ISSN 0316-5078

INSTITUTE FOR ECONOMIC RESEARCH

Discussion Paper #873

Incomplete Enforcement with Endogenous Regulatory Choice

by

Devon Garvie
Queen's University

and

Andrew Keeler
University of Georgia

January 1993

Queen's University
Kingston, Ontario, Canada K7L 3N6

INCOMPLETE ENFORCEMENT WITH ENDOGENOUS REGULATORY CHOICE

Devon Garvie*
Queen's University

Andrew Keeler
The University of Georgia

January, 1993

Abstract
This paper extends the economic literature on the incomplete enforcement of social regulation by incorporating regulatory choice in an institutional environment of limited regulatory resources and powers. We show how regulatory decisions determine the structure of incentives faced by regulated firms. Our results indicate that the expense of monitoring relative to the regulator's power to levy penalties helps to explain the differences between "compliance" and "deterrence" enforcement styles. We find that in most circumstances firms with higher abatement costs will receive a larger share of regulatory resources and thus face higher penalties than firms with lower costs.

* The author would like to acknowledge financial support for this research from the Principal's Advisory Research Committee at Queen's University, Kingston, Canada.

We would like to thank Anthony Fisher, Michael Hanemann, Mitch Renkow, Suzanne Scotchmer and Roger Ware for useful comments.

1. Introduction

This paper studies the enforcement of pollution control regulations. We extend existing work by introducing endogenous regulatory choice of monitoring strategies and enforcement expenditures in an institutional environment of limited regulatory powers and resources. In much of the existing literature on incomplete enforcement, firm penalties have been specified as a function of the firm's choice about the extent of pollution abatement but have been independent of regulatory action [for example, see Viscusi and Zeckhauser (1979), Harford (1978), and Jones (1989)]. This approach has been useful in examining the optimal selection and severity of policy instruments, but it ignores the process of regulatory choice and action and thus fails to take account of the important interactions between regulators and firms.[1] Contributions have also been made in understanding regulatory choice of monitoring in an environment where regulatory budgets are limited and monitoring firms is costly [Russell, Harrington, and Vaughan (1986), Harrington (1988), Jones and Scotchmer (1990), Harford and Harrington (1991), Harford (1991), Veljanovski (1984)]. There is an externally-imposed limit on penalties in this line of research: in almost all circumstances the regulator either implicitly or explicitly chooses the maximum fine level to achieve the greatest deterrence possible.[2] Firms found in violation are penalized by fines imposed at no

---

1. We define the regulator as the agency with specific responsibility for monitoring firm compliance and taking enforcement actions to levy penalties against noncompliant firms.

2. The exception is Harrington (1988), where a first offense carries no fine in order to maximize an additional aspect of that first offense - being reclassified as a likely offender and facing the maximum fine and an increased monitoring probability in subsequent periods. Harford and Harrington (1991) and Harford (1991) found that maximum fines were likely to be imposed at all times when minimizing control costs is a component of regulatory objectives. There are several reasons why maximum fines may not be optimal in all circumstances. Stigler (1970) stressed the role of progressively larger fines for more severe offenses to preserve marginal deterrence, and Polinsky and Shavell (1979) found that non-maximum fines could be optimal when offenders are risk-averse.

explicit cost to regulatory agencies. Further, penalties depend on a dichotomous compliance/noncompliance determination and are independent of the extent of the violation.

None of the existing models of enforcement adequately explain why different pollution media, different industries, and different firms producing the same pollutant are treated so differently by enforcement authorities. We believe that a fundamental explanation of enforcement lies in the nature and determinants of regulatory decisions on how aggressively to pursue sanctions against firms found out of compliance. Imposing penalties on firms requires regulatory choices about negotiation, documentation, publicity, and litigation. The type and extent of regulatory action determine the penalties faced by firms in the form of fines, production holdups, and the costs of engaging in quasi-legal and legal proceedings.

Melnick (1983) documents the complex set of choices facing regulators when penalizing stationary air pollution sources. The EPA can negotiate a compliance schedule, begin administrative proceedings, file suit under Section 113 of the Clean Air Act, or restrict itself to putting pressure on state enforcement agencies. Yeager (1991) documents a similar set of discretionary choices for enforcement of the Federal Water Pollution Control Act; negotiation, administrative action, and civil and criminal prosecution all have very different implications for EPA resource expenditures and for costs imposed on the firms found in violation of their permits. These authors and others [DiMento (1986), Hawkins (1984)] stress the importance of judicial attitudes, higher level support from the executive branch, and public opinion in determining the regulator's chance of a favorable ruling in litigation or administrative review and in the size of financial penalties imposed or the severity of injunctions granted. It is clear from experience that imposing penalties is expensive[3] and is dependent on institutional factors largely outside the control of both firms and regulators. Expenditures on these enforcement activities compete with monitoring as an effective use of the limited regulatory budget.

Taking explicit account of regulatory choices of monitoring and enforcement activities under budget limitations enables us to provide an economic interpretation of an important distinction that has consistently been drawn in the non-economic literature on enforcing social regulation: the difference between "deterrence" and "compliance" regimes. Observers have defined deterrence regimes as applying to discrete episodes of pollution wherein the intent of firms to actively conceal their actions from regulatory knowledge makes the discovery and verification of violations both difficult and expensive. When violations are discovered, they are met with little tolerance and are vigorously prosecuted.[4] Compliance regimes are characterized by more frequent contact between firms and regulators and by relative tolerance for intermittent noncompliance or some level of continuing noncompliance as part of a negotiated schedule for improvement.[5]

In this paper we investigate the interaction of regulators and firms in an environment where the courts (and other institutions) have a significant influence on penalties. Our model embodies an improved way of thinking about and representing penalty functions in the study

3. Any number of authors have documented the severity of resource constraints facing environmental regulators [Melnick (1983), DiMento (1986), Hawkins (1984), Yeager (1991)].

4. Hawkins (1984) discusses the episodic character of violations in deterrence systems. Reiss (1984) and DiMento (1986) discuss the high probability of legal and criminal sanctions which can occur under such a system.

