended to increase the penalty relative to the monitoring probability. Is the claim of monitoring and enforcement being perfect substitutes supported by the empirical literature that has been presented so far, i.e. are there indications that regulated agents react similarly on increases either in the monitoring probability or in enforcement? And do enforcers seem to increase penalties relative to the monitoring probability?

Unfortunately, no general results can be derived from the empirical cases: partly because the majority of studies investigate the impacts only of either monitoring or enforcement on the plant's compliance behaviour. Partly because the few studies that investigate both instruments come to contradictory results. Indeed, Gray and Deily (1996) find a stronger effect of total monitoring and enforcement actions (including inspections and enforcement) on the firm's compliance than of monitoring only. This, however, does not come as a surprise, given that they compare the effect of one type of action against that of the totality of monitoring and enforcement actions. In order to compare the relative impacts, the effects of enforcement would have had to be investigated in isolation as well. In Nadeau's (1997) study enforcement yielded slightly higher effects than monitoring, and Shimshack and Ward (no date) found that at the margin a fine was about 4 times as effective at inducing compliance as an inspection. Opposite to these results, Gray and Shadbegian (2000) found comparable impacts of both inspections and enforcement activity.

But there is yet another reason for why it is difficult to draw conclusions about whether or not monitoring and enforcement are substitutes even from the latter three studies. What Becker (1968) suggested was to increase enforcement while simultaneously decreasing the monitoring probability. However, no study was able to estimate effects of increasing either variable while proportionally decreasing the other.

3.3 What is the effect of monitoring and enforcement activity on plants' self-reporting behaviour?

The normative monitoring and enforcement literature has made further suggestions about how the administrative costs of enforcement could be reduced. One of these is to involve firms into the monitoring of their emissions or compliance status and to oblige them to report the results to the inspectors. To the enforcer, self-reporting by plants is a low-cost alternative to increase his informational basis about the plants' environmental performance and compliance status, which can have direct implications for the allocation of monitoring resources. Particularly in a setting of limited monitoring and enforcement budgets this can help improve the efficiency of enforcement decisions. Of course, this depends also on the resources enforcers need for processing the firms' self-monitoring reports, and on the quality and correctness of the issued reports, but to this day the quality issue has not been the primary focus of econometric studies. A further interesting question, nevertheless, is whether monitoring and enforcement activity positively influence the plants' self-monitoring behaviour in terms of the frequency of reporting.

A number of empirical studies try to assess in how far inspections can induce firms to comply with reporting requirements. These establish that inspections do indeed drive self-reporting of plants. Magat and Viscusi (1990) study a case where the industry was required to regularly report their pollutant discharge levels and where inspections constituted not so much the primary basis for checking compliance with the regulation, but rather a secondary source of information, permitting to test whether firms reported accurately. While these inspections should therefore be an incentive to correctly report, the authors do not study the quality aspect of self-reports. Testing the impact inspections had on the frequency of reporting these authors find that inspections did indeed serve the purpose to induce firms to report their discharge levels regularly: the effect of inspections on reporting was statistically significant and positive. The impact of monitoring on self-reporting was also studied by Laplante and Rilstone (1996). Predicting the probability that a plant reports in a simple binary choice model of reporting as a function -amongst other things- of cumulated inspections, the authors find that inspections have a strong effect on reporting. Also in this investigation, the act of reporting therefore was not random.

A positive impact of inspections on self-reporting behaviour of plants is also found by Helland (1998b). This author yields partly significant results for a positive effect of one, two or six quarter lagged inspections. It is not astonishing that those inspections which actually detected a violation yielded stronger effects than inspections which did not detect any violation[29]. While the effect of one-quarter lagged inspections (that did detect a violation) on self-reporting is greater than that of two-quarter lagged inspections, a comparatively high coefficient was found for the six-quarter lagged inspection variable. The author interprets this as a proof for his assumption that firms submit reports in order to signal the enforcer their willingness to co-operate and primarily do so (here over a year and a half) in cases where violations are particularly difficult to correct.

3.4 Alternative enforcement tools

While the more traditional view of regulation has focused on state-plant interactions, a rapidly growing number of more recent empirical studies suggest that community pressure or market forces can serve as an alternative, or complement, to more traditional formal monitoring and enforcement activity. The general idea is that neighbouring ***communities*** may have a powerful influence on the environmental performance of plants. Where formal regulators are present, they may actually use the political process to influence the tightness of enforcement, or citizens may directly complain towards enforcers. Where formal regulators do not exist, informal regulation may be implemented through community groups or NGOs (Afsah et al., 1996). As far as ***market influences*** are concerned, these can be multiple. The impact of firm specific environmental news on the image or the market value of firms may work its way through various channels by affecting decisions of suppliers, investors, stockholders as well as clients and consumers. Foulon et al. (2000) suggest that information disclosure programmes, revealing the environmental performance of polluters, can notably play a role as alternative enforcement means where agency budgets are limited and enforcers are reluctant to use stringent enforcement measures. By giving roles to additional actors, markets and the community, the enforcer can gain leverage in that such instruments are able to put additional pressure on firms, notably reputational, financial and judiciary pressures.

Empirical studies of community pressure and market forces frequently focus on developing countries. In that they try to assess whether market forces and community pressure may constitute an alternative to formal public enforcement in countries characterised by lower environmental standards and/or poor enforcement of environmental regulations. The majority of studies in this area estimate the impact of community pressure or public disclosure programmes on environmental performance, without assessing the impact on compliance (cf. for example Wang, 2000; Afsah et al., 2000). Similarly, studies assessing stock market fluctuations as reaction to the publication of positive or negative firm-related environmental information do frequently not go as far as estimating whether these market reactions eventually lead to environmental improvements (cf. for example Lanoie et al., 1997; Dasgupta et al., 1997). Presented below are only the results of studies which use econometric methods to assess the influence of community pressure, market forces and information disclosure on plant compliance with environmental regulation, or the impact of community pressure on

[29] Inspections that did not detect a violation were only partly significant with a one-quarter lag.

enforcer activity. Two of these focus on plant compliance (Foulon et al., 2000; Dasgupta et al., 2000), while a third focuses on enforcement as dependent variable (Pargal et al., 1997). Table 9.4 indicates which industrial sectors and countries are covered by the studies.

Table 9. 4: Industrial sectors, countries and dependent variables studied

Author	Industrial sector	Country	Regulatory issue	Dependent variable
Foulon et al. (2000)	Pulp and paper industry	British Columbia/Canada	Water pollution regulation	Compliance
Dasgupta et al. (2000)	Manufacturing industry	Mexico	Environmental regulation	Compliance
Pargal et al. (1997)	Industrial plants	India	Water pollution	Enforcement

Foulon et al. (2000) are interested in the performance of public information disclosure programmes relative to inspections and penalties with respect to the compliance behaviour of polluters in Canada. Here, the Environment Ministry regularly publishes lists of firms that do either not comply with existing regulation or whose environmental performance causes concern. Additionally, the Ministry takes legal action for firms violating regulation. Foulon et al.'s (2000) model describes compliance, amongst other things, as a function of traditional enforcement (number of prosecutions faced in a year, and total amount of fines imposed per plant in a given year), and public disclosure (listing of non-compliance, concern). Model estimations show that the public disclosure strategy, particularly a listing on the 'out of compliance' list had an additional and larger impact on the compliance status than traditional enforcement measures (orders, fines and penalties). Also, this impact was more immediate: while fines had a lagged influence only, the contemporaneous 'out of compliance variable' was significant. The authors, thus, suggest that lists of polluters may provide a stronger incentive than conventional enforcement measures for a quick response to damage. Nevertheless, they do not believe that information strategies can necessarily replace traditional enforcement practices and consequently argue in favour of a mix of both approaches, where information disclosure would be used as a complementary policy instrument. The impact of community and market pressures is also studied by Dasgupta et al. (2000). These authors try to assess whether next to formal regulatory activity local community action, the plant status as 'publicly traded', a multinational ownership, and export links to the OECD show impacts on emissions (reductions) and with this on compliance (increase) for plants belonging to Mexico's manufacturing industry. The results are striking in that all variables indexing OECD linkages, foreign ties and indirect community pressure had no significant effect on greater environmental effort. The only variable showing a positive significant effect is the public trading of the firms' stocks, from which the authors conclude that public scrutiny does in fact promote stronger environmental policies.

A further study which takes into account community pressure is provided by Pargal et al. (1997). Using district development as a proxy for community pressure these authors analyse its impact on inspections and emissions in India. Underlying this is the assumption that the valuation of damage, the demand for environmental quality and the community negotiation power increase with the level of development. More developed areas are then expected to be more likely to complain or make use of political or other channels to pressure regulators to take action against polluters. Also, neighbours and NGOs might try to pressure plants directly. Community pressure is also expected to affect the potential penalty firms face for pollution, and by this the emission level. The authors find that while district development shows a strong positive relationship with inspections, there are no significant results for its impact on emission levels. Contrary to this, in a survey conducted by the authors, plants identified pressure from neighbours and NGOs as important in inducing abatement. From this the authors conclude that the community pressure that exists might be rather channelled through the formal enforcement mechanism than through direct negotiation with plants. The result may also be affected by the plant sample, which includes no small but only medium sized and large plants, which may be less susceptible to direct pressure. What the study does however suggest is that community pressure may indeed induce enforcers to inspect plants more frequently.

In summary, it has so far been difficult to give econometric evidence for direct positive effects of market forces and community pressure on plant compliance, while strong effects could be shown for information disclosure strategies. Whereas Foulon et al. (2000) show the effectiveness of public information disclosure programmes in increasing plant compliance, Dasgupta et al. (2000) find only weak indications that market pressure might improve compliance. Pargal et al. (1997) finally believe that community pressure in India was rather directed at enforcers than directly at polluting firms, and that it induced enforcers to inspect plants more frequently.

3.5 Summary

Summing up, the majority of studies yield results in support of what is suggested by the normative theory on monitoring and enforcement. Both inspections and actual enforcement action, in the majority of cases, are shown to improve compliance, even if model issues are crucial to specific results. Few significant results were only found in the Foulon et al. (2000) study. Inspections are furthermore found to induce self-reporting by plants, which is important in that it can help increase the information available to enforcers in a cost-saving way. Of primary interest with respect to an implementation gap, however, is the regulator's effectiveness in increasing compliance. From what was said so far it can be concluded that monitoring and enforcement are indeed effective tools in increasing firm compliance with environmental regulation.

Nevertheless, a number of the studies reviewed in the present chapter provide evidence for the occurrence of a certain implementation gap by giving indications on the compliance rate of plants in their sample. To give just some examples, Helland (1998a) finds a non-compliance rate of 11% for the US pulp and paper industry with respect to the Clean Water Act, while Magat and Viscusi (1990) find a non-compliance rate of 25% for the same industry. Harrison (1995) reports a compliance rate for Canadian paper mills of 69%. These data indicate that there is indeed a discrepancy between initial objectives and final outcomes of environmental policies studied by the econometric analyses.

In view of the overall finding that the enforcer's activity improved compliance, it can be assumed that it is not the effectiveness of the tools monitoring and enforcement *per se*, but rather the selection of plants that are targeted or not by monitoring and enforcement activity which may explain an implementation gap. This means that overall compliance of firms can be expected to depend on whether or not the tools monitoring and enforcement are actually applied. Whether the econometric studies support the assumption that inspectors monitor plants, and enforce policy, only in a selective way, and according to which patterns they use these tools, is investigated in the following section.

But next to the question of whether monitoring and enforcement are complete or not (i.e. whether they contribute to an implementation gap), for an economic assessment of enforcer behaviour efficiency considerations also need to be taken into account. As Nadeau (1997: 77) points out, effectiveness of enforcement actions does not necessarily imply that an increased monitoring and enforcement activity is desirable. It is desirable only as long as benefits in the form of reduced harm to the environment and health are not outweighed by increasing compliance costs to plants and increasing monitoring and enforcement costs to the administration. What empirical studies have to say about the efficiency of patterns of enforcer behaviour is a further issue of the following section.

4 According to which rules are monitoring and enforcement applied in practice?

The normative economic theory models enforcers as maximising welfare when implementing environmental legisltaion. However, already Harrington's (1988) model presented in chapter 8 diverges from pure welfare maximisation by modelling a cost-effective instead of an efficient monitoring and enforcement scheme. More empirical studies further suggest that enforcers might not always strictly follow what is written in regulations. Instead, they might be sensitive to particular problems faced by a plant when trying to reach compliance, or they might have purely personal (political) objectives, which guide their behaviour. It has also been widely acknowledged that enforcement budgets are limited. This implies that regulators may not be able to monitor all plants and force all regulated agents into compliance, even if they wanted to. Instead, they have to allocate their resources to perform a limited number of enforcement activities. How they do this will be crucial to environmental and compliance outcomes and to the efficiency of the policy objectives actually implemented.

