Institute for International Economic Policy Working Paper Series
Elliott School of International Affairs
The George Washington University

The Desirability of Forgiveness in Regulatory Enforcement

IIEP-WP-2008-14

Arun Malik
George Washington University

February 4, 2007

Institute for International Economic Policy
1957 E St. NW, Suite 502
Voice: (202) 994-5320
Fax: (202) 994-5477
Email: **iiep@gwu.edu**
Web: **www.gwu.edu/~iiep**

The Desirability of Forgiveness in Regulatory Enforcement*

Arun S. Malik

Department of Economics
George Washington University
1922 F Street, NW
Washington, DC 20052, USA
Tel +1 202 558 2208
Fax +1 815 361 9126
amalik@gwu.edu

February 4, 2007

I present a model that explains two common features of regulatory enforcement: selective forgiveness of noncompliance, and the collection of information on a firm's compliance activities and not just its compliance status. I show that forgiving noncompliance is optimal if the information on a firm's compliance activities constitutes sufficiently strong evidence of the firm having exerted a high level of compliance effort. The key benefit of forgiving noncompliance is a reduction in the probability with which the firm needs to be monitored. If fines are costly, a further benefit is a reduction in fine costs.

Keywords: enforcement of regulation, selective enforcement, forgiving noncompliance

JEL Classification: L51, K42, K32, D86

*I thank Roberto Samaniego and Jorge Soares for helpful comments and suggestions on an earlier version of this paper. I am also grateful to participants of the 2006 South Asia Meetings of the Econometric Society for their comments. Any remaining errors are my own.

1. Introduction

Economic models of regulatory enforcement typically assume that penalties for detected noncompliance are certain and swift. Regulations are rigidly enforced and noncompliance when detected is never forgiven. In the language of the policy literature on regulation, economic models typically embody the "deterrent approach" to enforcement, first formalized by Becker (1968) and Stigler (1970). Regulatory scholars have argued that this approach has proven ineffective at securing compliance, and needlessly fosters an adversarial relationship between regulators and the regulated (Bardach and Kagan 1982). The approach advocated by a number of scholars is the "cooperative approach." This approach favors selective enforcement that takes into account the circumstances of observed violations (Scholz 1984) [1] In particular, under the cooperative approach, "regulators are more inclined to interpret rules flexibly, particularly when they believe firms are making good faith efforts to comply" (Harrison 1995). The cooperative approach, it is argued, is more likely to result in cost-effective enforcement and desirable regulatory outcomes (Bardach and Kagan 1982).

Cross-country comparisons of "regulatory styles" in the enforcement of health, safety, and environmental regulations indicate that the cooperative approach is common in Western Europe, Japan and Canada, whereas the deterrent approach is common in the U.S. (Harrison 1995, Kitamura 2000, Hawkins 1984). Even in the U.S., studies of enforcement find evidence of selective enforcement in a range of regulatory settings. For example, guidelines for enforcement of water pollution regulations, explicitly recommend that agencies be willing to forgive short-term violations, and that criminal fines be reserved for willful or negligent violations (Russell et al. 1986, Chapter 3, and Code of

[1] As Scholz and others have observed, the differences between the two approaches are not always clear cut. There is also considerable variation in terminology: the cooperative approach is also referred to as the "compliance" or "bargaining" approach, and the deterrent approach as the "sanctioning" or "rule-oriented" approach, among other possibilities. A third approach that has gained favor in recent years is "responsive regulation" (Scholz, Ayres and Braithwaite 1992, Nielsen 2006). This approach is, roughly speaking, an amalgam of the first two: the regulator first adopts a cooperative posture, but resorts to deterrence if the firm proves recalcitrant.

Federal Regulations 2004).[2]

In this paper, I show that a simple deterrence-based model of regulatory enforcement can be extended to allow for selective enforcement that takes into account a firm's efforts to comply. In the model, the regulator forgives noncompliance if it is able to obtain sufficiently strong evidence that the regulated firm had exerted a high level of compliance effort. This evidence takes the form of an additional signal of the firm's effort acquired by the regulator at some cost. This, possibly multi-dimensional, signal represents the results of inspections of a firm's records, interviews with its employees, or evaluations of its processes and procedures for achieving compliance. Collection of such information is a common element of enforcement practices.

The key benefit of forgiving noncompliance is, rather surprisingly, a reduction in the probability with which the firm needs to be monitored. The additional signal of the firm's compliance effort, coupled with selective forgiveness, allows for an increase in the power of the regulator's penalty scheme. Thus, the model provides a simple efficiency-based justification for selective enforcement based on an assessment of a violator's efforts to comply. This is accomplished in the context of a (noncooperative) principal-agent model with moral hazard.