5. Kagan (1988) and Reiss (1984) both define compliance regimes as embodying these characteristics.

of implementing social regulation. We find that regulatory choices of monitoring and enforcement activities where resources are limited depend crucially on the responsiveness of penalties to regulatory effort as well as to the cost of monitoring. We find that regimes with "compliance" characteristics are consistent with weak regulatory powers and relatively inexpensive monitoring. "Deterrence " regimes are consistent with expensive monitoring and the existence of strong societal sanctions for violations. Increased regulatory budgets without other institutional change will tend to make compliance regimes more effective but will not change their character.

When the population of regulated firms is characterized by differential costs of abatement, the regulator faces two budget allocation decisions: how to allocate scarce resources between firm types as well as between regulatory activities. The allocation of regulatory resources between firm types depends upon marginal compliance costs and the rate of change of these costs with respect to pollution levels. Firms with higher costs will tend to attract more regulatory effort and face higher penalties than firms with lower costs.

The paper proceeds as follows. In section 2 we characterize the environmental enforcement process. We analyze incomplete enforcement in a population of homogeneous firms in section 3. In section 4 we focus on differential regulatory treatment of a population of heterogeneous firms. Our conclusions and suggestions for extensions are presented in section 5.

2. Characterizing the Environmental Enforcement Process

We model environmental enforcement as a process involving three groups: the environmental protection agency or regulator, the judiciary and a population of n identical firms. We specify the goal of regulatory agencies responsible for monitoring and enforcement as minimizing pollution above firms' allowed standards. We do not model a policy process which compares environmental costs and benefits and chooses effluent levels. Rather, by characterizing the regulator as a violation-minimizing policeman, we can focus on the behavior of regulatory agencies whose primary goal is enforcement and not social welfare maximization.[6] We assume that the regulator and firms interact strategically in a Stackelberg subgame with complete information. The regulator commits to an enforcement policy and firms respond by choosing levels of noncompliance with the law.[7]

We model the judiciary as a non-strategic participant or a black-box technology in the environmental enforcement process. The role of the judiciary is to make decisions regarding the nature and extent of firms' deviation from the law: judicial decisions result in fines, $F(\cdot)$, being levied against noncompliant firms.[8] The relationship between the judiciary and the firms is characterized by incomplete information, that is, the judiciary does not observe firm behavior. A key role of the regulator is to resolve the information asymmetry between these two parties.

Let x be the amount of pollution in excess of the legal standard produced by a single

---

6. Other research has investigated enforcement with at least some attention to social welfare maximization, both in terms of setting standards in terms of monitoring probabilities and (costlessly imposed) fines [Viscusi and Zeckhauser (1979), Veljanovski (1984), Jones (1989), Jones and Scotchmer (1990), Harford and Harrington (1991), Harford (1991)]. While enforcement authorities will, to some extent, balance the costs and benefits of strict compliance in making enforcement decisions [Hawkins (1986); Kagan (1989)], enforcement authorities will generally seek stricter compliance than they are able to achieve.

7. We believe that the government and regulatory agencies, because of their power of taxation, law-making and coercion, can be sensibly thought of as being able to commit relative to private firms.

8. We recognize that the firm will usually elect to make expenditures to defend itself in court, to negotiate for a settlement, to mitigate adverse publicity with ad campaigns, etc. The function F() gives the firm's optimal combination of fines paid and other expenses undertaken to mitigate penalties given regulatory choices.

firm. The regulator can influence firm behavior by taking actions which cause the firm to face (expected) fines for the pollution it produces. Levying a sanction on any individual firm requires two separate expenditures on the part of the regulator: one to determine the firm's pollution level (monitoring) and one to bring about a fine given that monitoring reading (enforcement). We assume that the monitoring technology produces perfect measurements.

We differentiate monitoring from other enforcement activities. Monitoring is the process of discovering and verifying a firm's actual pollution levels.[9] It is required for third party verification, that is, to produce evidence that a firm is in violation of the law. Given perfect monitoring, x is both the actual amount of excess pollution and the regulator's produced evidence. The regulator must also incur costs in enforcing environmental regulations. Enforcement comprises the notification, administrative, and litigation-related expenditures which, contingent on a monitoring reading, affect firm profitability. Let α be the probability of monitoring and E be the regulator's expenditures on enforcement effort.

Monitoring serves at least two functions in the regulatory process. First, it is a necessary input into penalties: unless a firm has been found and documented to be in violation, the regulator will have no case and will be unable to litigate or negotiate with firms from a position of strength. Second, if a firm knows more about its behaviour and/or characteristics than the regulator, then it may have incentives to misrepresent itself to the regulator. Monitoring (at least partially) resolves the asymmetric information problem. Improving regulatory knowledge allows the regulator to better target its enforcement

9. The difficulty and expense of accurate monitoring are discussed in depth in Russell, Harrington, and Vaughan (1986). Their analysis (along with that of Harrington (1989), Jones and Scotchmer (1990), and Veljanovski (1984)) explores how a budget-constrained regulator will allocate its monitoring budget to achieve its goals.

resources.[10] Although monitoring certainly does have a value in revealing compliance information about firms, it is also true that regulatory authorities frequently know firm behavior without a formal monitoring procedure. Thus, in this paper we focus on monitoring as a verification of violation.

The key to our investigation of endogenous regulatory choice is our representation of the fine function. Judicial fines are based upon evidence of firm violation and regulatory enforcement expenditures: $F = F(x,E;\eta)$, where η is a vector of judicial parameters. The judicial parameters, η, reflect the socio-legal institutional environment within which firms and the regulator operate. For example, a pattern of legal decisions in enforcement cases in favor of firms will reduce the marginal effect on fines of regulatory effort. Similarly, enabling legislation with streamlined procedures for levying penalties will increase the responsiveness of fines to E. Similar considerations apply to the penalty function and firms' choices of pollution levels. Fines are not levied if a firm is in compliance with the standard or if the regulator does not prepare a case against a firm, that is, $F(x \leq 0,E) = F(x,0) = 0$. Fines are assumed to be strictly convex in the level of firm violation and strictly concave in regulatory enforcement effort: $F_x > 0$, $F_{xx} > 0$, $F_E > 0$ and $F_{EE} < 0$.[11] We also assume

10. Monitoring also allows better targeting of monitoring resources in subsequent time periods [see Harrington (1988), Harford and Harrington (1991), and Harford (1991)].