The objective of this section is to analyse according to which rules enforcers allocate their resources, whether these rules can explain an implementation gap, and which factors determine the regulator's decision-making. Are there empirical results suggesting that enforcers follow efficiency considerations, seeking to maximise welfare, when allocating their enforcement resources? And if enforcers follow welfare considerations, are there indications that imperfect monitoring and enforcement may actually make regulatory output more efficient than the objectives embodied in laws? In this context it should be recalled that imperfect compliance resulting from imperfect enforcement is not *a priori* inefficient from an economic point of view. This is particularly the case when the policy involves command-and-control regulation, a policy instrument that has been criticised by economists for its uniform character, thus failing to account for the polluters' heterogeneity. But as enforcement resources are limited, uniform legal requirements do not automatically entail uniform environmental responses in the field, as was shown in chapters 6 and 8. Therefore, it is in order to also investigate whether econometric studies give indications that imperfect monitoring and enforcement, in reality, lead to improvements in the efficiency of policy objectives.

4.1 A stylised set of behavioural patterns that may underlie enforcer behaviour

Before entering into a presentation of econometric studies that investigate according to which criteria enforcers allocate their limited enforcement budgets -generally referred to as 'targeting'- it is useful to classify possible behavioural patterns of enforcers. Such a typology was suggested by Dion et al. (1998), who distinguish three 'pure' types of targeting (cf. Box 9.1).

Box 9. 1: Targeting patterns according to Dion et al. (1998)

Firstly, the enforcer may be a ***'benevolent maximiser of welfare'***, as assumed by the normative environmental economics literature. The welfare maximiser can be expected to allocate his resources to those plants that have low marginal abatement costs[30] or that cause high environmental damages, in order to reach the maximum environmental benefit at minimum compliance cost.

[30] Note that Dion et al. (1998) believe the actual sign of the relationship between plant compliance costs and enforcement activity to be an empirical issue which can vary. According to these authors, the regulator might also direct his resources towards high cost plants which may have a greater incentive to avoid compliance.

Box 9.1 continued

Secondly, there is what one may call the ***'bureaucratic enforcer'***. He takes the law as the truth, maximises the rate of compliance, and concentrates his enforcement efforts on those plants that are less likely to comply, for example because of their particularly high compliance costs. This may also imply monitoring and enforcement activity in dependence of the compliance history of plants. Such a strategy would presume that compliance is equally desirable regardless of the impact of the plants' emissions on the environment, and regardless of the firms' compliance costs.

Finally, the enforcer may be ***'politically biased'***, seeking, instead of welfare maximisation, to maximise certain (personal) political objectives or (net) political support. The reference model here is the positive theory of regulation initiated by Stigler (1971) and Peltzman (1976). If the enforcer is politically biased, he will target polluters that are less likely to reward him with political benefits, or, put the other way round, less likely to politically harm him. Alternatively, the enforcer might also try to gain support from an environmentally aware community, which may result in 'visible' monitoring and enforcement actions.

As to the efficiency impacts of the stylised behavioural patterns, one can assume that under an unbiased ***welfare maximising strategy***, provided the enforcer is not constrained by any informational or budget limits, the resulting implementation gap would be economically efficient. Opposite to this, an implementation gap resulting from the ***maximisation of personal or political objectives*** can be expected to be inefficient in an allocative sense of balancing costs and benefits of environmental regulation.[31] The outcome may however be compatible with normative macro level political objectives, for example if enforcers tolerate non-compliance of large employers in high-unemployment areas who would cede operations, and make redundant their employees, as a result of enforcement measures. Finally, note that while the ***'bureaucrat'*** aims at minimising the implementation gap, such a gap may still occur where the enforcement budget it limited. But where compliance maximisation is successful, resulting in complete compliance, this result is not necessarily economically efficient.

Before presenting throughout the following sub-sections evidence found for the different types of enforcer behaviour, we list in Table 9.5 the types of variables frequently studied to test for the influence of alternative enforcer behaviour. Note that not all variables unambiguously relate to one specific type of enforcement pattern. A number of variables may support different or even contradictory hypotheses and interpretations in specific studies. Ambiguous are above all the variables compliance history, the impact of environmental preferences, and variables relating to plant employment.

The empirical studies, firstly, differ in their interpretations of compliance history, i.e. whether they assume enforcers who react to this variable to be driven by concern about efficiency or rather by a compliance maximising behaviour. It could be judged as efficiency enhancing if interpreted as a focus on plants that repeatedly violate standards and therefore cause enduring and with this potentially increasing environmental damage.[32] But it could also indicate bureaucratic behaviour, if interpreted as a focus on previously non-compliant plants without paying attention to related benefits and costs.

Secondly, ambiguity between variables giving evidence either for welfare or politically biased enforcer behaviour concerns environmental preferences and plant employment. While high plant employment in the majority of studies is interpreted as a proxy for the political power of plants vis-à-vis enforcers, there is also one study which interprets it in the sense of a firm's viability to pay for compliance. Indications that enforcement measures are in line with environmental preferences, finally, are frequently interpreted as enforcers taking account of perceived benefits of pollution reduction. However, there are

[31] Except if the regulator reacts to local preferences and pressures which reflect efficiency considerations.

[32] This obviously depends also on the type of pollutant, for example whether it accumulates in the environment.

also cases where proxies for environmental preferences are interpreted in enforcers being sensitive to political support or opposition by their constituencies.

Table 9. 5: Variables and their interpretation in terms of targeting strategies

Independent variable	Expected sign of influence on monitoring and enforcement
Efficiency consideration/welfare maximisation	
Damage variables	
Increasing degradation of environmental quality	+
Environmental and health damages	+
(Repeated) violation of standards (and hence compliance history) (1)	+
Increasing exposure of population or ecosystems to pollution	+
Environmental preferences/willingness to pay for improved environment (2)	
(Average) household income	+
Level of community development	+
Size of environmental movement	+
Abatement costs (often reflected in indirect variables)	-
Plant age (inverse to plant vintage, important for remaining life span of plant)	-
Plant capacity (economies of scale)	+
Firm's capacity to pay for abatement	
Firm's economic viability (sometimes represented by employment) (3)	+
Politically biased behaviour	
Firm's capacity to afford the necessary abatement investment (3)	+
Inability to pay abatement (Risk of plant closure and related employment effects) (3)	-
Plant employment (3)	-
Unemployment rate of regions	-
Resistance by plants towards enforcement	
Often indicated by compliance costs	-
Political representatives on higher administrative/political levels	Sign depends on specific case and theory tested
Environmental preferences/willingness to pay for improved environment (2)	
(Average) household income	+
Level of community development	+
Size of the environmental movement	+
Bureaucratic/compliance maximising behaviour	
Compliance history (1)	
Preceding non-compliance	+
Previous compliance	-

(1) Basic assumption behind the enforcement behaviour of a compliance maximiser, but sometimes also interpreted in terms of efficiency, assuming that plants repeatedly violating standards cause higher environmental damage. - (2) Sometimes interpreted as indicating benefits of pollution reduction, sometimes as political influence in terms of pressure on enforcers. – (3) Sometimes interpreted in terms of efficiency, sometimes in terms of political opposition to be expected.

4.2 What are the enforcers' objectives?

Dion et al.'s (1998) classification of possible enforcer behaviour raises the question of whether such behaviour, in practice, is unambiguous. Do the empirical studies suggest that enforcers follow unbiased objectives? The answer is no. Amongst the studies reviewed here, there is not one example indicating that enforcers follow either only political concerns, only efficiency concerns, or objectives solely based on previous compliance history.[33]

To start with, it seems sensible to give a short overview over the hypotheses tested by the different econometric studies that investigate the factors influencing enforcement behaviour (cf. Table 9.6). The authors do in fact focus on quite different and specific questions. Moreover, they frequently include control variables which could support enforcer behaviour alternative to the one explicitly tested. The table gives an overview

[33] In the following, support for a certain type of regulatory behaviour means that at least one of the respective independent variables tested yielded significant results. It does not mean that all variables which were tested and which could indicate the specific enforcer behaviour have in every case led to significant results.

over the explicitly tested assumptions and over other variables included that allow for alternative interpretations of enforcer strategies. As can be seen, a number of studies explicitly test for competing behavioural theories of enforcers, such as Dion et al. (1998), Oljaca et al. (1998) and Kleit et al. (1998), while others focus primarily on one type of explanation (Dion et al., 1998; Gray and Deily, 1996; Helland, 1998a; Mixon, 1995; Pargal et al., 1997). Two studies, finally, explicitly analyse the firm-regulator relationship (Helland, 1998b; Oljaca et al., 1998), an issue which cannot easily be subsumed under the 3 stylised patterns of enforcer behaviour suggested, but probably comes closest to a compliance maximising strategy.

Table 9. 6: Explicitly and implicitly estimated patterns of enforcer behaviour

Source	**Political behaviour**	**Welfare considerations**	**Bureaucratic behaviour**	**Co-operative firm-regulator relationships**
Dion et al. (1998)	Explicitly: economic theory of regulation (Stigler, 1971)	Explicitly: public interest theory of regulation (Posner, 1974)	Implicitly: past compliance status	-
Gray and Deily (1996)	Explicitly: political support maximisation (Stigler, 1971; Peltzman, 1976)	Implicitly: damage variables included, compliance costs	Implicitly: past compliance status	-
Helland (1998a)	Explicitly: political support maximisation (Peltzman, 1976)	Implicitly: major past violations (proxy for damage)	Implicitly: probability of violations	-
Helland (1998b)	Implicitly: shut-down point, profit, gross return	Implicitly: significant violations, effluent discharges (proxy for damage)	Explicitly: Harrington (1988) model, state-dependent pollution control enforcement	Explicitly (part of model): self-reporting to change inspection status
Oljaca et al. (1998)	Implicitly: company size	Explicitly: does the extent of violations affect the penalty? (proxy for damage)	Explicitly: does compliance history matter?	Explicitly: do co-operative relationships increase the weight given to firms' costs and reduce the fine?
Mixon (1995)	Explicitly: regulatory capture/business interest group influences/rent seeking (Stigler, 1971, Peltzman, 1976, Tollison, 1982 and Tullock, 1989); Party/Cabinet influence through EPA on states	Implicitly: damage variable	-	-
Kleit et al. (1998)	Explicitly: legislative dominance theory (Weingast and Moran, 1983); rent extraction theory (McChesney, 1987 and 1991)	Explicitly: (legal and) environmental severity of violation	Implicitly: past enforcement actions and violations	-
Pargal et al. (1997)	Explicitly: are monitoring and enforcement affected by community characteristics acting as proxies for political power	Implicitly: use of water for bathing or industrial purposes (proxy for damage); age of plant (possible proxy for compliance costs)	-	-

Table 9.7 summarises the sectors and countries to which the data studied in the econometric models refer, as well as the variables whose occurrence the authors wish to explain.

Table 9. 7: Industrial sectors, countries and variables explaining enforcement studied

Author	Industrial sector/polluters	Country	Dependent variable
Oljaca et al. (1998)	Industry in general	Georgia/USA	Amount of paid fine
Dion et al. (1998)	Pulp and paper industry	Canada	Number of inspections
Gray and Deily (1996)	Integrated steel industry	USA	Inspections + fines; Inspections
Helland (1998a)	Pulp and paper industry	USA	Probability of inspection
Helland (1998b)	Pulp and paper industry	USA	Frequency of inspection
Mixon (1995)	Urban centres	USA	Probability of EPA citation; Degree of citation
Kleit et al. (1998)	Industry in general	USA	Probability of penalty imposition relative to compliance order Probability of penalty; Height of penalty
Pargal et al. (1997)	Industrial plants	India	Number of inspections

Taking a general view on the empirical studies, the results of Oljaca et al. (1998) come the closest to an unbiased, ***welfare based enforcer*** behaviour. Two remarks seem necessary though. Firstly, this study did not explicitly study politically biased enforcer behaviour. The results should therefore not be overstated. The only political variable tested is plant employment, indicating company size. While it could serve as a proxy for a plant's power vis-à-vis enforcers, in this study it was interpreted as indicating the financial viability of plants and with this their ability to pay for penalties without risk of bankruptcy and lay offs of employees. Consequently, fines were expected to rise with plant employment and the results were as expected. Secondly, ***compliance history*** also had an influence on inspections and general enforcement.