The paper is related to several different strands in the existing literature on regulatory enforcement. To the extent that it provides a justification for selective enforcement, it is related to work by Garvie and Keeler explaining the variation in enforcement styles in the United States. They analyze a model in which fines for noncompliance depend on the budget-constrained regulator's enforcement effort and on the characteristics of the judicial process. They show that the regulator's optimal enforcement style depends on the cost of monitoring and on the responsiveness of judicially-determined fines to the regulator's enforcement effort. The optimal enforcement style

[2]Further evidence of selective enforcement in the U.S. can be found in Harrington (1988), who describes studies of U.S. environmental enforcement that reveal penalties are rarely imposed when violations are discovered. This is not simply a manifestation of lax enforcement, since compliance rates appear to be high. Hunter and Wasserman (1996) present similar results, and report that for 30 percent of the pollution violations examined, no action was taken by regulators beyond telephone calls and meetings with firms. Lofgren (1989) in a study of OSHA enforcement of safety and health regulations, reports that violations routinely go unpunished, yet OSHA enforcement has been deemed quite effective (Weil 1996). For evidence of selective enforcement in Europe, see, for example, the fascinating account by Hawkins (1984) of field-level environmental enforcement in Britain.

resembles the cooperative approach when monitoring is inexpensive and the responsiveness of fines is limited (rendering fines more costly for the regulator); and resembles the deterrent approach when monitoring is expensive and the responsiveness of fines is high.

The paper is also related to work by Jost (1997) on optimal enforcement when penalties are subject to judicial appeal. He considers a model with deterministic compliance in which a regulator must allocate its limited budget between two activities: monitoring, which yields an imperfect signal of a firm's compliance status, and investigations, which reveal compliance status perfectly. Investigations are conducted only when a firm appeals penalties for noncompliance in court. The appeal can result in penalties being waived.

The paper is also related, in spirit, to work by Nyborg and Telle (2004) that offers an explanation for the warnings frequently issued by regulators upon detecting noncompliance. In their model, a budget-constrained regulator attempts to control the behavior of firms that may be noncompliant by mistake. Firms detected in noncompliance are issued warnings and can come into compliance by incurring a "verification cost" in addition to the cost of compliance they would have incurred had they been compliant from the outset. In this setting, warnings can reduce the probability of accidentally switching from a full-compliance equilibrium to a no-compliance equilibrium.

The paper is arguably most closely related to a body of literature that attempts to explain the paradox of widespread compliance with environmental regulations despite infrequent punishment of noncompliance. Harrington (1988) shows that this paradox may be explained by a state-dependent enforcement strategy in which penalties depend not only on current noncompliance but also on past noncompliance. Livernois and McKenna (1999) provide an alternative explanation based on a model in which firms must self-report their compliance status. Heyes and Rickman (1999) argue that the paradox may reflect "regulatory dealing," with the regulator choosing to forgive a firm's noncompliance in one context (e.g., air pollution control), in exchange for compliance in another (e.g., water pollution control). In all three of these papers, the optimal penalty for some types of noncompliance may be zero, in other words, some types of noncompliance may be forgiven.

The model presented here is similar to those in the literature just mentioned in that it yields conditions under which setting the fine for noncompliance equal to zero is optimal. But it differs in a number of important respects. The model explicitly attempts to capture the key principle underlying selective enforcement: that enforcement should take into account the circumstances of a violation, in particular the extent of a violator's efforts to comply. This is not true of any of the existing literature.

The model also differs from those in the existing literature in that it incorporates, and justifies, the practice of collecting information not just on a firm's compliance status, but also on a variety of other measures.[3] In the U.S., for example, "compliance monitoring" by environmental regulators typically includes inspection of production and abatement equipment, review of a firm's records, and interviews with its employees (Hunter and Waterman 1996, 39-46). Hawkins, and Nyborg and Telle report similar practices in Britain and Norway. This additional information makes selective enforcement possible in the model.

Consistent with the notion that firms may be noncompliant despite good faith efforts to comply, compliance is modeled as being probabilistic. This reflects the fact that a firm's compliance status often depends on factors that are inherently difficult for it to control. For example, equipment malfunctions or input quality variations may result in output that does not meet product quality or product safety standards. In the case of pollution control regulations, violations can also occur because of treatment process upsets (Russell et al. 1986, Chapter 3). More generally, compliance can be probabilistic in any setting in which firms can make mistakes. Nyborg and Telle point out that mistakes could arise from agency problems within the firm—the level of compliance effort desired by the firm's management may not be exerted by subordinates.

A final feature of the model is that it allows for the possibility that the regulator may only be able to observe the firm's compliance status imperfectly.[4] This is true, for example, when compliance with product quality or product safety standards is determined by examining a sample of a firm's output. In the case of environmental regulations, limitations of pollution measurement techniques can result in a firm's compliance status being observed imperfectly (Russell et al., 1986,

[3] An exception is Malik (2006), who presents a model of environmental regulation that incorporates information on firm behavior other than its observed emissions. The focus of his paper is the design of optimal environmental regulation rather than selective enforcement of regulations.

[4] Harford (1991) and Jost (1997), among others, allow for imperfect observation of compliance.

Chap. 3). More generally, if noncompliance is a fugitive event, establishing that it occurred at a later date is likely to be subject to error. For example, a temporary violation of a workplace safety regulation may be difficult to detect after the fact. Similarly, an accidental discharge of pollution into a water body may be difficult to detect, or attribute, once the pollutant has dispersed.