11. See Harford (1978), Jones (1989), Kambhu (1989), and Shaffer (1990) for discussion about the consequences of the second derivative of the fine function with respect to violation size. Our assumption implies that larger violations bring increasingly severe penalties at the margin, holding regulatory effort equal. We also assume that greater regulatory effort results in higher penalties, but there are diminishing returns to regulatory effort. Diminishing returns is a necessary assumption to avoid an outcome where the maximum deterrent is the regulator's threat of putting all regulatory resources into a single firm; such a result is analogous to an arbitrarily high penalty.

that the fine function exhibits the following property: $F_{xEx}/F_{xx} \geq F_{xE}/F_x$.[12]

The regulator announces a type-specific monitoring and enforcement strategy to minimize industrial noncompliance $\sum_i x_i$ subject to a budget constraint and firms' regulatory reaction functions. Firms choose levels of noncompliance to minimize expected costs of pollution control given the regulator's monitoring and enforcement strategies $\{\alpha, E\}$. Firms face two sets of pollution control costs - abatement costs and penalties. The firm's cost of abatement to achieve a given pollution level is represented by C(x). Abatement costs are the firm's least-cost combination of treatment equipment, process changes, and production cutbacks; the abatement cost function reflects the minimum subtraction from unregulated profits for a given effluent level. We assume that the abatement cost function is convex: $C_x(x) < 0$ and $C_{xx}(0) \geq 0$. The second set of costs are those associated with expected penalties imposed on firms as a result of violating pollution control laws. A firm's total expected cost associated with any pollution level is $C(x) + \alpha F(x, E; \eta)$.

To summarize, the timing of the environmental protection game is specified as follows: the regulator, acting as a Stackelberg leader, announces monitoring and enforcement strategies $\{\alpha, E\}$; firms choose a level of noncompliance x; a number of firms, αn, are monitored; monitored firms found in noncompliance with the law are taken to court; the regulator expends E preparing and presenting each actionable case; the court levies a fine F(x,E) against firms based on the level of violation and the strength of the regulator's case.

12. This restriction ensures that enforcement is a *normal* regulatory instrument, that is, enforcement expenditures per firm will not decline as budgetary resources per firm increase. It also ensures that $d\xi/dx \geq 0$, that is, a constant or increasing elasticity of the marginal fine function.

3. Enforcement with Homogeneous Firms

Suppose that the regulator must enforce an environmental standard among a population of n identical firms. The regulator recognizes that a firm's choice of pollution above the standard or the level of noncompliance will be influenced by the monitoring and enforcement policies announced by the regulatory agency. Consequently, the regulator will incorporate a representative firm's optimizing condition, or *regulatory reaction function*, into its own decision problem.

A representative firm's decision problem can be written as

$$\min_{x} \quad C(x) + \alpha F(x,E) \tag{1}$$

The firm's first-order condition is

$$C_x(x) + \alpha F_x(x,E) = 0 \tag{2}$$

A firm chooses the optimal level of noncompliance by equating the marginal cost of compliance with the expected marginal costs of noncompliance with the standard. Rewrite equation (2) as $f(x,\alpha,E;\theta,\eta) = 0$. Given that the second-order condition, $f_x > 0$, holds, (2) implicitly defines the firm's unique equilibrium strategy x^*, where $x\colon \{\alpha,E;\theta,\eta\} \rightarrow R^+$.

A firm's optimal noncompliance level is decreasing in the probability of monitoring and the level of enforcement expenditures. To see this, note that firm responses to changes in regulatory strategies are given by $x_\alpha = -f_\alpha/f_x < 0$ and $x_E = -f_E/f_x < 0$. Monitoring and enforcement activities have an *indirect* and a *direct* effect on expected fines. Both activities indirectly reduce expected fines by inducing firms' to increase compliance with the law - a pollution mitigating effect. Conversely, both activities directly increase expected fines - a penalty enhancing effect. The overall effect will always be to reduce noncompliance.

The regulator's design problem can be written as

$$\min_{\{\alpha,E\}} \quad nx^{*} \tag{3}$$

subject to

$$x^{*} = x(\alpha,E;\theta,\eta) \tag{4}$$

$$\alpha n\left[M + E\right] \leq B \tag{5}$$

where M is the cost of monitoring a firm and B is the regulator's fixed budgetary resources. The regulator chooses monitoring and enforcement strategies, $\{\alpha, E\}$, to minimize industrial noncompliance with environmental standards subject to a representative firm's regulatory reaction function given by equation (4) and a budget constraint given by equation (5). We assume that budgetary resources are sufficiently limited and/or the cost of monitoring sufficiently high to restrict attention to an interior monitoring solution, $0 < \alpha < 1$.

The first-order conditions for the regulator's design problem are given by

$$x_{\alpha} + \phi(M + E) = 0 \tag{6}$$

$$x_{E} + \phi\alpha = 0 \tag{7}$$

$$-\frac{B}{n} + \alpha(M + E) = 0 \tag{8}$$

where ϕ is the Kuhn-Tucker multiplier on the budget constraint. Equations (6) - (8) implicitly define the regulator's unique equilibrium monitoring and enforcement strategies

$\{\alpha, E\}$, where α: $\{n, M, B; \theta, \eta\} \rightarrow [0,1]$ and E: $\{n, M, B; \theta, \eta\} \rightarrow R^{+}$.

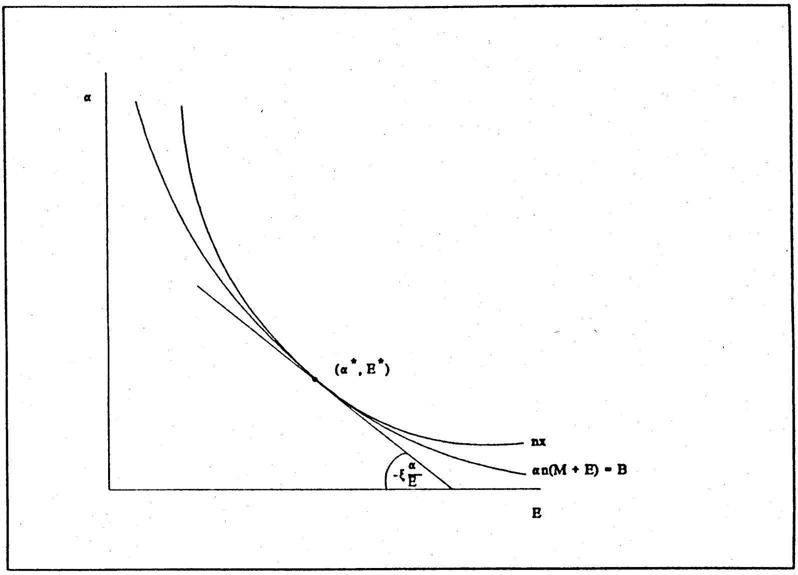

Figure 1. Optimal Regulatory Strategies

The optimal regulatory strategy is illustrated in Figure 1.[13] Rearranging equations (6) and (7) yields the following relationship:

$$-\frac{x_E}{x_\alpha} = -\frac{\alpha}{M + E} \tag{9}$$

The left-hand side of equation (9) is the slope of an iso-noncompliance curve nx^* - the

13. We assume that the following second-order condition for the required relative curvature of the two curves in Figure 1 holds:

$$\frac{2x_\alpha x_{E\alpha} - x_E x_{\alpha\alpha} - \frac{x_\alpha^2 x_{EE}}{x_E}}{x_E^2} > \frac{2(M + E)}{\alpha^2}$$

combinations of α and E to which firms react with identical noncompliance levels. The slope of the iso-noncompliance curve is the marginal rate of substitution between the two regulatory activities. The right-hand side of equation (9) is the slope of the regulator's budget constraint, which defines the feasible combinations of α and E. The regulator chooses monitoring and enforcement strategies so that the iso-noncompliance curve lies tangent to the budget constraint.

Define $\xi = E\, F_{xE}/F_x$ as the elasticity of marginal penalties with respect to regulatory enforcement expenditures. ξ reflects the power of enforcement expenditures to alter firm incentives. ξ can be interpreted as the "returns to scale" from regulatory enforcement activity. It reflects judicial willingness to impose large fines or support the regulator on statutory interpretations as well as the strength of public opinion and political support for strict enforcement. Higher values of ξ increase (reduce) the effectiveness of an additional dollar spent on enforcement (monitoring) activities, resulting in increased enforcement expenditures and a lower monitoring probability.[14] The slope of the line tangent to the iso-noncompliance curve and the budget constraint at the equilibrium point is - $\xi\ \alpha/E$.

The equilibrium characteristics of the regulatory monitoring and enforcement strategies can be obtained from equations (6) - (8). If the budgetary resources of the regulatory agency are increased, then the regulator will monitor a higher proportion of firms and either increase or keep constant per-firm enforcement expenditures. Firms will respond to increases in budgetary resources by reducing noncompliance. An increase in the number

14. To see this, define $x_\alpha^* = -\dfrac{F_x}{f_x}$ *and* $x_E^* = -\alpha \cdot \dfrac{F_{xE}}{f_x}$. Note that $\partial x_\alpha/\partial\xi < 0$ and $\partial x_E/\partial\xi > 0$.

of firms in the regulated population will lead to a lower proportion of firms being monitored while per-firm enforcement expenditures will either decline or remain constant. The level of firm noncompliance will correspondingly increase with the number of firms. If the cost of monitoring increases, then fewer firms will be monitored and the regulator will increase expenditures on enforcement activities.

Proposition 1: *The optimal regulatory strategies $\{\alpha^*, E^*\}$ with homogeneous firms are characterized by the following equations:*

$$\alpha^* = (1 - \xi)\frac{B}{nM} \tag{10}$$

$$E^* = \frac{\xi}{1 - \xi}M \tag{11}$$

Proof: See Appendix.

The optimal rules for regulatory monitoring and enforcement policies are simple functions of ξ. As the regulator's support from the judiciary and other institutions gets stronger (reflected through increases in ξ), the iso-noncompliance curves become steeper and the budget constraint remains unchanged. The equilibrium shifts to the southeast in Figure 1 and corresponds to a lower noncompliance level. Maximum deterrence is achieved by monitoring fewer firms and aggressively pursuing sanctions against noncomplying firms.

Increases in monitoring costs leaves iso-noncompliance curves unchanged but flattens and shifts the budget constraint towards the origin; the equilibrium again shifts southeast in Figure 1 to more of a low-monitoring, aggressive sanctions enforcement regime characterized by a higher noncompliance level. This situation fits the stylized facts of deterrence regimes quite well. When either the harmful activity or the identity of the perpetrator can be hidden from regulators, successful monitoring becomes quite expensive. When the environmentally damaging activity is regarded as criminal (for example, the "moonlight dumping" of hazardous wastes) penalties have frequently included criminal prosecution, sizeable fines, and damage assessments [Russell (1990)]. The regulator gets greater support from enabling legislation and the legal process for these more serious offenses.

Suppose that ξ is low because of ambivalence about the desirability of strict enforcement relative to the harm of plant closings, political power of the regulated industry, or a high probability of leniency due to unclear or ambiguously worded regulations. The regulator will respond to low values of ξ by choosing more frequent contact with firms (higher α) but less aggressive sanctioning activities (lower E). Less expensive monitoring also leads to a high-monitoring, low-sanction regulatory regime. This describes a compliance regime, where contact with regulators is fairly frequent and low sanctions are observed for violations [Hawkins (1984)].

The budgetary allocation rules are also simple functions of ξ. Let μ_M and μ_E be expected expenditures on monitoring and enforcement, respectively. Equations (10) and (11) can be rewritten as expected budgetary allocation rules $\{\mu_M^*, \mu_E^*\}$ as follows:

$$\mu_M^* = (1 - \xi)B \tag{12}$$

$$\mu_E^* = \xi B \tag{13}$$

The regulator will allocate a larger share of the budget to enforcement activities the more responsive the fine function is to these expenditures.

The optimal regulatory strategies and budgetary allocation rules are cost contingent provided that judicial support depends on the level of noncompliance. In other words, if ξ is a function of x, then the regulatory policies will depend on the cost parameters. However, if ξ is independent of x, then the optimal regulatory strategies and budgetary allocation rules will be *independent of firms' cost parameters*. To see this, define a multiplicative fine function $F = G(x)H(E)$ so that $\xi = E \cdot H_E / H$ and is constant with respect to x. If the fine function is multiplicative, then the regulator's optimal enforcement expenditure per firm will also be independent of budgetary resources and the number of firms.