Compliance history frequently plays a role also in other studies. Past violations or past enforcement actions towards a plant yielded positive significant impacts on enforcement behaviour also in Helland (1998a and 1998b), Gray and Deily (1996) and Kleit et al. (1998). Unfortunately, as outlined above, one cannot unambiguously establish whether this indicates concern about efficiency or rather a compliance maximising behaviour. In the following, positive significant impacts of compliance history (such as past violations) on enforcement activity are subsumed under welfare considerations, except in cases where the authors explicitly interpret their results in terms of bureaucratic behaviour.

The majority of cases studied gives evidence for a ***mix of politically biased and welfare related enforcer behaviour***. This holds for Dion et al. (1998), Gray and Deily (1996), Helland (1998a), Mixon (1995), Pargal et al. (1997) and Kleit et al. (1998). It also holds for Helland (1998b), although this author explicitly studies empirical evidence for the Harrington model and therefore for bureaucratic behaviour.

Taking a closer look at the types of variables that yield significant estimation impacts on monitoring and enforcement, one finds some patterns, but the heterogeneity of variables actually investigated by the different studies implies that such results can only remain indicative. One or several ***damage variables*** were included in all studies, and at least some of these yielded significant results in all cases. This holds no matter whether their influence on the probability of inspections or general enforcement is concerned (Dion et al., 1998; Gray and Deily, 1996, Helland 1998a; Helland 1998b; Pargal et al., 1997), or whether the factors specifically influencing the probability and exact height of penalties are concerned (Oljaca et al., 1998; Mixon, 1995; Kleit et al., 1998). ***Compliance cost*** data did not yield significant results in the only study which explicitly includes such data

(Gray and Deily, 1996)[34]. Also in Pargal et al. (1997) plant age -a possible proxy for the impact of compliance costs on the viability of plant operations- did not reach significant levels either.

Variables indicating ***benefits of pollution reduction*** were generally not taken into consideration. This is probably explained by the difficulty to obtain related data. Some variables, such as citizen complaints, average household income and the development index of communities might however be proxies for the willingness to pay for reductions in pollution. But these variables are frequently interpreted in terms of political influence on enforcers (e.g. in Helland, 1998a; Pargal et al., 1997). Average household income had a positive effect on enforcement in Helland (1998a and 1998b), while it did not yield significant results in Dion et al. (1998). Also incidences of citizen complaints, studied by Oljaca et al. (1998), did not show significant results, while the development index of communities studied by Pargal et al. (1997) had a significant positive impact on inspections. Other ***political variables*** frequently studied refer to regional unemployment and plant employment. Although mostly significant (cf. Dion et al., 1998; Gray and Deily, 1996; Helland, 1998a; Oljaca et al., 1998), the sign of their impact is not always identical (this is further discussed in sub-section 4.3). Variables indicating that enforcement might result in plant closure were studied by Gray and Deily (1996) and Helland (1998a and 1998b). These were always significant and indicated that enforcement was weakened to avoid such effects, and stronger for firms not in direct danger of closure. Finally, variables indicating political interactions between higher level politicians and enforcement agencies led to heterogeneous results and are discussed in more detail below.

Helland (1998b) more specifically tests in how far Harrington's (1988) targeting model, which comes close to the bureaucratic enforcer view, correctly describes enforcer behaviour. Next to variables relating to environmental damage and political variables as discussed just above, this author also includes a number of variables motivated by the Harrington model in order to test to which extent and according to which criteria regulators target non-compliant plants. His inspection equation captures variables indicating whether and how frequently plants were in violation in preceding quarters, whether or not plants failed to report and how long ago they had last been inspected. Finally, financial and time constraints to the enforcer are reflected in two variables covering a measure for the enforcement budget available to the enforcer and the number of plants. Helland (1998b) finds some evidence for targeting in the sense of enforcers maximising compliance rates: firms who were self-reporting more frequently than the average firm had a lower number of inspections, while plants having failed to report two quarters ago were more likely to be inspected. Furthermore, there are indications that agencies divided plants in high and low frequency groups for inspections, where plants with recent (one-period lagged) violations were more likely to be inspected. But as noted above, the results give indications that the targeting objective was not an unbiased one in the above-suggested way of maximising compliance under a budget constraint. Instead, the author finds evidence also for politically biased and welfare-maximising behaviour of enforcers. Based on this he suggests that Harrington's targeting model correctly describes the plant-regulator interaction only for a subset of plants: those that are targeted for violations, while there exists also a group of plants which is targeted according to their political sensitivity.

Summing up, the empirical evidence suggests that enforcers follow several objectives. Importantly, welfare considerations always seem to have played at least some role in determining enforcement strategies. This is crucial to an assessment of the efficiency of enforcement and the resulting policy outcomes. However, enforcers were in the large

[34] Note that while the authors did therefore not find indications that enforcers consider the compliance costs of plants, enforcers seem to consider the risk of shut down of operations. Therefore, where costs induced a risk of plant closure, this was taken into consideration by the enforcement strategy.

majority of cases also influenced by personal or political objectives. Therefore, in order to come to clearer results about the efficiency properties of enforcement and the potentially resulting implementation gap, it seems sensible to ask whether the enforcer objectives, despite not being unbiased, showed a dominance of one or the other behavioural function. Furthermore, in the following sub-sections, we investigate whether differences in enforcer strategies are visible across industrial sectors, and, more importantly, across different countries or geographical zones.

As to the first question of whether the econometric studies give clear indications about the relative importance of welfare and political considerations in enforcer objectives, its assessment is made difficult by the fact that most empirical studies present only the coefficients of different variables, and not their marginal impacts, which can more easily be compared. Amongst the studies reviewed here, there are only two which actually specify the marginal impacts of different independent variables on the dependent variable. As one of these -Helland (1998a)- furthermore indicates marginal effects only of the political variables whose impact on inspections it studies, a presentation of such results remains anecdotal. Oljaca et al. (1998) publish the average impacts[35] of each independent variable whose effect on enforcement (penalties) they analyse (cf. Table 9.8; note that only significant results are presented in the table). But this study cannot tell much about the relative importance of alternative behavioural enforcement patterns either, as it does not explicitly test politically biased behaviour and only includes one variable that could function as a proxy for such behaviour. Estimating the determinants of penalty functions, the authors investigate the relationship between penalties and characteristics of firms, environmental violations and the firm-regulator relationship respectively (cf. Table 9.6 above).

Table 9. 8: Average impact of independent variables on the penalty

Major violations	\$ 9373
Minor but repeated violation	\$ 5640
previous violator	\$ 5616
no emergency response	\$ 4907
violator employing 20 to 99 workers	\$ 5175
violator employing more than 100 workers	\$ 5721
Plant from paper industry	\$ 10974

Source: Oljaca et al. (1998), 261

Two variables relate to the damage of pollution: those indicating major violations and repeated although minor violations. The variable indicating that a firm did not promptly report a violation -so that no immediate remedial action could be taken- is part of a set of variables indicating firm co-operativeness. Two variables refer to plant employment, where a positive relationship with enforcement is assumed to indicate that enforcers, when deciding about the height of a fine, pay attention to what a plant is able to pay. Employment variables could be interpreted in the sense of political considerations of enforcers, but Oljaca et al. (1998) interpret them as indicating welfare considerations of enforcers (cf. the remarks with respect to Table 9.5).

Apart from the variable indicating that a firm belonged to the pulp and paper industry, an issue discussed further below, the damage variable ‘major violation’ clearly shows the highest impact on the fine (cf. Table 9.8). This gives a high influence to welfare considerations in the overall determination of penalties. Enforcers, furthermore, seem to pay attention to a firm’s ability to pay when issuing fines. As the data show, very small firms received lower penalties. The results are less clear though with respect to larger firms: large firms were not treated much differently than medium-sized firms. The importance of compliance history -increased fines for repeat offenders, as suggested for example by Harrington (1988) and Harford and Harrington (1991) discussed in the

[35] The authors present the variables’ impacts in terms of the difference that each regressor makes in the expected fine, denoted as the average impact on the expected fine.

previous chapter- finally, emphasises the importance of understanding enforcement as a continuing interaction between regulators and enforcers. No significant results were found for whether violations were accidental or intentional, for whether these were detected through citizen complaints; and when a firm did report a violation this had no significant impact on the penalty either (Oljaca et al., 1998).

4.3 Does the affiliation to specific industrial sectors matter for enforcement strategies?

Another relevant question with respect to the effectiveness and efficiency of enforcement is whether enforcers adjust their behaviour -in terms of the relative influence they give to political or compliance cost and damage variables- to the specific situation of industrial sectors. It would be of interest to know whether enforcers put more emphasis on political factors or on environmental factors depending on the specific situation of a plant. One might wonder, for example, whether damage variables play a relatively more decisive role than political considerations in the case of strong polluters, and whether employment effects count relatively more in economically threatened industries.

What makes inter-industry comparisons difficult is, firstly, the limited number of cross-sector studies that include sector variables as arguments in the equations determining enforcer behaviour. Secondly, a discussion about differences pertaining to different industrial sectors is limited by the restricted number of different sectors for which such econometric tests were undertaken. As can be seen in Table 9.7 above, one of the most popular study objects have been water and air regulation directed at the North American (US and Canadian) pulp and paper industry. Amongst the papers presented here, two refer to the US American (Helland, 1998a and 1998b) and one to the Canadian pulp and paper industry (Dion et al., 1998), while one study deals with the US integrated steel industry (Deily and Gray, 1996). Of the remaining studies three investigate industry in general, but only Oljaca et al. (1998) to some extent include sector variables. While Kleit et al. (1998) study penalties imposed on industrial polluters, and Pargal et al. (1997) inspections towards industrial plants in India, they do not differentiate between sectors. Mixon (1995) finally, does not study industry but the enforcement of carbon emission standards towards urban centres.

Comparable to a certain extent are the studies of Dion et al. (1998), Deily and Gray (1996), and partly of Helland (1998a), whereas Helland (1998b) studies a more specific question (cf. sub-section 4.2). As was outlined before, all three studies found indications that welfare related as well as political variables influenced the enforcer’s behaviour. Given that two studies investigate the pulp and paper industry, which is known to be a large polluter, whereas the other one studies the integrated steel industry, and thus a sector in decline (although also polluting), the question is whether similar variables affected enforcer behaviour in a different way across these two sectors.

All three studies find that damage and employment related variables affect enforcement, but employment variables partly show inverse impacts in these studies (cf. Table 9.9). Dion et al. (1998) find that ***plant employment*** has a positive impact on enforcement, in line with their expectations that the more important the plant is for the local labour market, and thus the more visible, the higher the probability of inspection. The opposite holds for Deily and Gray’s (1996) and Helland’s (1998a) estimations. Here the results support the assumption of the regulator’s sensitivity to the firms’ political power or the adverse local impact at stake if the plant closes, which are assumed to increase with plant employment. Directly comparing their results to Deily and Gray’s (1996), Dion et al. (1998) offer a possible explanation for these contradictory findings. They suggest that regulators may undertake more monitoring (studied by them) in visible plants while at the same time being reluctant to impose enforcement measures (studied by Gray and Deily) on plants more likely to challenge the regulator or whose closure might disrupt the local labour market. As a critique to this explanation, it should be noted that Deily and Gray (1996) found the negative impact of plant employment with respect to both of their variables, that covering total enforcement (inspections and fines), but also that covering inspections only. Given moreover that Helland (1998a) and Deily and Gray (1996) study the same industrial sector, one cannot conclude that enforcers pay attention to large employers only in the case of an industry subject to contraction.