The basic model of regulatory enforcement is presented in the next section. In Section 3, I then characterize the optimal enforcement policy under the assumption that fines are costless transfers. I show that forgiveness can reduce the probability with which the firm must be monitored, yielding savings in expected inspection costs. I also identify a sufficient condition for forgiveness to be socially desirable. The two subsequent sections consider extensions of the basic model. Section 4 introduces fine costs and shows that these costs render the regulator more inclined to forgive noncompliance. Self-reporting by the firm of its compliance status is introduced in Section 5. I show that forgiveness may be even more likely in the presence of self-reporting. The final section contains some concluding remarks.

2. Model

Compliance with a regulation is modeled as a binary event. The probability the firm is in compliance $p(a)$ depends on the level of compliance effort it undertakes $a \in [0, \overline{a}]$, with $p(a) \in [0, 1)$, $p' > 0$ and $p'' < 0$. Thus, compliance is determined not only by the firm's effort level, but also by other random factors. Higher effort levels increase the probability of compliance, but never ensure it. The firm's effort costs are given by a smooth, strictly increasing and strictly convex function $C(a)$.

The firm is monitored by the regulator with some probability m. Monitoring of the firm is a prerequisite to assessing its compliance status and acquiring any other signals of its compliance effort. The cost of a monitoring visit is given by c_m. Assessing the firm's compliance status during a monitoring visit entails an additional cost c_A. We can think of c_m as capturing the cost of simply visiting the firm, whereas c_A captures, for example, the cost of measuring the firm's pollution emissions. These two types of costs are typically lumped together in models of enforcement. The reason they are separated here will become evident further below. To simplify exposition, I will use the term *inspection cost* to refer to the sum $(c_m + c_A)$.

Assessments of the firm's compliance status may be subject to error. The firm may be perceived to be in noncompliance even though it was actually compliant, and vice versa. Let $\theta_{N|N}$ denote the probability that the regulator correctly concludes that the firm is noncompliant, and let $\theta_{N|C}$ denote the probability that it does so incorrectly. The probabilities $\theta_{C|C}$ and $\theta_{C|N}$ are defined analogously. The accuracy of the regulator's assessment technology is higher the larger is $\theta_{N|N}$ and the smaller is $\theta_{N|C}$. If assessments are perfectly accurate, $\theta_{N|N} = 1$ and $\theta_{N|C} = 0$. The difference $\theta_{N|N} - \theta_{N|C}$ (which equals $\theta_{C|C} - \theta_{C|N}$), is a measure of the accuracy of the assessment technology. As is plausible, I assume the assessment technology has some discriminating power, so the probability of correctly observing the firm's compliance status is larger than the probability of incorrectly observing it. Thus,

$$\alpha \equiv \theta_{N|N} - \theta_{N|C} \in (0, 1] \tag{1}$$

The compliance status perceived by the regulator represents the first signal of the firm's effort level acquired by the regulator. If the firm is perceived to be in compliance, no fine is imposed and no additional signals are acquired the regulator. If the firm is perceived to be in noncompliance, the regulator conducts an "investigation" and acquires another (verifiable) signal of the firm's effort level at a cost c_s. This signal s is drawn from a finite set S. The set S may be multidimensional, and it may contain both quantitative and qualitative information, as is plausible for the outcome of an investigation of firm behavior.[5] Two different types of variables may be contained in the signal s: (1) variables that are a deterministic function of effort level a but can only be observed by the regulator with error, and (2) variables that are inherently randomly related to a. An example of the first type of variable would be the regulator's observation of the degree to which the firm followed procedures consistent with achieving compliance, based on inspections of the firm's records and interviews with its employees. An example of the second type of variable would be the state of equipment needed to achieve compliance, specifically whether or not the equipment was in working order. Equipment break downs are presumably random, with the probability of one occurring depending on the firm's compliance effort.

[5]See Sinclair-Desgagne (1994) for the theory of multi-signal principal agent problems.

The probability of observing a particular signal $s \in S$ is given by $q(s|a)$, which is assumed to be a continuously differentiable function of a, with[6]

$$q(s|a) > 0 \ \forall\{s,a\}\, ; \quad \text{for every } a\, , \ q_a(s|a) \neq 0 \text{ for some } s. \tag{2}$$

The first assumption is a standard one and ensures that the moral hazard problem is not trivial. The second assumption ensures that for every effort level, at least some signals are informative. Since $\{q(s|a), s \in S\}$ is a probability distribution, the following relationships must also hold:

$$\sum_{s \in S} q(s|a) \equiv 1, \quad \sum_{s \in S} q_a(s|a) = 0 \ \ \forall a. \tag{3}$$

The latter equality together with the second assumption in (2) implies that for every effort level a, $q_a(s|a) < 0$ for some signals and $q_a(s|a) > 0$ for others.

The fine imposed on the firm for noncompliance can depend on the signal observed, and is written $f(s)$. Following Harrington, and Heyes and Rickman, among others, there is an exogenously specified maximum fine F. This is typically determined by legislation, the firm's assets, or what courts are likely to consider reasonable. I assume the magnitude of F is such that it does not induce the firm to shut down, yet it is large enough to ensure the existence of a solution to the regulator's problem. Initially, fines are treated as costless transfers, I introduce fine costs in Section 4.