If ξ is relatively unresponsive to the level of violation, then changing the regulatory budget will change the frequency of monitoring but have little effect on the resources spent on enforcement. In the case of a multiplicative fine function, the budget constraint will shift upwards in Figure 1, and lie tangent to a curve representing lower noncompliance levels at an equilibrium with a higher monitoring probability and an unchanged level of per-firm enforcement expenditures. Deterrence regimes and compliance regimes will both be characterized by higher monitoring probabilities in response to budget increases, but the character of the regime, as defined by the aggressiveness of enforcement, will not change.

4. Enforcement with Heterogeneous Firms

In this section we consider a population of firms with heterogeneous abatement costs. Relaxing the assumption of identical firms creates an additional decision problem for the regulator: differential firm treatment. The fixed budget has to be apportioned not only between monitoring and enforcement activities but also among different classes of firms. We differentiate firms solely on an abatement cost basis; the vector of cost technology parameters, θ, is type-dependent.[15] We restrict attention to two firm types, θ_L and θ_H, where the subscripts L and H denote low abatement costs and high abatement costs, respectively. We define a high-cost firm as the firm type with higher marginal abatement costs at the amount of pollution allowed by the standard. The industrial population consists of $n\lambda_L$ low-cost firms and $n\lambda_H$ high-cost firms. The identity of firms is assumed known to the regulator.

The regulatory strategy is a quadruple: $\{\alpha_L,E_L,\alpha_H,E_H\}$. The regulator's design problem can be written as

$$\min_{\alpha_L,E_L,\alpha_H,E_H} \quad n\left[\lambda_L\, x_L^* + \lambda_H\, x_H^*\right] \tag{14}$$

subject to

$$x_L^* = x_L(\alpha_L,E_L;\theta,\eta) \tag{15}$$

$$x_H^* = x_H(\alpha_H,E_H;\theta,\eta) \tag{16}$$

$$n\left[\alpha_L\lambda_L\left(M + E_L\right) + \alpha_H\lambda_H\left(M + E_H\right)\right] \leq B \tag{17}$$

The regulator will choose type-dependent monitoring and enforcement strategies to minimize industrial noncompliance subject to low- and high-cost firms' regulatory reaction functions given by equations (15) and (16) and the budget constraint given by equation (17). The regulatory reaction functions are defined such that the sum of marginal costs of compliance

15. It is certainly possible that firms face fine functions with type-dependent socio-legal parameters, η_k, depending upon factors like sensitivity to publicity or political strength. We assume that η is constant for all firms. One can interpret this assumption as equal treatment before the courts or judicial fairness.

and expected marginal costs of noncompliance are equated across firm types in equilibrium. The first-order conditions for the regulator's design problem are given in the Appendix.

The main features of the solution are as follows: the regulator will choose monitoring probabilities to equate the weighted marginal noncompliance with respect to monitoring across firm types, where the weights are given by the inverse of the per-firm regulatory expenditure. The regulator will choose enforcement expenditures so as to equate the weighted marginal noncompliance with respect to enforcement expenditures across firm types, where the weights are given by the inverse probability of being monitored.

Alternatively, we can interpret the regulator's decisions as follows: when choosing monitoring and enforcement strategies the regulator is simultaneously dividing the budget between firm types. Let $B = B_L + B_H$, where $B_i = n\lambda_i\alpha_i(M + E_i)$ for $i = L,H$ are the expected budgetary resources allocated to the population of i-type firms. The regulator will choose type-dependent monitoring and enforcement strategies, $\{\alpha_i, E_i\}$, so that the iso-noncompliance curve, $n\lambda_i x_i$, lies tangent to the regulator's budget curve, B_i, for that industry group. The regulator effectively *separates* the twin problems of allocating budgetary resources across firm types and across regulatory activities. Separation is possible because of the assumptions of complete information and nonstrategic interaction between firms.

Proposition 2: *The optimal regulatory strategies $\{\alpha_L^*, E_L^*, \alpha_H^*, E_H^*\}$ with heterogeneous firm types are characterized by the following equations:*

$$\alpha_i^* = (1 - \xi_i)\,\frac{B_i}{\lambda_i n M} \tag{18}$$

$$E_i^{*} = \frac{\xi_i}{1 - \xi_i} M \tag{19}$$

where $B_i = \frac{\lambda_i}{\Lambda}\left[F_{x_i}F_{x_j x_j}B - \lambda_j n M \Xi_i\right]$, $\Lambda = \lambda_L F_{x_L} F_{x_H x_H} + \lambda_H F_{x_H} F_{x_L x_L}$, $\Xi_i = \frac{C_{x_i x_i} F_{x_j}}{1 - \xi_i} - \frac{C_{x_j x_j} F_{x_i}}{1 - \xi_j}$,

and $i,j = L,H$ *for* $i \neq j$.

Proof: See Appendix.

The solutions for the optimal regulatory strategies for heterogeneous firms appear similar to those derived for homogeneous firms and exhibit the identical equilibrium properties with respect to own parameters and the exogenous variables. In particular, the type-elasticity of the marginal penalty function, ξ_i, plays the same role in determining monitoring and enforcement strategies as in the previous section. However, regulatory strategies now depend, in a complex way, upon both the socio-legal parameters and firms abatement technology parameters.

The term which summarizes the interaction between the socio-legal and technology parameters is B_i, the type-contingent enforcement budget. We can decompose B_i into three terms: a scaling factor, λ_i/Λ, that depends on the relative curvature of the fine functions and population proportions; the budget weighted by the product of the first derivative of one firm's fine function and the second derivative of the other firm's fine function, $F_{x_i}F_{x_j x_j}B$; and a term involving weighted slopes of marginal cost curves, $\lambda_j n M \Xi_i$.