Table 9. 9: Variables influencing enforcement behaviour with respect to individual industrial sectors

Author	Industrial sector/polluters	Country	Independent variable	Impact on dependent variable
Dion et al. (1998)	Pulp and paper industry	Canada	Higher pollution flows major violations (check) higher plant employment high regional unemployment	Increase inspections Increase inspections Increases inspections Decreases inspections
Helland (1998a)	Pulp and paper industry	USA	Higher pollution flows Major violations (check) higher plant employment High local unemployment Large employers in high unemployment area probability of plant closure	Increase inspections Increase inspections Decreases inspections no significant result Decreases probability of inspections decreases inspections
Gray and Deily (1996)	Integrated steel industry	USA	Higher emission levels plant location in non-attainment area higher plant employment high county unemployment probability of plant closure	Increase inspections Increase inspections Decreases inspections; and inspections+ fines Increases inspections and enforcement decreases inspections and enforcement

Opposite results are also found with respect to the ***regional unemployment*** variable. In Dion et al.'s (1998) case, the higher the level of unemployment in the region, the lower the probability of inspections. As a counterintuitive result, Gray and Deily (1996) find that firms situated in counties with higher unemployment rates faced more inspections and enforcement. While not offering any explanation for this result in their 1996 study, the authors suggested one in their earlier study, which came to the same result: 'to the extent that high-unemployment areas tend to be more populous or more polluted, the benefits from reducing emissions in such areas may be greater' (Deily and Gray, 1991: 270)[36]. If this were true, enforcers would from a certain threshold onwards be motivated by environmental damage considerations at the expense of negative employment effects, and this even in a declining industry. This finding is however in contrast also to Helland's (1998a), whose data did not yield any significant results for the level of local unemployment on its own, while large employers located in a high unemployment area were subject to a lower probability of inspections. Finally, both Helland (1998a), for the US pulp and paper industry, and Gray and Deily (1996), for the US steel industry, found that variables indicating the ***possibility that a plant might close*** affected enforcer behaviour negatively.

Concluding, the -admittedly few- studies investigated do not give any indication that enforcer behaviour follows different logics across different industrial sectors. More studies, preferably cross-industry studies, are needed to investigate this issue in more detail. The only cross-industry study amongst those covered here that actually differentiates between sectors gives indications that fines differ with sector affiliation (Oljaca et al., 1998). The mean fine for the pulp and paper industry -paid with respect to water quality regulation violations in Georgia- was twice as high as for any other industry category. Which characteristics or reasons actually determine the different penalties is however not discussed by the authors. It might be firm size, damage, or any other plant related reason.

[36] Cited by Dion et al. (1998: 14).

4.4 Are there differences between geographical zones?

The major problem with respect to a study of whether there are empirical indications that enforcer behaviour differs across geographic zones, again, are the limitations set by available econometric studies on the enforcement of environmental regulation. The large majority of the studies available focus on the United States or Canada, and only one on India (cf. Table 9.10), while European studies are inexistent. This table also summarises the earlier findings with respect to patterns of enforcement behaviour. On first sight, there are no visible differences between different countries. Almost in all cases welfare and political concerns were identified as having influenced enforcement strategies.

Table 9. 10: Geographic zones studied and patterns of enforcement behaviour

Author	Industrial sector/polluters	Country	Enforcement behaviour
Oljaca et al. (1998)	Industry in general	USA	Welfare (+ compliance history matters)
Dion et al. (1998)	Pulp and paper industry	Canada	Welfare/political bias
Gray and Deily (1996)	Integrated steel industry	USA	Welfare/political bias (+ compliance history matters)
Helland (1998a)	Pulp and paper industry	USA	Welfare/political bias (+ compliance history matters)
Helland (1998b)	Pulp and paper industry	USA	Welfare/political bias/bureaucrat
Mixon (1995)	Urban centres	USA	Welfare/political bias
Kleit et al. (1998)	Industry in general	USA	Welfare/political bias (+ compliance history matters)
Pargal et al. (1997)	Industrial plants	India	Welfare/political bias

The earlier finding that higher plant employment had a positive impact on enforcement in Canada (cf. Dion et al., 1998), while it had a negative impact in the USA (cf. Gray and Deily, 1996; Helland, 1998a) suggests that US American enforcers might be more sensitive to the political power of regulated actors than their Canadian counterparts. Opposite results between the two countries were also found with respect to the effect of high regional unemployment on enforcement. While this variable decreased the probability of inspections in Dion et al.'s (1998) Canadian case, it increased inspections and enforcement in Gray and Deily's (1996) American case. Further differences across the econometric studies are found in the impact of average household income or the regional development index. Average household income showed a positive impact on enforcement in Helland's (1998a and 1998b) American studies, but did not yield significant results in the Canadian case (cf. Dion et al., 1998). The comparable variable development index in the Indian study (cf. Pargal et al., 1997), however, did influence the number of inspections positively. Again, one might interpret that the Canadian inspectors might pay less attention to political pressures that might be exerted on them, but insignificant results do not allow to draw conclusions.

The only conclusion possible is that the empirical evidence currently available is far too limited in number, and too much restricted to a few countries, to compare enforcement styles across geographic zones. Furthermore, differences in specifications of variables and enforcement functions and in estimation methods across available studies show the need for cross-country or cross-continental studies in order to further assess this question.

As far as possible differences between Canada and the USA are concerned, Harrison (1995) asks whether different enforcement styles have an impact on compliance outcomes. This author investigates the relative effectiveness of the supposedly adversarial and inflexible US enforcement style as compared to the supposedly more co-operative, negotiation-based Canadian approach at the example of compliance of the pulp and paper industry with environmental regulation in both countries. With this she reacts to authors that have criticised the US enforcement style for being inefficient in treating minor and significant violations equally and counterproductive in creating antagonistic relationships between regulators and regulated actors. Unfortunately, however, Harrison's (1995) study diverges from the rest of studies presented here: while she analyses the impact of plant characteristics on compliance in Canada with help of econometric methods, she only makes interference about the effectiveness of different enforcement styles by comparing compliance rates between the US and Canada. There is hence no econometric modelling of enforcement styles and her study does not shed much light on the different patterns of enforcement with respect to plant employment and household income outlined before. At best it could explain that Canadian enforcers may be more sensitive to a possible aggravation of regional unemployment as result of strict enforcement measures, although both countries' inspectors seem to weaken the stringency of enforcement measures where firms are threatened by plant closure (cf. sub-section 4.3). Finding significantly lower compliance rates of the pulp and paper industry with environmental regulation in Canada as compared to the US, Harrison (1995) argues in need for pooled data analysis of industries across countries as well as for more critical examination of the effectiveness of co-operative regulatory regimes.

4.5 Alternative Political Behaviour Theories

The majority of studies presented so far and which explicitly test politically biased enforcer behaviour focused on the very general Stigler-Peltzman framework. Kleit et al. (1998) test two more specific theories when estimating the effect of political factors on enforcement in the form of penalties. These are the legislative dominance theory and the rent extraction theory (developed by Weingast and Moran and by McChesney respectively), both referring to the impact of oversight committees, a typically American issue. Kleit et al. (1998) review the assumptions of these theories as follows (cf. Box 9.2).

Box 9. 2: Summary of major assumptions for Kleit et al.'s (1998) study

The legislative dominance theory

This theory sees the representative system in the US as creating a 'political market' in which politicians try to maximise their own political support. In order to appease their constituency, they strategically place themselves on those committees overseeing agencies that are important to their district of representation. This oversight allows the politician a certain amount of control over regulation. Through various incentives towards agencies -e.g. budgets, sanctioning of uncooperative agencies, reappointment of agency members- politicians further their interests. The theory also assumes that legislators act solely to *represent the interests of their districts*.

For Kleit et al.'s (1998) study, evidence for this theory would be provided by indications that committee members try to maximise their political support by protecting companies from penalties, for example by pressuring enforcers either not to issue fines at all or to issue lower fines.

The rent extraction theory

In contrast to the previously presented theory, the rent extraction theory argues that politicians may threaten to, or actually do, act against other actors to serve *their own interests*, independent of their constituents' concerns. This is based on the idea that quasi-rents (tangible or intangible capital returns to assets) exist in the industrial market, which politicians try to extract through threats and/or actual regulation. Firms are willing to pay private rents (a part of their producer surplus) so that regulators let them earn returns on capital.

Furthermore important is the assumption that a threat of rent extraction may be sufficient for the politician to gain the rents he seeks. The rationale is as follows: the politician issues a threat, the target responds, and the threat may then be removed, i.e. the politician does not necessarily need to take any punitive action against the target. Applied to enforcement, politicians could use bureaucratic agencies to make threats to producers, while bureaucratic agencies, because of their regulatory powers, may in turn use producers to extract rents from legislators in the form of legislative favours. They can do so by threatening a producer who is represented by a legislator. The legislator might then in turn ask the agency for relief on the producer's behalf, which the agency might accept in exchange for legislative favours.

In Kleit et al.'s (1998) study, if rent extraction was occurring, a threatened penalty to the firm should be at a higher level when a firm is represented by a member on a committee overseeing the enforcement agency. There should, however, be no impact on the penalty eventually issued. Instead, the threat represented by the initial penalty should be alleviated by a political payoff, which eliminates any effect on the actually issued penalty.

Kleit et al. (1998) wish to assess, firstly, whether political factors play a role for the probability of penalties being imposed, as well as for the exact height of penalties. If so, they wish to secondly assess whether such political influence rather serves the interest of the district represented by the politician who is placed in an oversight committee, or whether it serves pure personal rent extraction objectives of the politician. The relevant data consist of a non-random sample of compliance orders and penalties. Estimated are, firstly, the probability of the discrete choice between issuing a penalty or only a compliance order and, secondly, the height of the penalty. For the latter, two measures of penalty are used, the initially assigned (threatened) dollar amount (penalty 1), and the penalty eventually applied (penalty 2), which may be lower than the initial penalty as most respondents appeal penalties. The explanatory political variables cover the district's representation by members of different oversight committees.

In their model estimating those variables which influence the issuing of a penalty, the authors find some evidence for the assumption of the *rent extraction theory* and against the legislative dominance theory: companies in districts represented by committee members are more vulnerable to penalties than companies outside these districts. The subsequent model estimating the height of the penalty comes to comparable results as far as the initial penalty is concerned. The results with respect to eventually issued penalties are less clear. While there are indications that producers are threatened with rent extraction, it is not clear who actually makes these threats (committee members or enforcers). However, no evidence was found for the legislative dominance theory (Kleit et al., 1998).

4.6 What do the studies allow to say about the efficiency of monitoring and enforcement and of a possible implementation gap?

The empirical findings suggest that there are three broad reasons for why monitoring and enforcement are not systematically applied. The first has to do with the administrative costs of monitoring and enforcement and the fact that enforcement budgets are generally limited. Limited enforcement budgets as reason behind an implementation gap is the most obvious where enforcer behaviour is driven by bureaucratic motives of compliance maximisation. A second reason for selective enforcement are efficiency considerations, where enforcers may either for reasons of cost allocation and minimisation, or for macro-political reasons, target those plants for which the costs of enforcement are expected not to outweigh the benefits. The third reason are personal-political motives of regulators who may react to pressures from below from the regulated community and their constituencies or from above from higher level politicians. In how far an implementation gap is efficient, obviously depends both on the initial efficiency properties of the environmental policy implemented and on the relative importance of political and efficiency considerations in the determination of enforcement strategies.

The literature gives indications that enforcer behaviour did always take into account efficiency considerations, at least to some extent, although their importance relative to that of differently motivated considerations could not be assessed. This finding suggests that enforcement in the studied cases was at least not completely inefficient and that it is possible that it actually may have improved policy outcomes relative to the initial policy objectives. But as enforcement behaviour was never only determined by welfare consideration, and to some extent always also by political considerations, it can be assumed that enforcement outcomes were not as efficient as they could theoretically have been. Taking both findings together, for the time being, it is not possible to come to clear results about the efficiency properties of enforcement and of the potentially resulting implementation gap in the studies reviewed.

How do the authors themselves estimate the efficiency of the enforcement strategies they identified? Only few of the studies reviewed explicitly address this question. It is obvious that a final assessment of efficiency requires a high level of information: knowledge about the optimality of the initial policy objectives, of enforcement costs, compliance costs and environmental benefits related to improved compliance. The only study taking the further step of assessing costs and benefits of enforcement is Magat and Viscusi (1990), who offer an exploratory cost-benefit analysis, which remains necessarily imprecise as no data were available on the compliance costs entailed by an inspection. These authors approximate the costs of inspections that were covered by their empirical study of enforcement of water pollution regulation towards the pulp and paper industry in the US. They compare these estimates to available assessments of benefits from reductions in BOD discharges and to approximated estimates for compliance costs. As a result the authors suggest that if inspections led firms to make substantial capital investment, the costs of compliance were likely to outweigh the benefits, whereas once this investment made, costs of continuing compliance were more likely to pass the cost-benefit test.