As is common in principal-agent models, I assume the regulator is able to commit to its enforcement policy at the start of the game. Furthermore, the firm's cost function $C(a)$, the probability functions $p(a)$ and $q(s|a)$, and the characteristics of the assessment technology, $\theta_{N|N}$ and $\theta_{N|C}$, are assumed to be common knowledge. The model differs from a typical principal-agent model by the absence of a participation constraint. By assumption, the regulation is mandatory, and the firm cannot exempt itself from being regulated. The participation constraint is replaced by a limited

[6] As implied by the above formulation, the firm's (true) compliance status and the signal s are assumed to be conditionally independent given the firm's effort level. This assumption, that the signals of the agent's effort observed by the principal are conditionally independent, is common in the related literature on optimal auditing, e.g., see Dye (1986). The assumption is relevant only when the regulator's assessment technology is imperfect. When the technology is perfect ($\alpha = 1$), s is acquired by the regulator only when the firm is actually in noncompliance.

liability constraint, which reflects the above-mentioned restrictions on the maximum fine for noncompliance.[7] Inclusion of a participation constraint restricting the magnitude of the firm's total costs (which consist of effort costs plus expected fine costs) would not mitigate the desirability of forgiveness. As shown below, forgiveness invariably lowers the firm's expected fine costs. Therefore, for a given effort level, a participation constraint that is binding when the regulator never engages in forgiveness would not be binding when the regulator engages in forgiveness; similarly, a constraint that is not binding in the first scenario also would not be binding in the second one.

3. Costless Fines

The firm's problem is to choose the effort level that minimizes the sum of its effort costs and expected fine costs:

$$\min_{a \geq 0} \quad C(a) + m[1 - p(a)]\theta_{N|N} \sum_{s \in S} q(s|a)f(s) + mp(a)\theta_{N|C} \sum_{s \in S} q(s|a)f(s). \tag{4}$$

The third term reflects the possibility of the firm being erroneously fined for noncompliance. The term drops out if compliance assessments are perfectly accurate. For $a > 0$, the first-order condition for the firm's problem is

$$C'(a) + m \left\{ [\theta_{N|N} - p(a)\alpha] \sum_{s \in S} q_a(s|a)f(s) - p'(a)\alpha \sum_{s \in S} q(s|a)f(s) \right\} = 0. \tag{5}$$

[7] The absence of a participation constraint and the inclusion of a limited liability constraint are quite common in principal-agent models of regulation. See, for example, Innes (1999) and Malik (1993). This type of model structure is plausible to the extent that regulation of the firm is socially desirable and F is not so large as to induce the firm to shut down.

This condition constitutes the incentive compatibility constraint faced by the regulator.[8] Note that the expression in braces must be negative for the condition to hold. The sign of the first term in braces is determined by the sign of the sum in this term, since $\theta_{N|N} - p(a)\alpha \equiv \theta_{N|N}[1 - p(a)] + \theta_{N|C}p(a) > 0$.

The regulator's problem can be divided into two stages, as in Grossman and Hart (1983). In the first stage, the benevolent regulator minimizes the expected social costs of inducing the firm to select some effort level $a > 0$. In the second stage, the regulator then chooses the effort level that minimizes total expected social costs, which consist of the social costs captured in the first stage objective function plus the social costs associated with noncompliance and the firm's effort costs. Given the focus of this paper, I will restrict attention to the first-stage problem. The second stage simply calls for balancing the expected marginal social cost of inducing higher effort and the marginal benefit of lower expected noncompliance.

The regulator's first-stage problem can be written

$$R(a) \equiv \min_{m \geq 0,\, f(s) \geq 0} \quad m[c_m + c_A] + m\left\{\theta_{N|N}[1 - p(a)] + \theta_{N|C}p(a)\right\} c_s, \tag{6}$$

subject to the incentive compatibility constraint (5), and a limited liability constraint,[9]

$$f(s) \leq F \quad \forall s. \tag{7}$$

The two terms on the RHS of (6) capture expected inspection costs and expected investigation costs, respectively. Note that the objective function and constraints are linear in the decision variables.

[8] The condition is both necessary and sufficient for a unique optimum if the firm's objective function is strictly convex,

$$C''(a) + m\left\{[\theta_{N|N} - p\alpha]\sum_{s \in S} q_{aa}f - 2p'\alpha\sum_{s \in S} q_a f - p''\alpha\sum_{s \in S} qf\right\} > 0.$$

The first term on the LHS and the third term in braces clearly have the correct sign. This is also true of the second term in braces given (3) and the result obtained below that $f(s) = F$ unless $q_a(s|a)/q(s|a)$ is positive and sufficiently large, in which case $f(s) = 0$. The sign of the first term in braces is ambiguous, but the inequality will clearly hold even if the term is negative, provided it is not too large.

[9] In general, a constraint $m \leq 1$ is also required. To avoid clutter, I have omitted this constraint, and assume that the parameters of the problem ensure a solution with $m \in (0, 1]$.

As a result, the Kuhn-Tucker conditions for the problem are both necessary and sufficient.