Per-firm enforcement expenditures again depend upon the judicial or other institutional support for enforcing pollution control laws. If ξ depends upon firm compliance behavior, then the regulator will expend more resources enforcing compliance among the firms that yield a higher return to the regulatory dollar. Note, however, that enforcement expenditures are interdependent: ξ_i depends on the j-type firm's technological parameters since x_i is a function of B_j. Institutional support for an i-type firm is constant or increasing with a j-type firm's noncompliance behavior if and only if the following inequality holds:

$$\frac{F_{x_i E_i x_j}}{F_{x_i x_j}} \geq \frac{F_{x_i E_i}}{F_{x_i}} \tag{20}$$

If ξ is independent of firm compliance behavior and constant across firm types, then the optimal per-firm enforcement strategy is identical to that obtained with homogeneous firms: enforcement expenditures are independent of cost parameters and equal across firm types. If the elasticity of the marginal fine function is relatively unresponsive to the level of noncompliance, then there will be a narrower range of enforcement expenditures among different firms than the range of monitoring probabilities. When ξ varies greatly with pollution levels, a much greater disparity between enforcement expenditures will be optimal.

Referring to Figure 2, when E is fixed, the regulatory budget constraint in monitoring probability space is linear. The slope of the budget line is given by the negative ratio of the population proportions, $-\lambda_H/\lambda_L$. Totally differentiating the industry noncompliance curve, we obtain a slope of $-\lambda_H x_{H\alpha_H}/\lambda_L x_{L\alpha_L}$. The optimal combination of monitoring frequencies is given by the tangency point of the budget line and the industrial noncompliance curve.

If the proportion of high-cost firms in the industrial population declines, then the regulator will shift budgetary resources from high- to low-cost firms. The marginal impact of the regulatory dollar is, however, higher for low-cost firms. Hence, the regulator can increase the frequency of monitoring both firm types and, correspondingly, reduce the level

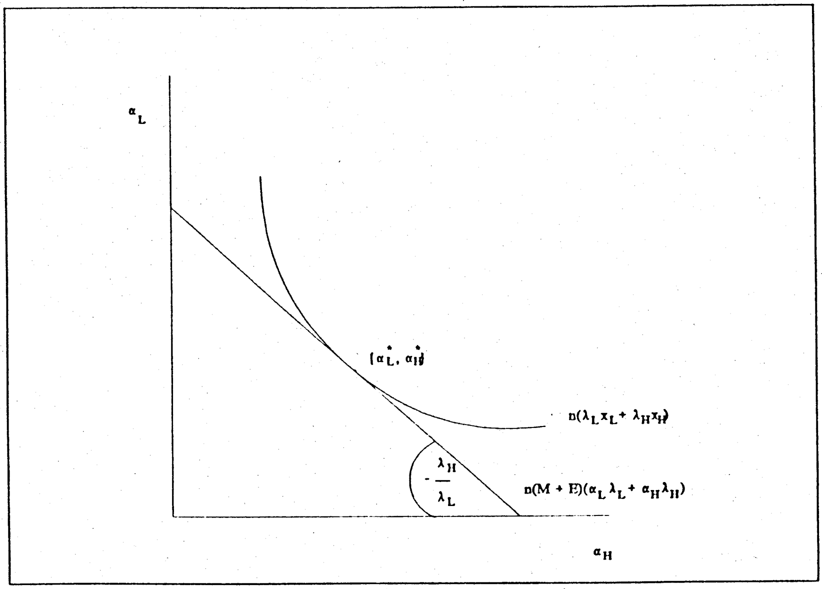

Figure 2. Optimal Monitoring Strategies for Heterogeneous Firms.

of industrial noncompliance. In Figure 2, the slopes of the budget line and the iso-noncompliance curve decline, shifting the equilibrium to the northeast.

The regulator allocates the budget between firm types and between regulatory activities. The key determinants of budget allocation are the degree of institutional support for enforcement against different levels of noncompliance and the relationship between firms' technological ability to abate pollution. The regulator's micro or intrafirm budgetary allocation rules are given by:

$$\mu_{M_i}^* = (1 - \xi_i)B_i \tag{21}$$

$$\mu_{E_i}^{*} = \xi_i B_i \tag{22}$$

for i=L,H.

The regulator's macro budgetary allocation rules $\{\mu_M^*, \mu_E^*\}$ that divide the budget between the two regulatory activities are defined by the following equations:

$$\mu_M^* = (1 - \xi_L)B_L + (1 - \xi_H)B_H \tag{23}$$

$$\mu_E^* = \xi_L B_L + \xi_H B_H \tag{24}$$

If institutional support for enforcing compliance declines for either or both firm types, then regulatory resources will be shifted to monitoring activities. If ξ is constant across firm types, then the macro rules are independent of firm cost parameters and are identical to the rules obtained with homogeneous firms.

An Example of Optimal Monitoring Strategies

The choice of monitoring strategies depends largely on B_i. An example is provided to further comparative static properties and relative magnitudes of monitoring strategies for heterogeneous firms. We assume quadratic costs and a Cobb-Douglas fine function:

$$C_i(x) = K_i - c_i x_i + \frac{\omega_i}{2} x_i^2 \tag{25}$$

$$F(x,E) = \frac{A}{2} x^2 E^{\beta} \tag{26}$$

where $c > \omega$ and $0 < \beta < 1$. The parameter A reflects the responsiveness of the fine function to changes in both choice variables.

The optimal type-dependent monitoring strategy is given by equation (27):

$$\alpha_i^* = (1 - \beta)\frac{B_i}{\lambda_i n M} \tag{27}$$

where $B_i = \frac{\lambda_i}{\Delta}\left[\sqrt{c_i}B - \lambda_j n M \frac{b}{(1-\beta)}\left(\sqrt{c_j}\,\omega_i - \sqrt{c_i}\,\omega_j\right)\right]$, $\Delta = \lambda_L\sqrt{c_L} + \lambda_H\sqrt{c_H}$

and $b = \frac{1}{A}\left(\frac{1-\beta}{\beta M}\right)^{\beta} = \frac{1}{F_{xx}}$.