There is also only one study which discusses the efficiency of enforcement outcomes relative to the initial policy objectives (Dion et al., 1998). More specifically, these authors discuss the effects of selective monitoring and enforcement on the efficiency of the implementation of uniform standards. Environmental regulations imposing uniform standards on polluters have been criticised by economics, since such standards ignore that plants generally face non-uniform abatement costs and non-uniform marginal damage functions. The discretion which enforcers may have in allocating their monitoring and enforcement resources, discussed throughout preceding chapters, implies that the monitoring and enforcement activities performed by local regulators ultimately determine, or at least influence, the extent of pollution control undertaken and with this the level of compliance with the regulation. Dion et al. (1998) argue, in line with our earlier suggestions, that as a consequence of discretion in implementation, uniform standards may turn out less detrimental than they seem at first sight, and believe that this is actually what happened in the case they empirically studied: not only was monitoring responsive to the heterogeneity in local conditions faced by plants that were subject to uniform regulation. What is more, enforcers did follow efficiency criteria when allocating their limited enforcement resources. A plant's probability of being inspected rose, for example, with the likely environmental damage caused by the plants' emissions. Note that Dion et al. (1998) therefore give an explicit example for one of our basic hypotheses: next to flexibility by local implementers foreseen in environmental legislation, local enforcers can make use of a second type of flexibility by allowing for a certain degree of non-compliance. Pushing their judgement even further, these authors argue that it may actually be less costly to set sub-optimal uniform standards and let enforcers take care of local specificities, than setting standards that correctly reflect such specificities without leaving room to enforcers to deviate from the standards. Furthermore they consider impossible that regulators correctly specify standards for every possible set of circumstances, which are not even necessarily foreseeable. In their view, thus, implementation of standards needs to be flexible and allow for effective standards to account for changing conditions.

5 Conclusions

Two central results can be drawn from the empirical studies reviewed in this chapter. Firstly, there is evidence that, in general, monitoring and enforcement activity is effective in reducing pollution and deterring non-compliance. The effectiveness of inspections and penalties in general is therefore not at cause when looking for explanations for an implementation gap. Rather the existence of such a gap is explained by a non-systematic application of monitoring and enforcement. Local enforcers hence do indeed enjoy some level of discretion to selectively enforce policy objectives. It is here that the studies reveal the reasons for imperfect enforcement and therefore the explanation for an implementation gap. Secondly, enforcers were found to not only follow welfare considerations when implementing policies, as has been assumed by normative environmental economics, but to rather follow multiple objective functions when determining their enforcement efforts and targeting specific plants. Although most studies found evidence of regulators partly allocating their limited enforcement budgets in line with efficiency criteria, more often than not the enforcement behaviour was determined by political considerations as well. Furthermore, enforcers were in no case shown to follow an unbiased compliance maximising strategy. Consequently, one can assume that enforcers in the studies reviewed here have used their discretion in a partly, although not completely, efficient way. There is also some evidence that local discretion resulting in an implementation gap has helped to improve the efficiency of an inefficient policy by adjusting the sub-optimal policy to the local context during the implementation step.

Two further findings are retained from this chapter for our econometric study of France: from a methodological point of view it is important to note that enforcement and compliance decisions are likely to be interdependent. Consequently, econometric models should take account of this endogeneity in order to arrive at unbiased results. And with respect to the scope of the analysis it is important to cover a large set of explanatory variables, including political ones, to allow for a test of alternative behavioural patterns when studying enforcement.

Last but not least, there are a number of issues that are relevant to the econometric study of the French case but which have not been dealt with in the literature reviewed here. The regulated owners of the municipal waste incinerators are political actors, while, in the literature discussed here, the polluters were generally private firms. Potentially, this feature may have an impact both on the compliance behaviour of the polluter and on the behaviour of the enforcers. Furthermore, the French case deals with initial compliance, whereas other econometric studies of compliance and enforcement have focused on continuing compliance. And finally, most of the studies reviewed here were set in the North-American and not in a European context. There might be differences in enforcement styles between continents.

Chapter 10 An Empirical Evaluation of Initial Compliance and Enforcement in France

1 Introduction

Throughout the previous chapters it has been stressed that enacting a new environmental regulation is only the first step towards actual environmental improvement. In subsequent steps monitoring polluters' behaviour and eventually applying sanctions are needed to foster compliance. Within the overall policy process, the enforcement and compliance stages are the key moments at which the objectives of a policy become reality or not. As was shown in the previous chapter, a major source of possible discrepancy between expectations and outcomes is that resources devoted to monitoring and enforcement activities are limited. In France, about 700 staff are dedicated to controlling 63,000 industrial sites, which are subject to environmental permitting (http://www.environnement.gouv.fr/actua /cominfos/commai99/instaclassee.htm). Given this scarcity, enforcement authorities have to allocate their effort to perform a limited number of monitoring and enforcement activities. As a result enforcement is imperfect and full compliance is unlikely to occur in practice.

The preceding chapters established that, from an economic point of view, imperfect compliance is not *a priori* inefficient if the policy involves inefficiencies. This is particularly the case when the policy involves a 'command-and-control' regulation. Given the range of possible enforcer strategies and polluter responses and their impacts upon efficiency outcomes, empirical studies of monitoring, enforcement, and compliance practices are necessary. This chapter uses econometric models to empirically evaluate the enforcement of and compliance with air emission standards in France, which were imposed by the 1989 European Directive on 'existing' municipal waste incineration plants. It attempts to answer the following set of questions: what is the compliance behaviour of (publicly owned) incinerators? Does the enforcer seek to either maximise welfare, political support or compliance rate? And was the policy outcome more efficient than the theoretical regulation embodied in the Directive? The answer depends on both the actual compliance behaviour of the regulated agent and the enforcement behaviour of the regulator.

Empirical studies of enforcement and compliance were reviewed in the previous chapter. In this context, the present chapter brings new insights in several respects. Firstly, it studies a problem of ***initial compliance***: the Directive basically obliged the incinerators either to close down or to invest in expensive abatement equipment in order to comply. In all previous contributions, the focus is on ***continuing compliance***, i.e. on whether the polluter complies with emission standards once investments are made. In the case of initial compliance, the issue for the enforcer is not primarily monitoring. Non-/compliance is highly visible, as it refers to the application or the lack of application of relevant abatement technology. The issue here lies in the deterring ability of the penalties necessary to foster investments. As a consequence, other contributions focus on monitoring variables, typically the number of inspections, whereas the present study investigates the occurrence of enforcement measures.

The second new aspect of the present study is the nature of the compliance decision-maker: domestic waste incinerators are systematically owned and controlled by municipalities even though they might be privately operated. Such hybrid polluting entities are not exceptional in the transport, water, energy or waste sectors. The normative literature dealing with monitoring and enforcement has paid less attention to this aspect: one has routinely assumed that regulated entities exhibit a foreseeable profit maximising behaviour. This assumption might be legitimate in the case of firms, for instance, but may be less appropriate when the regulation targets public utilities. In the case studied here, the ***polluter is a political agent*** with preferences, which may *a priori* differ from the (private) industrial decision-makers', which have been for the most part investigated in other studies. The focus on municipalities is relevant: they hold a central legal responsibility for waste as well as water management in many countries. As such, they are frequently the targets of environmental policy. While the few other econometric studies that treat the question of differences between the behaviour of public and private sector plants found that public plants are less likely to comply[37] and possess relatively more bargaining power than their private counterparts[38], they never investigated in detail what rationales their behaviour followed. Note furthermore that these studies are all set in Asia. One might expect incinerators owned by municipalities to be politically biased when making their compliance decisions. Lastly, this study analyses ***European data*** in contrast with the previous American and Asian studies. As it seeks *inter alia* to identify political motives, the fact that political and administrative systems in Europe differ from the American model may lead to different results.

The rest of the chapter is organised in the following way. Section 2 includes a short reminder of the regulation of municipal waste incineration in France, while section 3 presents the data set. Section 4 comprises the compliance and enforcement models as well as the estimation results. Section 5 concludes.

2 The regulation and municipal waste incineration in France - A reminder

2.1 The Directive 89/429/EEC and the French municipal waste incineration sector

The European Directive 89/429/EEC on atmospheric emissions aims to limit emissions from 'existing' municipal waste incineration plants. It constitutes a classical command-and-control regulation. The Directive establishes emission limit values for a range of pollutants, including heavy metals, HCL, HC, SO_2, dust, CO and organic compounds. Emission standards as well as compliance deadlines have been differentiated for various sizes of plant, with stricter requirements and shorter deadlines for large incinerators, qualified as those having a capacity of strictly greater than 6 tonnes per hour. The number of incinerators in the early 1990s amounted to approximately 60 large plants (capacity > 6 tonnes per hour), and approximately 254 smaller plants. The analysis in this chapter only deals with the capacity group of large incinerators.

Large plants were required to meet the standards by 1 December 1996. Compliance requires heavy investments in either dry scrubbers at low temperature, semi-dry or wet scrubbers or condensation for the abatement of HCL, HF and SO_2 emissions, and two- or three-stage electric filters and fabric filters for the abatement of dust and heavy metal emissions. The emissions of organic compounds are addressed by adjusting the temperature of combustion and post-combustion (Milhau and Pernin, 1994). There are no big cost differences between these alternative technologies when considering overall costs, i.e. taking operation costs into account (Pernin, 1997; Milhau and Pernin, 1994).[39]

A crucial feature of this policy is that compliance cost involved is very high relative to total incineration cost. Bringing a plant into compliance costs from 230,000 to 600,000 €

[37] Cf. Hartman et al. (1997). This author studies determinants of pollution abatement by pulp and paper mills in Bangladesh, India, Indonesia and Thailand. He found that state owned plants, despite being on average the largest plants, have invested less in pollution abatement than private enterprises

[38] Wang et al. (2001) and Wang and Wheeler (2000) found that firms from the private sector appear to have less bargaining power than state owned enterprises when it comes to the enforcement of pollution charges in China. As a result, charges for pollution levied on public plants tended to be lower than those levied on private plants. Finally, Wang (2000), also studying China's pollution charge system, seems to provide evidence that the bargaining power of state owned plants was superior to that of private plants, but that state owned plants faced greater community pressure and informal regulation. Note that while these studies deal with a pollution tax, it was differentiated between emissions below and above the respective emission standard. With this, the Chinese pollution charges can be vaguely compared to penalties.

[39] However investment costs for a given capacity level considerably vary between incineration plants. The differences may be explained by factors like the space available to install the cleaning technology (Pernin, 1997).

for one tonne per hour of capacity.[40] This represents about twenty percent of total incineration costs (waste collection excluded) (Milhau and Pernin, 1994). This level of cost may perfectly justify a radical compliance strategy: that of closing down the plant. It has been actually observed in many instances. This raises an econometric problem that will be addressed below.

Incineration plays an important role for municipal waste management in France. 33.6 percent of total municipal waste was incinerated in 1993. It is the second most important treatment mode after landfilling (61.4 percent) (ITOM 6, 1995). The organisation of municipal waste incineration is the prime legal responsibility of municipalities in France. However municipalities can delegate the operation of incinerators (like other waste management installations) to private companies. In 1993, about 55 percent of the waste incinerators were privately operated (ITOM 6, 1995). Nevertheless whoever runs the plant, incineration is financed through local taxes and the municipalities are ultimately responsible for compliance.

2.2 The enforcer

In France, practical implementation of the EU Directive relies upon the 'Inspections des Installations Classées' (inspectorate of classified installations), which is a regional emanation of the Ministry of the Environment. Generally, once a plant has been identified as non-compliant, the enforcer initiates an administrative proceeding involving two steps. In the first step, the enforcer issues a 'mise en demeure' in which an ultimate date for compliance is set. If the plant remains non-compliant at this date, it is either closed down or a second step is launched (called 'procédure de consignation'). This second step consists in imposing the plant owner a monetary payment equal to the cost of compliance. When compliance is achieved or the plant closed down, the money is reimbursed minus eventually a monetary penalty.

3 The data set

Data are available on 63 large 'existing' municipal waste incinerators. An inventory initially established by the French Ministry of the Environment (MATE) in April 1998 and updated in May 1999 serves as reference data base for information on compliance behaviour and simple plant characteristics such as treatment capacity and vintage. It is based on information the plants were required to provide to the DRIRE. Missing data were added from two further inventories: ADEME (1998) and ITOM (1995), describing the French incinerator park in 1998 and 1993 respectively. The latter two inventories served as reference database for information on the status of the plants' operators (private/public) and on whether or not the plant incinerated waste with recovery of energy (cf. also annex 6.C). Information on political data was taken from online databases (http://dna.scrutins.net/elections /historique/dna/select.html;). Since data about inspections and informal enforcement activity were not available, we focus on formal enforcement steps taken. Information on such measures, applied towards 14 large incinerators since 1997, was likewise provided by the MATE.