The structure of the objective function in (6) reflects the assumption that the regulator acquires the signal s only if the firm is perceived to be in noncompliance. Thus, acquisition of s is contingent on the firm's perceived compliance status. One could instead assume that acquisition of s is not made contingent on the firm's perceived compliance status. It is easily verified that the regulator's costs would be higher in this case given the assumption that $\alpha > 0$, i.e., that the probability of correctly detecting noncompliance is higher than the probability of incorrectly detecting it.[10] Thus the regulator is better off making investigations contingent on the firm's perceived compliance status, as one would expect.

3.1 Optimal Fines and Inspection Probability

To put the effects of forgiveness in perspective, let us first consider a setting in which the regulator never conducts investigations and never forgives noncompliance. The regulator simply imposes a fine f whenever noncompliance is perceived. In this setting, the incentive compatibility constraint reduces to

$$C'(a) - mp'(a)\alpha f = 0, \tag{8}$$

and the second term in the regulator's objective function in (6) drops out. For this case, the optimal enforcement policy is easily characterized:

Result 1 *When the regulator never forgives noncompliance, the firm is monitored with probability $\overline{m}^* = C'(a)/p'(a)\alpha F$, and the maximum fine is imposed whenever noncompliance is perceived, $\overline{f}^* = F$.*

This policy reflects the conventional result in the literature on crime and punishment: to set fines at their maximum level if they are costless transfers (Becker 1968). This allows the probability of

[10]For a given value of a, the optimal values of m and $f(s)$ are no different when acquisition of s is not contingent on the firm's perceived compliance status. This stems from the fact that: (i) the incentive compatibility constraint remains the same, and (ii) the regulator's objective function in (6) changes only to the extent that the expression in braces is replaced by one. A comparison of the modified objective function and the original one yields the conclusion that, given $\alpha > 0$, the regulator's costs are lower when acquisition of s is contingent on the firm's compliance status.

the fine (m) to be reduced, lowering expenditures on inspections.

I now turn to the consequences of forgiveness. The primary objective is to demonstrate that for some signals s, setting fines for noncompliance equal to zero, i.e., forgiving noncompliance, can be optimal *given* that an investigation is conducted. Whether or not it is socially desirable to conduct costly investigations is a question that is deferred to Section 3.3.

Given the linearity of the regulator's problem in (5) – (7), its solution can be partly characterized by examining the objective function and constraints. The optimal value of the monitoring probability is determined by the incentive compatibility constraint (5), which can be rewritten

$$m^* = \frac{C'(a)}{-[\theta_{N|N} - p(a)\alpha]\sum_{s\in S} q_a(s|a)f^*(s) + p'(a)\alpha\sum_{s\in S} q(s|a)f^*(s)}, \tag{9}$$

where $f^*(s)$ denotes the optimal values of the fines. The optimal fines can be characterized using the relevant Kuhn-Tucker conditions:

$$\Phi_f \equiv \mu m \left\{[\theta_{N|N} - p(a)\alpha]q_a(s|a) - p'(a)\alpha q(s|a)\right\} + \lambda(s) \geq 0, \quad \Phi_f \cdot f(s) = 0 \quad \forall s, \tag{10}$$

where μ and $\lambda(s)$ are the Lagrange multipliers associated with the the constraints in (5) and (7), respectively. It is not difficult to show that μ must be positive;[11] $\lambda(s)$ is by construction non-negative.

Since $f(s)$ does not appear in (10), the optimal fines will, in general, take on values of either zero or F.[12] If the expression in braces in (10) has a negative sign, the first condition in (10) will hold as an equality with $\lambda(s) > 0$, and the optimal fine will equal F. The expression in braces obviously has a negative sign when $q_a < 0$; it may also have a negative sign when $q_a > 0$. However,

[11] If μ were negative, (10) would imply that $f(s) = F$ when the expression in braces is positive, but if this were true then (5) could not hold because the expression in braces in (5) would be positive.

[12] Depending on the exogenously given values of $\theta_{N|N}$, α, $q(s|a)$, $q_a(s|a)$, $p(a)$ and $p'(a)$, the expression in braces in (10) could take on a value of zero for some s. In such cases, the magnitude of the fine is irrelevant, for ease of exposition, I assume it is set equal to F.

it has a positive sign when $q_a > 0$ and

$$L_S(s,a) \equiv \frac{q_a\,(s|a)}{q(s|a)} > \frac{p'(a)\alpha}{\theta_{N|N} - p(a)\alpha} \equiv -L_N(a,\alpha). \tag{11}$$

If the above inequality holds for some signal s, (10) implies it is optimal to set $f(s) = 0$ and forgive noncompliance. Let $S_0 \subset S$ denote the set of signals for which (11) holds.