The choice of monitoring strategies depends upon the relationship between the intercepts and slopes of the firms' respective marginal cost functions. To see this, we can write the probability of monitoring a high-cost firm as a function of a low-cost firm's monitoring probability as follows:

$$\alpha_H^* = \sqrt{\frac{c_H}{c_L}}\,\alpha_L^* + b\left(\sqrt{\frac{c_H}{c_L}}\,\omega_L - \omega_H\right) \tag{28}$$

If the intercept of a high-cost firm's marginal cost curve, c_H, increases relative to a low-cost firm's, then the regulator must reallocate monitoring resources toward the high-cost firm to maintain an efficient allocation. Suppose that the regulator does not adjust the monitoring probabilities. The marginal costs of compliance are greater than the expected marginal costs of noncompliance at the statusquo, inducing high-cost firms to increase their pollution level. But at the new equilibrium $x_{L\alpha_L} > x_{H\alpha_H}$ - hence, the regulator is allocating too few resources to H-type firms, resulting in their overproduction of pollution. The regulator must raise the expected marginal penalty faced by a high-cost firm to minimize industrial noncompliance. Thus, the regulator will increase α_H and reduce α_L. As a result of the increase in c_H, B_H and total industrial noncompliance are higher. Referring to Figure 2, the slope of the budget line remains the same while the industrial noncompliance curve flattens and shifts down and to the right, shifting the equilibrium to the southeast.

An increase in the slope of the marginal abatement cost curve has the opposite effect on the regulator's allocation decision. If there is no change in regulatory strategies, then at the statusquo level of pollution the marginal costs of compliance are less than the expected marginal costs of noncompliance. H-type firms will respond by reducing noncompliance, resulting in $x_{L\alpha_L} < x_{H\alpha_H}$. The regulator must increase the level of noncompliance of H-type firms by reducing the slope of the expected marginal penalty function. This is achieved by increasing α_L and reducing α_H. The reallocation of budgetary resources results in lower levels of noncompliance from both firm types. Referring to Figure 2, the slope of the iso-noncompliance curve increases while the slope of the budget constraint remains unchanged, shifting the equilibrium to the northwest.

A low-cost firm will be monitored with a higher frequency than a high-cost firm if and only if the following condition holds:

$$\alpha_L^{*} F_{xx} < - \frac{\sqrt{c_H}\,\omega_L - \sqrt{c_L}\,\omega_H}{\sqrt{c_H} - \sqrt{c_L}} \tag{29}$$

Note that the condition given by equation (29) can hold if $c_H > c_L$ and $\omega_H > \omega_L$[16]. Larger values of ω are associated with abatement cost structures where marginal changes in pollution

16. Jones and Scotchmer (1990) find that higher-cost firms face lower monitoring probabilities. Their result relies on the assumptions that a high-cost firm's marginal cost curve cuts a low-cost firm's marginal cost curve from above and that firms have identical pollution levels when marginal costs equalled zero. Equation (29) would be violated under these assumptions.

cause larger changes in *marginal* abatement costs. Higher monitoring probabilities for low-cost firms will occur only when marginal abatement costs are much more responsive to changing pollution levels for high-cost firms. The larger the difference in intercept parameters, the greater the difference in slopes necessary to produce this result. If $\omega_H \leq \omega_L$, high-cost firms will always face higher monitoring probabilities.

5. Conclusion

In this paper we develop a model of enforcement in which penalties for noncompliance are determined by strategic interaction between a regulatory agency and an industry. We demonstrate that the balance between monitoring and sanctioning activities depends crucially on the power of the regulator's enforcement technology. This power, which can be measured by the elasticity of changes in the firm's perceived marginal penalties with respect to regulatory effort, reflects the political, legal, and social attitudes toward the regulation being enforced. The model indicates that severe offenses will be characterized by low detection rates and vigorous enforcement while pollution regarded more ambiguously by society at large will be characterized by frequent contact with regulators and relatively mild sanctioning activity. This result is consistent with observed behavior.

The paper shows how the penalties faced by firms for violating pollution limits depend on their own violations, regulatory authority and budget constraints, and the regulator's choices about monitoring and enforcement expenditures. Our analysis indicates that if the consensus about the desirability of improved environmental quality continues to broaden and noncompliance with effluent regulations and operating permits is seen more as criminal and less as the subject of negotiation, then environmental enforcement will tend toward more of a deterrence regime. Where such a consensus fails to take hold, most likely in areas where there is broad disagreement about the desirability of environmental regulation or the definition of compliance behavior, compliance regimes are likely to continue.

In our model, when the regulatory agency must enforce environmental laws on firms with different abatement cost structures, they choose strategies to equalize marginal noncompliance across firm types. The model indicates that firms with more expensive abatement costs will tend to be monitored more frequently and face higher penalties than low-cost firms. Our result suggests that in a world of privately informed firms, firms may want the regulator to believe they have low compliance costs in order to face reduced regulatory effort.[17]

This paper establishes some of the tradeoffs in regulatory strategy when levying penalties is costly and compliance is not an either/or proposition. Interesting extensions of our model include incorporating optimal standard-setting and extending the model to a multi-period game. Our model could also be extended to the case where firms have better information about their costs than the regulator. Monitoring would serve both as a necessary input into sanctioning and as a resolution of asymmetric information.

APPENDIX

Proof of Proposition 1: Dividing equations (6) and (7) in the text yields

17. Note that firms have an incentive to overstate costs in the literature on optimal rule-setting in environmental regulation with asymmetric information [see Baron (1985) and Spulber (1988)].

$$\frac{x_\alpha}{x_E} = \frac{M + E}{\alpha} \tag{A1}$$

Equation (A1) can be rewritten as follows

$$\frac{F_x}{F_{xE}} = M + E \tag{A2}$$

Rearranging equation (8) in the text yields

$$M + E = \frac{B}{\alpha n} \tag{A3}$$

Substituting (A3) into (A2) yields the solution for the optimal monitoring strategy

$$\alpha^* = \frac{F_{xE}}{F_x}\frac{B}{n} \tag{A4}$$

Substituting (A4) into (A3) yields the solution for the optimal enforcement expenditure strategy

$$E^* = \frac{F_x}{F_{xE}} - M \tag{A5}$$

Multiplying the right-hand sides of equations (A4) and (A5) by E/E and rearranging yields the solutions in terms of ξ.

Q.E.D.