Table 10.1 provides the descriptive statistics of the sample. It is worth noting that the independent variables include several political variables. While in other econometric studies such variables were primarily included with an attempt to explain enforcer behaviour, in our case they may likewise be determinants of the plant's compliance behaviour, given that the French incinerators are owned by municipalities and thus by political actors. The political variables used here are above all two variables describing the importance of an environmentally aware public in the respective municipality.

[40] Costs are increasing in the order dry, semi-dry and wet processes. Average investment costs for the specific steps of off-gas treatment are: 500 kF (76.22 k€) per t/h for the de-dusting facility (electric or fabric filter), 750 kF (114.34 k€) per t/h per reactor (wet stage, semi-dry reactor), between 1 and 2 MF (0.15 – 0.3 M€) for the water treatment step of the gas washing and 500 kF to 1 MF (76.22 k€ to 0.15 M€) for the system of self-monitoring (Pernin, 1997).

GREEN is a proxy for the influence of the green electorate in the district and LEFT is an index variable equal to one if the political coalition of the municipality is left oriented (the ecologist party is part of the left coalitions in France). A further variable that might reflect political considerations of the enforcer is PRIV, describing whether or not an incinerator's operation was delegated to a private company or taken on by the municipality itself. One might imagine that enforcers are more reluctant to enforce regulations towards publicly operated plants. A second hypothesis is that private operators, often belonging to large companies that are also the waste treatment technology providers, care for the image of waste incineration and consequently have a stronger self-interest in compliance than public operators.

Table 10. 1: Descriptive statistics

Variable	Mean	Standard deviation	Definition
Independent variables			
PLANTCAP	20.2897	17.6320	Plant capacity in tonnes of waste incinerated per hour
VINTAGE	77.603	7.2613	Plant vintage
PRIV	0.7460	0.4388	Equal to one if the plant operator is private, zero otherwise
ENERG	0.8413	0.3684	Equal to one if the plant is equipped with an energy recovery system, zero otherwise
LEFT	0.5873	0.4963	Equal to one if the city is led by a political coalition from the left in 1995
GREEN	3.6119	0.9906	Percentage of votes in the local district for the green candidate in the 1995 presidential elections
CLOSED	0.1746	0.3827	Equal to one if the plant has been closed down, zero otherwise
NEW	0.2381	0.4293	Equal to one if plant included a new oven, zero otherwise
Compliance variables			
COMP	0.4444	0.5009	Equal to one if the plant was compliant with the Directive in 1996
LAG1COMP	0.5714	0.4988	Equal to one if the plant was compliant with the Directive with a one year lag, zero otherwise, one year lagged COMP
COMPMULTI	0.8095	0.9480	Equal to zero if the plant was not compliant in 1996, one if it has complied by closing down, two if the plant has complied by making a pollution abatement investment
LAG1COMPMULTI	1.0476	0.9576	One year lagged COMPMULTI
DELAY	1.3333	1.4811	Compliance delay (relative to 1996), year
LAG1DELAY	0.7778	1.1136	One year lagged DELAY (relative to 1997)
Enforcement variable			
ENF	0.5000	0.5189	Equal to one if the incinerator is subjected to a "Procédure de Consignation", zero if subject to a "Mise en Demeure"

It is useful to further discuss compliance and enforcement variables. The first point is that compliance has been clearly imperfect. The rate of compliance with the 1996 deadline was only 44% even though the one-year lagged compliance rate rose up to 57%. A second point, directly following from the first, is that compliance was delayed. Table 10.2 presents the cumulative distribution of the variable DELAY describing compliance delay in years. This variable has a mean of about 1.3 years and a maximum of 5 years indicating that full compliance was expected to be achieved in 2001.

Table 10. 2: Distribution of the variable DELAY

DELAY	Nb of incinerators	Percent
0	28	44.44
1	8	12.70
2	14	22.22
3	6	9.52
4	5	7.94
5	2	3.17
Total	63	100.00
DELAY= compliance delay with the 1996 deadline, in years		

Thirdly, compliance was achieved through either investment or plant closure. This is illustrated in Table 10.3, where the distribution of the variable COMPMULTI is reported. At the end of the compliance period (2001), 11 incinerators out of 63 were closed down.

Table 10. 3: Compliance status in 1996

COMPMULTI	Nb. of incinerators	Percent
Non compliant	35	55.56
Closed down	5	7.94
Upgraded	23	36.51
Total	63	100.00

Finally, enforcement only concerns 14 incinerators, which were those non-compliant in 1998 and later, where 1997 is the year when formal enforcement measures were first taken.

4 Models and results

In this section we proceed in three steps. First, we discuss compliance models without controlling for the endogeneity of enforcement. Different compliance variables are studied and we address the possible endogeneity of plant closure decisions. Then we move on to discussing enforcement models, which again do not control for the endogeneity of compliance. Finally, we discuss endogeneity concerns, although simultaneous equation methods to jointly estimate compliance and enforcement variables could not be applied for reasons of data restrictions.

4.1 The compliance models

The models presented in this sub-section study the behaviour of the regulated agent, i.e. the municipalities owning the waste incinerators. The overall question is whether the plants, although publicly owned, behave in a profit maximising way, or whether different rationales apply to such agents' compliance decisions.

The basic model

Our first objective is to test the impact of a set of plant specific variables on the plant compliance decision represented by the qualitative variable LAG1COMP. The equation estimates the probability that the incinerator is compliant, using a logit model of the form:

$$\Pr\,[\text{LAG1COMP} = 1] = \Psi\,[X\beta]$$

where:

X is the row vector of observed characteristics of incinerators: (PLANTCAP, VINTAGE, LEFT, GREEN, PRIV)

β is the column vector of coefficients

Ψ [.] is the logistic density function.[41]

LAG1COMP is a discrete one-year lagged compliance indicator variable. The variable COMP was not used since models systematically failed to reach statistical significance

[41] Tests were also carried out with the probit model, which led to similar results.

for this variable. This suggests that the incinerators anticipate that the actual compliance deadline[42] is lagged in comparison with the official one. This is not surprising in the French context, where, as explained in chapter 7, enforcement measures generally only start to be applied once the formal deadline for compliance is passed.

PLANTCAP and VINTAGE[43], both continuous variables, are the plant capacity and vintage respectively. They are both expected to have positive signs if the incinerator behaves in a cost minimising way. This is so because of the existence of economies of scale in pollution abatement, which are generally assumed as important for waste incinerators (European Commission, 1997), and because the life span of the abatement investment is shorter in old incinerators. GREEN and LEFT are dummy variables aiming to test whether the municipalities are politically biased in their compliance decisions. Both expected signs are positive since GREEN is a proxy for the influence of the green electorate in the district and LEFT is an index variable equal to one if the political coalition of the municipality is left oriented. Finally PRIV is a control dummy variable stating whether the incinerator is privately operated. We would assume the latter variable to be positively correlated with the dependent variable as well. The reasons are that private operators frequently belong to large companies providing waste incineration technology which have to fear for their reputation, and because enforcers might be able to imply more pressure on private than on public operators.

Table 10. 4: Basic compliance logit model

LAG1COMP	Coefficient	Standard Error	W statistics
PLANTCAP*	0.0580	0.0299	1.94
VINTAGE**	0.1469	0.0512	2.87
PRIV	-0.5267	0.7780	-0.68
LEFT	-0.5465	0.6905	-0.79
GREEN	0.3197	0.3071	1.04

Sample size: 63
Log likelihood statistics = -32.735446
LR chi2 = 20.58
McFadden's R2: 0.239
* Significant at the 10% level.
** Significant at the 1% level.

The results from this estimate are presented in Table 10.4. First note that the two abatement cost variables PLANTCAP and VINTAGE are both significant and their coefficients exhibit the expected positive signs. Neither the political variables LEFT and GREEN nor PRIV are significant. All in all, this basic model suggests that incinerators, although publicly owned, classically minimise compliance costs.

The multiple choice compliance model

However, this model neglects an important aspect: it does not account for the fact that a significant proportion of the plants closed down during the compliance time span. Plant closure could be a sensible compliance strategy in that compliance investment cost represents about twenty percent of the total incineration cost of an *average* incinerator. The consequence is that the compliance decision is a multiple choice involving three options: non-compliance, compliance through investment in abatement technology, and compliance by closing down. Since we study only attributes of the decision-maker, this suggests the use of a multinomial logit model of the form:

[42] That is the deadline in terms of the date from which on enforcement measures are applied.

[43] The year the plant was first put into operation.

$$\Pr[\text{LAG1COMPMULTI} = \text{j}] = P_j = \frac{e^{Y\beta j}}{1 + \sum_{-j=1}^{2} e^{Y\beta_{-j}}} \qquad with\ j = 1{,}2$$

$$\Pr[\text{LAG1COMPMULTI} = 0] = P_0 = \frac{1}{1 + \sum_{-j=1}^{2} e^{Y\beta_{-j}}}$$

where:

Y is a row vector of observed characteristics of incinerators: Y = (PLANTCAP, VINTAGE, PRIV, ENERG, LEFT, GREEN)

β_j is a column vector of coefficients which may change across choices

The only difference between Y and the vector of plant characteristics X used in the basic compliance model is the inclusion of ENERG. ENERG is a dummy variable pertaining to whether the incinerator is equipped with an energy recovery system. The variable ENERG controls for the influence of waste management policy that was implemented in parallel with the MWI Directive. The basic aim of this policy, initiated under the 1992 Law on Waste Management (Loi n° 92-646 du 13 juillet 1992), was to promote waste recycling and waste incineration with energy recovery at the expense of traditional incineration and landfilling. For the decision on how to comply with the Directive's air emission standards, this introduces an asymmetry between the alternatives abatement investment and plant closure. Incinerators with energy recovery are more valuable in the context of the new waste management policy and these plants are thus less likely to be closed down.[44]

Note that a further major difference with respect to the basic model is the expected influences of the cost variables PLANTCAP and VINTAGE. Expected impacts are presented in Table 10.5. In comparison with the basic model, the first difference is that PLANTCAP has no impact on compliance cost when compliance is reached through plant closure: there are no economies of scale in plant closure. The second difference lies in the differential impact of VINTAGE on closure and investment. Closure is cost efficient when the plant is old. Or put differently, when the business-as-usual closing date -the date at which the plant would have closed down in the absence of regulation- is close to the compliance deadline. The typical life span of an incinerator is about 30 years. In the extreme case the compliance deadline and the business-as-usual closing date coincide. In this case, compliance cost is nil. VINTAGE is thus expected to have a negative impact on the compliance through closure decision. As far as investment is concerned, the assumptions discussed in the basic model hold.

Table 10. 5: Expected impacts of the cost variables in the multiple choice compliance model

	Compliance through closure	Compliance through investment
PLANTCAP	0	+
VINTAGE	-	+

Reference choice = non-compliance

0, +, - mean no impact, a positive and a negative impact, respectively

[44] ENERG was therefore not included in the previous estimation, because it can be assumed not to influence the general probability of compliance, but rather the way how compliance is reached.

Table 10.6 presents the results of this estimation. Interpretation of the parameters of multinominal logit models requires caution. Note that:

$$ln\frac{P_j}{P_0} = Y\beta_j$$ and

$$\ln\frac{P_j}{P_{-j}} = Y(\beta_j - \beta_{-j}),$$

where j = 1,2 denotes the compliance decisions compliance through closure and compliance through investment respectively.

The coefficients β_j measure the impact of the attributes Y on the log-odds that the incinerator chooses j instead of 0 while the difference between β_j and β_{-j} measures the impact on the log-odds that he chooses j instead of -j.[45]

Note further that robust estimation was applied. Robust estimators are less sensitive to violations of model assumptions than so called 'best' estimators, specifically with respect to assumptions regarding a normal distribution of the error term. While the standard case assumes for example that the independent variables and the error term are independent and identically distributed, robust estimation is not dependent on these assumptions. The other side of the coin is that they also tend to be less precise.