The LHS of the inequality in (11) is simply the likelihood ratio associated with the signal s. The RHS is the negative of the likelihood ratio associated with the signal "noncompliance perceived," which I denote by N: $-L_N(a,\alpha) \equiv -dln\{[1-p(a)]\theta_{N|N} + p(a)\theta_{N|C}\}/da$ The magnitude of this likelihood ratio is positively related to the accuracy of the regulator's assessment technology, with

$$-L_N(a,\alpha) \Rightarrow 0 \text{ as } \alpha \Rightarrow 0, \qquad -L_N(a,\alpha) \Rightarrow \frac{p'(a)}{1-p(a)} \text{ as } \alpha \Rightarrow 1. \tag{12}$$

$-L_N(a,\alpha)$ is a measure of the evidence conveyed by perceived noncompliance of the firm having chosen a *low* effort level. Analogously, $L_S(s,a)$, the LHS of (11), is a measure of the strength of evidence conveyed by the signal s that the firm had chosen a *high* effort level. Thus, the inequality in (11) indicates that the regulator will choose to forgive noncompliance if, and only if, the evidence conveyed by the signal s of the firm having a chosen a high effort level is stronger than the evidence conveyed by perceived noncompliance of the firm having a chosen a low effort level. This finding is consistent with Holmstrom's (1979) result that additional signals of an agent's effort choice are valuable only if they alter the principal's *ex post* assessment of the agent's effort choice.

Since $-L_N(a,\alpha)$ approaches zero as the accuracy of the assessment technology decreases, the inequality in (11) will invariably hold for some s if the accuracy of the assessment technology is sufficiently low. In the limiting case where $\alpha = 0$, perceived noncompliance is ignored when determining the magnitude of the fine. The fine is based solely on the sign of $L_S(s,a)$, with a fine imposed if, and only if, $L_S(s,a) < 0$. These observations imply that the cardinality of S_0, $n(S_0)$, will be (weakly) larger the lower is the accuracy of the assessment technology.

It might seem implausible for α to take on values near zero, and for the regulator to virtually ignore direct assessment of a firm's compliance status when determining the magnitude of a fine.

However, this could well occur in some circumstances, especially when noncompliance is a fugitive event.[13]

The set of signals for which (11) holds, and forgiveness is optimal, also depends on the desired level of effort a, though not in an easily predicted manner. It is readily verified that the LHS of (11) is decreasing in a given the plausible assumption that $q_{aa} < 0$ when $q_a > 0$. However, the sign of the derivative of the RHS with respect to a is ambiguous given the corresponding assumption that $p'' < 0$.

It should be emphasized that inaccurate assessments are not necessary for forgiveness to be optimal. The inequality in (11) can hold when assessments are perfectly accurate ($\alpha = 1$). Whether or not it holds in this case depends on the relative magnitude of $q_a(s|a)/q(s|a)$ and $p'(a)/[1-p(a)]$. It is only if their relative magnitude is such that (11) does not hold for any signal s that $f(s) = F$ $\forall s$, and the optimal enforcement policy is no different than that described in Result 1.[14]

Therefore, we have:

Proposition 1 *The regulator will choose to forgive noncompliance for some signals s if they constitute sufficiently strong evidence of the firm having chosen a high effort level. The set of signals for which noncompliance is forgiven, S_0, is (weakly) larger, the lower is the accuracy of the regulator's assessment technology, α. For signals $s \notin S_0$, it is optimal to impose the maximum fine F for perceived noncompliance.*

The set of signals for which noncompliance is forgiven can be characterized, at least partially, if some structure is imposed on the set of possible signals S. Consider the simplest case where S is a vector of values ordered from smallest to largest. If we make the common assumption that $L_S(s, a)$ satisfies the monotone likelihood ratio property, so $L_S(s, a)$ is increasing in s, noncompliance will be forgiven (if it is at all) for large values of s, since they constitute stronger evidence of the firm

---

[13] Nyborg and Telle describe another possibility, drawn from Norwegian regulatory practice. Some regulations limit the amount of pollution a firm can emit over a period of a year. In such cases, the level of pollution observed during a monitoring visit may tell the regulator little about the firm's compliance with the annual limit. As a result, the regulator is more inclined to make inferences about the firm's compliance effort on the basis of other information collected during a monitoring visit.

[14] For this case, given (3), the denominator in (9) reduces to $p'(a)\alpha F$.

having chosen a high effort level.[15]

The structure of the optimal enforcement policy in Proposition 1 implies that the signals in S can be aggregated to yield a binary statistic that is mechanism sufficient (Demougin and Fluet 1998); that is, the binary statistic can be used to induce the firm to choose the desired effort level a at the same cost as that incurred when the signal s is observed. The relevant binary statistic is

$$Y = \begin{cases} 0 & if \quad s \in S_0 \\ 1 & if \quad s \in \overline{S_0} \end{cases} \tag{13}$$

where $\overline{S_0} = S \backslash S_0$ is the set of signals for which the maximum fine is imposed. The probability that $Y = 0$ is captured by $\sigma(a)$, where

$$\sigma(a) \equiv \sum_{s \in S_0} q(s|a). \tag{14}$$

By forgiving noncompliance when $Y = 0$ and imposing the maximum fine when $Y = 1$, the regulator can duplicate the outcome obtained when the signal s is observed. The usefulness of this binary statistic will become evident shortly. Note that its likelihood ratio is given by $\sigma_a(a)/\sigma(a)$.