Equilibrium Conditions for Heterogeneous Firms:

The first-order conditions corresponding to the regulator's design problem are:

$$x_{L\alpha_L} + \phi(M + E_L) = 0 \tag{A6}$$

$$x_{H\alpha_H} + \phi(M + E_H) = 0 \tag{A7}$$

$$x_{LE_L} + \phi\alpha_L = 0 \tag{A8}$$

$$x_{HE_H} + \phi\alpha_H = 0 \tag{A9}$$

$$-\frac{B}{n} + \left(\alpha_L\lambda_L(M + E_L) + \alpha_H\lambda_H(M + E_H)\right) = 0 \tag{A10}$$

Equations (A6) - (A10) implicitly define the unique equilibrium regulatory strategies $\{\alpha_L^*, E_L^*, \alpha_H^*, E_H^*\}$, where α_i: $\{n,M,B;\theta,\eta\} \rightarrow [0,1]$ and E_i: $\{n,M,B;\theta,\eta\} \rightarrow R^+$, for i = L,H. Note that the regulatory strategies for an i-type firm will depend upon a j-type firm's parameters, that is, there is technological interdependence among the strategies.

Proof of Proposition 2: Divide equations (A6) and (A8) and equations (A7) and (A9) to yield the following relationship

$$M + E_i = \frac{F_{x_i}}{F_{x_iE_i}} \tag{A11}$$

for i = L, H. Multiplying the right-hand side of equation (A11) by E_i/E_i and rearranging yields the optimal per-firm enforcement expenditure given by equation (19) in the text.

Substitute (A11) for i = L, H into equations (A6) and (A7) to yield

$$\frac{F_{x_iE_i}}{C_{x_ix_i} + \alpha_i F_{x_ix_i}} = -\phi \tag{A12}$$

Equate (A12) across firm types and rearrange to define α_j as a function of α_i

$$\alpha_j = \frac{F_{x_jE_j}}{F_{x_jx_j}F_{x_iE_i}}\left(C_{x_ix_i} + \alpha_i F_{x_ix_i}\right) - \frac{C_{x_jx_j}}{F_{x_jx_i}} \tag{A13}$$

for i,j = L, H and i $\neq$ j. Substitute equation (A11) for both firm types and equation (A13) for a j-type firm into the budget constraint, defined by equation (A10). After some manipulation we can derive the optimal solution for monitoring an i-type firm as

$$\alpha_i^{*} = \frac{F_{x_iE_i}}{\Lambda}\left[F_{x_jx_j}\frac{B}{n} - \lambda_j F_{x_j}\left[\frac{C_{x_ix_i}}{F_{x_iE_i}} - \frac{C_{x_jx_j}}{F_{x_jE_j}}\right]\right] \tag{A14}$$

where Λ is defined in the text. Define B_i as follows

$$B_i^{*} = \frac{\lambda_i}{\Lambda}\left[F_{x_i}F_{x_jx_j}B - n\lambda_j F_{x_i}F_{x_j}\left[\frac{C_{x_ix_i}}{F_{x_iE_i}} - \frac{C_{x_jx_j}}{F_{x_jE_j}}\right]\right] \tag{A15}$$

We can rewrite the optimal monitoring strategy defined by equation (A14) in terms of B_i as follows

$$\alpha_i^{*} = \frac{F_{x_iE_i}}{F_{x_i}}\frac{B_i}{\lambda_i n} \tag{A16}$$

Multiplying equation (A16) by E_i/E_i, substituting equation (A11) and some manipulation yields the desired result.

Q.E.D.

REFERENCES

Baron, David. 1985, Regulation of prices and pollution under incomplete information, Journal of Public Economics 28, 211-232.

DiMento, Joseph F., 1986, Environmental Law and American Business, Plenum Press, NY.

Harford, Jon D., 1978, Firm behavior under imperfectly enforceable pollution standards and taxes, Journal of Environmental Economics and Management 5, 26-43.

Harford, Jon D., 1991, Measurement error and state-dependent pollution control enforcement, Journal of Environmental Economics and Management 21, 67-81.

Harford, Jon D. and Winston Harrington, 1991, A reconsideration of enforcement leverage when penalties are restricted, Journal of Public Economics 45, 391-395.

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

Hawkins, Keith, 1984, Environment and enforcement: regulation and the social definition of pollution, (Clarendon Press, Oxford).

Jones, Carol A. and Suzanne Scotchmer, 1990, The social cost of uniform regulatory standards in a hierarchical government, Journal of Environmental Economics and Management 19, no. 1, 61-72.

Jones, Carol Adair, 1989, Standard setting with incomplete enforcement revisited, Journal of Policy Analysis and Management 8, no. 1, 72-87.

Kagan, Robert A., 1989, Understanding regulatory enforcement, Law & Policy 11, 89-119.

Kambhu, John, 1989, Regulatory standards, noncompliance, and enforcement, Journal of Regulatory Economics 1, no.2, 103-114.

Melnick, R. Shep, 1983, Regulation and the courts: the case of the clean air act (Brookings, Washington, DC).

Polinsky, A.M. and S. Shavell, 1979, The optimal tradeoff between the probability and magnitude of fines, American Economic Review 69, 880-891.

Reiss, Albert J., Jr, 1984, Selecting strategies of social control over organizational life, in Enforcing regulation, Keith Hawkins and John M. Thomas, editors (Kluwer-Nijhoff Publishing, Boston, Mass).

Russell, Clifford S., W. Harrington, and W.J. Vaughan, 1986, Enforcing pollution control laws (Johns Hopkins University Press, Baltimore, MD).

Russell, Clifford S., 1990, Monitoring and enforcement *in* Public policies for environmental protection, Paul R. Portney, editor (Resources for the Future, Washington, DC).

Shaffer, Sherrill, 1990, Regulatory compliance with non-linear penalties, Journal of Regulatory Economics 2, no 1 (99-103).

Spulber, Daniel, 1988, Optimal environmental regulation under asymmetric information, Journal of Public Economics 35, 163-181.

Stigler, G.J., 1970, The optimum enforcement of laws, Journal of Political Economy, 526-536.

Veljanovski, Cento G., 1984, The economics of regulatory enforcement", *in* Enforcing Regulation, Keith Hawkins and John M. Thomas, editors (Kluwer-Nijhoff Publishing, Boston, Mass).

Viscusi, W. Kip and R.J. Zeckhauser, 1979, Optimal standards with incomplete enforcement, Public Policy 27 no. 4, 437-456.

Yeager, Peter C., 1991, The limits of law: the public regulation of private pollution (Cambridge University Press, New York, NY).