Table 10. 6: Multinomial logit compliance model

Independent variables	**Coefficients**	**Robust standard Error**	**W statistics**
LAG1COMPMULTI = compliance through plant closure			
PLANTCAP	-0.0597	0.0815	-0.73
VINTAGE*	0.1241	0.0760	1.63
PRIV	-1.3543	1.2085	-1.12
ENERG	-0.5693	1.3149	-0.43
LEFT	0.3159	1.2481	0.25
GREEN	0.4818	0.4777	1.01
LAG1COMPMULTI = compliance through abatement investment			
PLANTCAP**	0.0755	0.0341	2.21
VINTAGE**	0.1314	0.0585	2.24
PRIV	-0.3890	1.0369	-0.38
ENERG***	20.6497	5.0631	4.08
LEFT	-0.5340	0.8463	-0.63
GREEN	0.0493	0.4590	0.11

Reference choice = non compliance
Sample size: 63
Log likelihood statistics = -42.068078
McFadden's R2: 0.290
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

Coming back to Table 10.6, the results are similar to those of the basic model in that neither the political variables LEFT and GREEN nor the operator status PRIV are significant. As to the cost variables' coefficients, they are in line with the expectations in

[45] One necessary condition for using a multinomial logit model is that the choice set is truly irrelevant (that is odds ratios are independent of other alternatives). We have used the IIA test suggested by Hausman and McFadden (1984) to check whether this condition holds. More specifically we have compared the coefficients of the full model with those of a restricted model in which we have dropped the alternative "compliance through investment" and the corresponding observations. We obtained a Chi2 statistics with 6 degrees of freedom equal to -20.46, which strongly supports the validity of the IIA assumption.

the equation estimating the log-odds of compliance through investment instead of non-compliance: investing incinerators seem to be minimising compliance cost. Also, ENERG performed as expected.

However the equation estimating the log-odds of compliance through closure instead of non-compliance diverges from the cost-minimising assumption. If PLANTCAP is expectedly non significant, VINTAGE has a positive impact on the log-odds that the incinerator is closed down instead of being non compliant: old incinerators are less likely to be closed down instead of remaining non-compliant.

How strong this effect is, is shown in Table 10.7, which reports predicted probabilities[46] of being compliant through plant closure or investment instead of non-compliant for the oldest (VINTAGE=1965) and the newest (VINTAGE=1990) incinerators. All other characteristics being equal, as a counter-intuitive result, the newest incinerator (VINTAGE=1990) has about thirteen times more chance of complying by plant closure than the oldest one (VINTAGE=1990), whereas the ratio is only 2.25 for the probability of complying through investment. This latter result is more in line with intuition.

Table 10. 7: Predicted impacts of VINTAGE on the probabilities of being compliant using the multinomial logit compliance model

	Predicted probabilities	Odd ratio
Pr(compliance through closure) :		
$Pold_1$	0.0183	$Pnew_1 / Pold_1 = 13.07$
$Pnew_1$	0.2392	
Pr(compliance through investment)		
$Pold_2$	0,2741	$Pnew_2 / Pold_2 = 2.25$
$Pnew_2$	0,6100	

P1old = Pr(compliance through closure) of the hypothetical oldest incinerator with VINTAGE=1965 and the average values for the other independent variables
P1new = Pr(compliance through closure) of the hypothetical newest incinerator with VINTAGE=1990 and the average values for the other independent variables
P2old = Pr(compliance through investment) of the hypothetical oldest incinerator
P2new = Pr(compliance through investment) of the hypothetical newest incinerator

Overall, this result is counter-intuitive. Taking into account the influence of enforcement can, however, provide a possible explanation. Old incinerators are those for which the business-as-usual closing date is closest to the regulatory deadline. Without any enforcement threat, these are the plants which would have the shortest compliance delay. The enforcer may anticipate this and target more recent plants betting on the fact that non-compliance of the oldest plants would only be a short temporary state before they are closed down anyway. To explore whether this explanation holds in practice, we first develop compliance models involving closure considerations and compliance delay and move then on to enforcement models.

A two-stage estimation of compliance

In the intuition that has just been developed, there is the idea that the incinerator has an incentive to delay compliance until the business-as-usual compliance date. This, however, is true only for the incinerator for which the first best compliance option is closure. In the data set, incinerators that were eventually closed down are characterised by the dummy variable CLOSED. One may want to include this variable in the compliance logit equation in order to control for this influence. But since plant closure is one of the possible compliance options, this involves the econometric problem of endogeneity. The solution to this endogeneity problem is to use an instrumental variable two-stage estimation. In the first stage, CLOSED is estimated using an instrument variable non-correlated with the compliance variable. An appropriate instrument variable is ENERG,

[46] Predicted over the above equations.

which is supposed to negatively influence CLOSED. Then, in a second stage, compliance is estimated using among the independent variables the probability of closure predicted in the first stage. Our estimation model is thus:

First stage equation: Pr [CLOSED = 1]= $\Psi\ [\alpha ENERG\ ]$

Second stage equation: Pr [LAG1COMP = 1]= $\Psi\ [X\beta + \delta PCLOSED\ ]$

where:

X is the row vector of characteristics (PLANTCAP, VINTAGE, LEFT, GREEN, PRIV)

PCLOSED is the estimated value of CLOSED obtained from the first stage equation

α, δ and β are coefficients and a column vector of coefficients, respectively

Ψ [.] is the logistic density function.

Table 10. 8: Two-stage logit compliance model

Independent variables	Coefficients	Robust standard Error	W statistics
FIRST STAGE = estimation of CLOSED			
ENERG***	-2.6672	0.7984	-3.34
SECOND STAGE = estimation of LAG1COMP			
PLANTCAP**	0.0567	0.0255	2.22
VINTAGE**	0.1224	0.0548	2.23
PRIV	-0.7025	1.0316	-0.68
LEFT	-0.3632	0.8212	-0.44
GREEN	0.0652	0.3953	0.16
PCLOSED*	-3.8002	2.3159	-1.64

Sample size: 63
Log likelihood statistics of stage 1 equation = -23.29075
Log likelihood statistics of stage 2 equation = -31.233082
McFadden's R2: 0.274
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

The estimation results of Table 10.8 show that the coefficient PCLOSED is significant and negative. This supports the assumption that, all other things being equal (and notably plant vintage in this model), expectation of a near plant closure leads the decision-maker to delay compliance longer than the one year already included in the dependent variable.

A Poisson estimation of compliance delay

While the previous estimation produced only a rough result of whether or not the expectation of a near plant closure may be an incentive to delay compliance, this issue can be further investigated and in more detail. Studying a new endogenous variable, LAG1DELAY, which measures the compliance delay (with respect to a one year lagged compliance deadline, i.e. assuming the compliance deadline were 1997), might allow us to come to more subtle results and to maybe detect political influences on the compliance behaviour.

Given that LAG1DELAY is asymmetrically distributed, a Poisson regression model seems appropriate. One can estimate an equation distributed around a Poisson distribution of the form:

$$\text{LAG1DELAY} = e^{Y\beta}$$

where:

Y is the row vector of observed characteristics (PLANTCAP, VINTAGE, PRIV, ENERG, LEFT, GREEN)

β is a column vector of coefficients

Note that PCLOSED is not included in the equation. The justification is that if the probability of closure may be assumed to have an influence on the probability of delaying compliance, there is no *a priori* link with the duration of this delay. The results of the Poisson model are presented in Table 10.9. As before, PLANTCAP and VINTAGE are both significant exhibiting negative signs in line with a cost minimising behavioural pattern. The novelty lies in the fact that GREEN is significant at the 10-percent level. The negative sign of the coefficient suggests that decision-makers in greener electoral districts are more prone to comply fast, which is not absurd.

Table 10. 9: Compliance delay Poisson model

LAG1DELAY	Coefficient	Standard error	W statistics
PLANTCAP*	-0.0288	0.0156	-1.85
VINTAGE**	-0.0643	0.0270	-2.38
PRIV	-0.1021	0.3808	-0.27
ENERG	0.2881	0.4247	0.68
LEFT	0.4128	0.3636	1.14
GREEN*	-0.2781	0.1658	-1.68

Sample size: 63
Log likelihood statistics = -71.316571
LR chi2 = 18.94
Mc Fadden's R2: 0.117
* Significant at the 10% level.
*** Significant at the 1% level.

Nevertheless, this is the first time that a political variable is significant. A possible explanation lies in the fact that we are dealing with initial compliance. Failing to comply in this context is a very strong and visible signal. It implies not to do anything, i.e. to neither invest in abatement measures nor to close the plant. By contrast, continuing compliance (the issue once initial compliance -through abatement investment- is achieved) consists in continuously meeting the emission limit value defined by the regulation, where deviation of actual emission levels from the standard is possible. Actual emission levels may for example fluctuate with the composition of waste inputs. Although deviation from the emission limit value can be significant after initial compliance is reached, the abatement level will generally not be zero, since investment in abatement technology was undertaken. Given this distinction, an enforcer allocating his scarce resources between initial and continuing compliance may be inclined to give priority to initial compliance and thus to obtaining -although possibly only after delays- full initial compliance.

In the data set used here, full compliance was expected to be obtained in 2001. In this context, the incinerators had no choice with respect to the initial compliance decision: enforcement concerned all plants that did not comply in 1998. The only degree of freedom plant owners had was the possibility to delay compliance. These considerations could explain the results of Table 10.9. The compliance delay was the only decision variable into which the incinerator (and the municipality, which in fact makes the decision) could introduce its preferences. At this stage, the argument is fragile since its core lies in the enforcer's preferences. We will now move on to the study of enforcement decisions.

4.2 A naive enforcement model

Formal enforcement measures were applied towards 14 large French incinerators that had not been compliant in 1998. 50% of these concerned the weaker measure ‘mise en demeure’, consisting in fixing an ultimate date by which the plant has to be brought into compliance, while in 50% of the cases the next and stricter enforcement step needed to be launched. This is the ‘procédure de consignation’ which imposes a monetary payment into the public treasury, equalling the necessary abatement investment costs, on the incinerators. The limited number of observations -14 enforcement cases- poses problems for the subsequent estimations.

We estimate a simple logit model to assess the influence of the plant characteristics VINTAGE and PLANTCAP on enforcement measures, in order to determine which rationale was underlying enforcement behaviour.

$$\Pr[ENF = 1] = \Psi[X\beta]$$

where:

X is the row vector of observed characteristics of incinerators: (VINTAGE, PLANTCAP)

β is the column vector of coefficients

Ψ [.] is the logistic density function.

ENF is the discrete enforcement variable, taking on the value one if the second, stricter enforcement step was applied, and the value zero, if only the first enforcement step was applied.

ENF is estimated only over VINTAGE and PLANTCAP because the number of observations (14) does not allow to include more independent variables. If the higher closure probability of relatively newer incinerators, compared to that of the older plants, we found earlier can be explained by enforcers targeting newer plants we would expect VINTAGE to have a positive coefficient. The suggested rationale behind this assumption was that incinerators close to the business-as-usual closing date can be expected to have the shortest compliance delay, even without an enforcement threat. A regulator efficiently allocating his limited enforcement resources might anticipate this and target more recent plants. More generally, the expected sign for both VINTAGE and PLANTCAP depends on the enforcer’s behaviour. If enforcers were behaving as welfare maximisers we would expect them to target plants for which compliance is attainable at relatively lower cost. Then both VINTAGE and PLANTCAP would have positive signs. But if enforcers were behaving as bureaucrats, maximising compliance, we would expect them to target high cost plants, and negative signs were to be expected for both independent variables.

Results of the estimation are presented in Table 10.10 below.

Table 10. 10: Naive enforcement logit model

ENF	Coefficient.	Robust Standard error	W Statistics
VINTAGE **	-0.3854	0.1723	-2.24
PLANTCAP	-0.0616	0.1279	-0.48

Sample size: 14
Log likelihood statistics = -6.7842
**Significant at the 5% level.

In contradiction to our earlier hypothesis, the result shows a negative sign for the variable VINTAGE. What we can conclude at this stage is that there is no evidence for the above stated hypothesis of enforcers targeting more recent plants. On the contrary, the regression results suggest that they target older plants, and enforcers therefore maximise compliance. PLANTCAP was not significant.

This result is illustrated by descriptive statistics presented in Table 10.11. The average capacity of the incinerators subject to the first as compared to the second enforcement step is rather similar. Opposite to this, the incinerators subject to stricter enforcement measures were on average older.

Table 10. 11: Descriptive statistics on enforcement

	1st enforcement step					2nd enforcement step				
	Observ.	Mean	Standard deviation	Min	Max	Observ.	Mean	Standard deviation	Min	Max
PLANTCAP	7	14	6.35	8	24	7	13.49	7.14	7	28
VINTAGE	7	77.6	6.88	70	88	7	71.4	1.81	68	73

4.3 Endogeneity issues

As was outlined in the preceding chapter, one might sensibly assume enforcement agencies not to apply enforcement measures towards all plants but rather to target specific plants, for example by focusing on those plants that can be expected to be non-compliant. Similarly, incinerators will integrate expected enforcement behaviour in their compliance decision. The anticipation, by each actor who makes a decision of the behaviour of the other actor, makes decisions interdependent, leading to an endogeneity problem. In order to take account of endogeneity we use two-stage estimation to assess, firstly, in how far incinerators took the probability of enforcement into consideration when making their compliance decision. We replicate the compliance models estimated in section 4.1 now including predicted enforcement. Secondly, we also use two-stage estimation to assess in how far enforcers took the probability of compliance into account when making their enforcement decision.