3.2 The Benefit of Forgiveness

Given the form of the regulator's objective function in (6), the benefit of forgiving noncompliance must be a reduction in the monitoring probability needed to satisfy the incentive compatibility constraint, reducing expected inspection costs, $m[c_m + c_A]$. This seems contrary to the conventional wisdom that inspection costs are reduced by setting higher fines. The seeming contradiction stems from the fact that, for a given monitoring probability, the additional information provided by the signal s combined with selective forgiveness, enables the regulator to increase the marginal benefit to the firm of choosing a higher effort level, as captured by the expression in braces in (5). In other words, the regulator is able to increase the power of the penalty scheme, allowing the regulator to

[15] A similar result can be obtained if s is a multidimensional signal and S is a lattice that has a component-wise partial ordering. If $L_S(s, a)$ satisfies the monotone likelihood ratio property (see Sinclair-Desgagne 1994), then forgiveness is associated with large values of the vector s.

satisfy the incentive compatibility constraint with a lower monitoring probability.

The reduction in the monitoring probability is characterized in

Result 2 *When forgiveness is optimal, it reduces the monitoring probability needed to satisfy the incentive compatibility constraint. The magnitude of this reduction is: (i) negatively related to the accuracy of the assessment technology,* α*; and (ii) positively related to the difference in likelihood ratios* $\Delta^{\sigma} \equiv \sigma_a(a)/\sigma(a) - [-L_N(a, \alpha)]$*, except when an increase in* Δ^{σ} *is accompanied by a proportionally larger decrease in the value of* $\sigma(a)[\theta_{N|N} - p(a)\alpha]$*.*

The proof is provided in the Appendix. Result 2 implies that, in general, the benefit of forgiveness is larger the less accurate is the regulator's assessment technology, and the larger is the difference in likelihood ratios $\Delta^{\sigma} \equiv \sigma_a(a)/\sigma(a) - [-L_N(a, \alpha)]$.

The reduction in the monitoring probability, together with the fact that noncompliance is not always punished, yields

Corollary 1 *The firm's expected costs are lower when the regulator engages in forgiveness.*

Thus, as we would expect, the firm is better off when the regulator selectively forgives noncompliance.

3.3 When are Investigations Desirable?

The above analysis demonstrates that investigations coupled with forgiveness can be socially desirable, because of their potential to reduce expected inspection costs. But investigations are costly to conduct. They will reduce the social costs of achieving a desired effort level only if the expected costs of conducting investigations are offset by the reduction in expected inspection costs. I now derive a simple sufficient condition under which this offset occurs.

Making use of (6) and Result 1, investigations coupled with forgiveness are desirable if, and only if, $R(a) < \overline{R}(a)$, where $\overline{R}(a) \equiv c\overline{m}^*$ denotes the regulator's indirect objective function when investigations are never conducted. Now consider the regulator's objective function when investigations are conducted, but noncompliance is forgiven only for the realization of s that yields the largest value of the likelihood ratio $L_S(s,a) \equiv q_a(s|a)/q(s|a)$; thus, for this realization, labeled s_M, (11) must hold. Let $R^M(a)$ denote the regulator's objective function for this case. Clearly, $R^M(a) \geq R(a)$, since it may be optimal to forgive noncompliance for other realizations of s. Therefore, $R^M(a) < \overline{R}(a)$ is a sufficient condition for investigations to be desirable.

Some algebra establishes that $R^M(a) < \overline{R}(a)$ if, and only if,

$$\frac{c_s}{c_m + c_A} < \frac{q_a(s_M|a)}{p'(a)\alpha} - \frac{q(s_M|a)}{\theta_{N|N} - p(a)\alpha}. \tag{15}$$

Let K denote the critical value of the cost ratio represented by the RHS of (15). Note that K does not depend on $C'(a)$ or F. Furthermore, from (11), K must be positive given the premise that it is optimal to forgive noncompliance when s_M is observed. The value of K will depend on the values of $\theta_{N|N}$ and $\theta_{N|C}$ (recall $\alpha \equiv \theta_{N|N} - \theta_{N|C}$). As shown in the Appendix, $\partial K/\partial\theta_{N|N} < 0$. This implies, roughly speaking, that investigations are more likely to be desirable, the lower is the regulator's ability to correctly detect noncompliance. Correspondingly, $\partial K/\partial\theta_{N|C} > 0$, i.e., investigations are more likely to be desirable the lower is the regulator's ability to correctly detect compliance. Comparing (11), with $s = s_M$, to the RHS of (15), we can see that the value of K is also related to the difference in likelihood ratios, $\Delta^S \equiv L_S(s_M,a) - [-L_N(a,\alpha)]$. In fact, we can write $K \equiv [q(s_M|a)/p'(a)\alpha]\,\Delta$. This expression reveals that an increase in Δ^S would result in a larger value of K, except in the case where the increase in Δ^S is accompanied by a proportionally larger reduction in the value of $q(s_M|a)/p'(a)\alpha$, a possibility that cannot be ruled out.