To this end, enforcement is predicted on the basis of a naive enforcement model, which only includes VINTAGE as a predictor.

$$\Pr[ENF = 1] = \Psi[X\beta]$$

where:

X is the row vector of observed characteristics of incinerators: (VINTAGE)

β is the column vector of coefficients

Ψ [.] is the logistic density function.

This yields PENF, which is then introduced in the compliance models in the second stage.

A compliance model taking into account predicted enforcement

Next to the impact of the set of plant specific variables on the plant decision represented by the qualitative variable LAG1COMP used earlier, we include the variable PENF. The equation again estimates the probability that an incinerator is compliant, using a logit model of the form:

$$\Pr[LAG1COMP = 1] = \Psi[X\beta]$$

where:

X is the row vector of observed characteristics of incinerators: (PLANTCAP, VINTAGE, LEFT, GREEN, PRIV, PENF)

β is the column vector of coefficients

Ψ [.] is the logistic density function.

We expect PLANTCAP, VINTAGE, LEFT, GREEN and PRIV to have the same signs as predicted under section 4.1. PENF should show a positive sign, indicating that the probability of enforcement induces the municipalities to bring their incinerators into compliance in order to escape the costs enforcement actions may cause them. Results of this regression are presented in Table 10.12.

Table 10. 12: Logit compliance model taking into account predicted enforcement

LAG1COMP	Coefficient	Robust Standard Error	W Statistics
PLANTCAP**	0.0558	0.0236	2.37
VINTAGE	0.0547	0.1528	0.36
PRIV	-0.4691	0.9642	-0.49
LEFT	-0.5971	0.8460	-0.71
GREEN	0.3849	0.3982	0.97
PENF	-2.0937	3.2747	-0.64

Sample size: 63
Log likelihood statistics = -32.5124
LR chi2 = 21.02
Mc Fadden's R2: 0.224
**Significant at the 5% level.

As in the earlier version of this model, PLANTCAP is significant and shows the expected positive sign. Note that this result here -unlike in the estimation under section 4.1- is unbiased. The estimation does not give any result for PENF. Econometric modelling therefore does not provide any evidence that incinerators take into account expected enforcement when making their compliance decision. However, the insignificant result may be due to the low number of enforcement cases in the sample.

The multiple choice compliance model taking into account predicted enforcement

Endogeneity issues were also introduced into the estimation taking account of the fact that plants, when making compliance decisions, were faced with a multiple choice between the options non-compliance, compliance through investment in abatement technology, and compliance by closing down. In order to model this multiple choice compliance decision, we again use a multinomial logit model of the form:

$$\Pr[\text{LAG1COMPMULTI} = \text{j}] = P_j = \frac{e^{Y\beta j}}{1 + \sum_{-j=1}^{2} e^{Y\beta_{-j}}} \qquad with\ j = 1,2$$

$$\Pr[\text{LAG1COMPMULTI} = 0] = P_0 = \frac{1}{1 + \sum_{-j=1}^{2} e^{Y\beta_{-j}}}$$

where:

Y is a row vector of observed characteristics of incinerators: Y = (PLANTCAP, VINTAGE, PRIV, ENERG, LEFT, GREEN, PENF)

β_j is a column vector of coefficients which may change across choices

Expected impacts of the variables used before are presented in Table 10.5 above. Additionally we include PENF. Results are presented in Table 10.13. As a first result, PENF was not significant for either compliance choice, compliance through plant closure instead of remaining non-compliant or through investment instead of remaining non-compliant. Again, we cannot positively establish that the incinerators took the probability of enforcement into account when making their compliance decision, but we cannot conclude either that they did not pay attention to it. Furthermore, as far as compliance through closure is concerned, VINTAGE is no longer significant, and neither is any other variable. As far as the choice compliance through investment is concerned, PLANTCAP remains significant and this result is now unbiased and so does ENERG, although the significance level of the latter is low. VINTAGE did no longer reach significant levels. Summing this up, we partly get contradictory results between the biased and the unbiased estimation.

Table 10. 13: Multinomial logit compliance model taking predicted enforcement into account

Independent variables	Coefficients	Robust standard Error	W statistics
LAG1COMPMULTI = compliance through plant closure			
PLANTCAP	-0.0624	0.0790	-0.79
VINTAGE	0.0850	0.2804	0.3
PRIV	-1.3192	1.3095	-1.01
ENERG	-0.6128	1.1976	-0.51
LEFT	0.2864	1.2620	0.23
GREEN	0.5209	0.4862	1.07
PENF	-0.8060	5.9389	-0.14
LAG1COMPMULTI = compliance through abatement investment			
PLANTCAP**	0.0745	0.0340	2.19
VINTAGE	0.0495	0.1625	0.3
PRIV	-0.3505	1.0428	-0.34
ENERG*	20.5479	13.0772	1.57
LEFT	-0.5905	0.8831	-0.67
GREEN	0.1355	0.5246	0.26
PENF	-1.9731	3.7701	-0.52

Reference choice = non compliance
Sample size: 63
Log likelihood statistics = -41.9119
Mc Fadden's R2: 0.293
* Significant at the 12% level.
** Significant at the 5% level.

Poisson estimation of compliance delay taking predicted enforcement into consideration

Finally, we also tried to estimate whether the probability of enforcement had an impact on the compliance delay LAG1DELAY. We would have expected the probability of enforcement to reduce the compliance delay. Poisson regression models did not yield significant results for PENF either and the results are therefore not presented here.

Considering alternatively descriptive statistics (Table 10.14), we find a reverse relationship between the probability of enforcement and the compliance delay. The probability of enforcement is clearly higher for the 27 plants which complied with one or more years delay with respect to the actual compliance deadline of 1997, than for the 36 plants which complied earlier. This could be interpreted as an indication of enforcers targeting those incinerators which are less likely to comply.

Table 10. 14 Descriptive statistics on the compliance delay in relation to predicted enforcement

	predicted enforcement				
Compliance delay in years with respect to 1997	Observations	Mean	Standard deviation	min	Max
0	36	0.25	0.31	0.003	0.95
1	14	0.52	0.29	0.004	0.91
≥ 2	13	0.48	0.30	0.005	0.88

An enforcement model taking predicted compliance into consideration

The enforcement model presented under section 4.2 did not allow explaining the earlier result of newer plants having a larger probability of closing down relative to older plants. Nevertheless, the negative coefficient for VINTAGE in this enforcement model might be explained by the overall lower possibility of newer plants to delay compliance. If enforcers took this into account, they might be driven to focus their effort on older plants in order to minimise the overall compliance delay. This would be in line with the bureaucratic enforcer type who maximises compliance.

We test this by estimating enforcement over predicted compliance also for a one-year lag (PLAG1COMP). We would expect PLAG1COMP to have a negative coefficient if enforcers were targeting those plants whose probability to comply is low. To this end we predict PLAG1COMP on the basis of a simple logit compliance model as given in:

Pr [LAG1COMP = 1]= Ψ [$X\beta$]

where:

X is the row vector of observed characteristics of incinerators: (PLANTCAP, VINTAGE, ENERG)

β is the column vector of coefficients

Ψ [.] is the logistic density function.

We then use this in a re-estimation of ENF in the logit model

Pr [ENF = 1]= Ψ [$X\beta$],

where:

X is the row vector of observed characteristics of incinerators: (VINTAGE, PLANTCAP, PLAG1COMP)

β is the column vector of coefficients

Ψ [.] is the logistic density function

which this time includes PLAG1COMP as additional explanatory variable. Results are presented in Table 10.15.

Table 10. 15: Enforcement logit model taking predicted compliance into account

ENF	Coefficient.	Robust Standard error	W Statistics
VINTAGE*	-0.3647	0.2010	-1.81
PLANTCAP	-0.0417	0.1472	-0.28
PLAG1COMP	-1.0700	3.5565	-0.30

Sample size: 14
Log likelihood statistics = -6.7612
LR chi2 = 5.89
Mc Fadden's R2: 0.303
* Significant at the 10% level.

PLAG1COMP does not yield significant results. We can therefore not say whether or not enforcers target those plants whose probability to comply is low. PLANTCAP was not significant either, while VINTAGE remained significant, again with a negative sign.

Contrary to this econometric model, descriptive statistics relating enforcement to the plants' probability to comply (Table 10.16) suggest, however, that enforcement was targeted at incinerators that were less likely to comply. The probability to comply in 1997 was higher for those plants that were only subject to the first enforcement step than for those that were subject to the second, stricter enforcement step.

Table 10. 16 Descriptive statistics on enforcement in relation to compliance

	1st enforcement step					2nd enforcement step				
	Obser-vations	mean	Standard deviation	min	max	Obser-vations	mean	Standard deviation	min	max
PLAG1COMP	7	0.52	0.24	0.12	0.79	7	0.32	0.21	0.049	0.63

5 Conclusions

We can assume that decision makers, incinerators as well as enforcers -if they are rationally behaving actors- take each other's expected behaviour into consideration when making their own decision. In econometric studies, this requires 2-stage estimation of simultaneous compliance and enforcement equations to fully enlighten the causality relationships between variables. Such models were applied but, because of the low number of observations referring to enforcement, we encountered problems in reaching significant results over two stages. This is why we also presented the results of non-simultaneous models. This has two implications. First, we cannot say anything about the relationship between more or less strict enforcement and compliance decisions. Second, caution is necessary when interpreting the results of these more 'naive' models, because they will tend to be biased. Our results are, nevertheless, a first test of the behaviour of the regulated municipal waste incinerators and the enforcers.

One interesting result uncovered by our analysis is that the incinerators followed behavioural patterns that are generally in line with cost-minimising behaviour. This is a striking result given that the plants are public monopolies, owned by a political actor which might be expected not to have profit maximisation as a first objective. More precisely, incinerators showed cost-minimising behaviour in their choices of compliance and investment in abatement technology. Contrary to this, political variables were, generally, not significant with respect to compliance patterns.

There is however one exception with respect to the non-significance of political variables: plants in 'greener' municipalities complied earlier, which may indicate some political bias of the regulated agent. This finding is important as we study initial compliance, which is easily visible to the enforcer and the public. The result suggests that municipalities indeed paid attention to the preferences of their constituencies. Interview results had suggested that political support interests, in several municipalities, were obstacles to early compliance where the public was not willing to bear increased local taxes and where no low-cost waste treatment alternative was available. The finding here suggests that the inverse case may also have occurred, i.e. that in greener municipalities political support and image reasons have induced the mayors to invest early in abatement equipment. This suggests implicitly that the public in such constituencies value environmental improvements more than tax increases.

We also tested for possible behavioural differences between publicly and privately operated plants with respect to compliance behaviour. Estimations, however, yielded no significant results. While these results do not allow us to conclude that the operator status played no role, it was not possible to positively establish a difference between the two types of incinerators. It should be noted that as our sample covered only publicly owned plants (all French municipal waste incinerators were publicly owned), it was not possible to estimate possible differences in the behaviour of plants according to ownership status as done in other econometric studies.

When studying the different compliance alternatives, we found that the expectation of a near plant closure made incinerators delay compliance longer than did plants which eventually invested in abatement technology. This may also be in line with efficiency considerations, as these incinerators obviously tried to keep as much as possible to their business-as-usual closure date. One result that is not directly reconcilable with efficiency is the fact that newer plants -in comparison to older plants- were shown to be more likely to close down. We had suggested that this might be explained by enforcers targeting newer plants, knowing that the business-as-usual closure date and, thus, the compliance delay were the shortest for the oldest incinerators. This was however contradicted by the naive enforcement model, indicating that enforcers target rather older plants.

This latter result indicates that enforcers tend to maximise compliance and hence behave as bureaucrats. The fact that econometric estimation indicates that older plants seem to face stricter enforcement measures suggests that enforcers targeted those plants which were less likely to comply. This result was also supported by descriptive statistics. Targeting, as was shown in chapter 8, is crucial for cost-effectiveness where enforcement budgets are scarce. Descriptive statistics finally indicate that where formal enforcement measures were applied, they were eventually effective in reaching initial compliance, although with varying delays.