These results are summarized more formally as follows:

Proposition 2 *Investigations, together with forgiveness, reduce social costs if the ratio $c_S/(c_m{+}c_A)$ falls below some critical value K. This critical value does not depend on the marginal cost of effort ($C'(a)$) or the maximum fine (F). It is a decreasing function of the accuracy of the regulator's assessment technology, α; and it is an increasing function of the difference in likelihood ratios $\Delta^S \equiv L_S(s_M,a) - [-L_N(a,\alpha)]$ except when the increase in Δ^S is accompanied by a proportionally larger decrease in $q(s_M|a)/p'(a)\alpha$.*

We can get some sense of the magnitude of K by calculating its value for hypothetical values of the underlying probabilities. The results of this exercise are presented in Table 1. For the initial, and baseline, calculation, I assume the regulator's assessment technology is perfect, so $\alpha = \theta_{N|N} = 1$. Let the probability of the firm being in compliance, $p(a) = 0.8$. Let the probability of s_M being observed, $q(s_M|a) = 0.2$, as is not unreasonable if the set of possible signals S has multiple elements, with s_M being the signal that conveys the strongest evidence of the firm having chosen a high effort level. Note, from (15), that K depends on the ratio $q_a(s_M|a)/p'(a)\alpha$. Let this ratio equal 4, which implies, given the other values assumed, that $L_S(s_M, a) = 4 \cdot [-L_N(a, \alpha)]$, ensuring that forgiveness is optimal when s_M is observed. For these values, $K = 3$, as shown in the first row of the table, i.e., investigations are desirable if the cost of conducting an investigation is no more than three times as high as the cost of conducting an inspection. The next three rows show the effects of changes in the ratio $q_a(s_M|a)/p'(a)$. When $q_a(s_M|a)/p'(a) = 1$, $K = 0$, which implies that investigations are not desirable, since $c_S/(c_m + c_A)$ cannot be negative. In fact, forgiveness is not optimal for this case (and this case alone) since $L_S(s_M, a) = -L_N(a, \alpha)$.

The second set of results show the effect of varying the probability of compliance, while holding other probabilities at their baseline values, and the third set shows the effect of varying the probability of observing s_M (which is also the probability of forgiveness). These variations result in values of K that range between 2 and 3.5. The last two sets of results show the effects of introducing modest errors in the assessment technology. K then takes on values between 3.3 and 6.0.

In practice, the relative magnitude of investigation costs and inspection costs ($c_S/(c_m + c_A)$) is likely to vary widely across regulatory settings. In some cases, investigation costs may well be lower than inspection costs, whereas in others they may be much higher. The results in Table 1 suggest that investigations can be socially desirable even when the cost of conducting an investigation is several times higher than the cost of conducting an inspection. This is particularly true when the regulator's assessment technology is imperfect.

4. Costly Fines

The model presented in the previous section embodies the common assumption that fines are costless transfers. But there are invariably administrative costs associated with collecting fines.

In addition, fines may be subject to costly administrative or judicial reviews. Accordingly, in this section I incorporate fine costs in the model. Specifically, I allow for a constant unit (or marginal) fine cost, c_f. It may be more plausible to assume that fine costs are fixed, and invariant with the magnitude of the fine. However, fixed fine costs make the analysis much more tedious, and yield results that are qualitatively identical to those with a constant unit fine cost.

I begin, once again, by considering the benchmark case in which investigations are never conducted and noncompliance is never forgiven. The regulator's (first-stage) objective function now has an additional term, capturing expected fine costs: $m[c_m + c_A] + m\left\{[1 - p(a)]\theta_{N|N} + p(a)\theta_{N|C}\right\} c_f f$. The constraints of the regulator's problem are unchanged. In particular, the incentive compatibility constraint is still given by (8). Intuition suggests that the presence of fine costs should alter the optimal enforcement policy and lower the optimal fine. This is not true, however. Fine costs have no effect on the optimal enforcement policy, regardless of their magnitude:

Result 3 *When the regulator never forgives noncompliance, the presence of fine costs does not alter the optimal enforcement policy: it is still optimal to impose the maximum fine F whenever noncompliance is perceived, and to monitor the firm with probability $\overline{m}^* = C'(a)/p'(a)\alpha F$.*

This result can be explained as follows. Suppose the fine is not set equal to its maximum value F. The regulator could then raise the fine to F and lower the monitoring probability so that the expected fine for noncompliance, mf, is unchanged. These changes would not affect the incentive compatibility constraint nor would they alter the magnitude of the second term in the regulator's objective function, since both are functions of mf. However, they would lower expected inspection costs (the first term in the regulator's objective function), lowering the total costs faced by the regulator. This result holds *a fortiori* if fine costs are fixed and do not vary with the magnitude of the fine.

I now turn to the model with forgiveness. Once again, fine costs result in an additional term in the regulator's (first-stage) objective function,

$$\min_{m\geq 0, f(s)\geq 0} \quad m[c_m + c_A] + m\left\{\theta_{N|N}[1-p(a)] + \theta_{N|C}p(a)\right\}c_S$$
$$+ m\left\{[1-p(a)]\theta_{N|N}\sum_{s\in S} q(s|a)f(s) + p(a)\theta_{N|C}\sum_{s\in S} q(s|a)f(s)\right\}c_f, \tag{16}$$

but leave the constraints of the problem, (5) and (7), unchanged. Accordingly, the optimal inspection probability is still given by the expression in (9), however, its magnitude will differ to the extent that the optimal fine scheme is altered. The Kuhn-Tucker conditions for the optimal fines can be written:

$$\Phi_f \equiv c_f + \mu\left\{\frac{q_a(s|a)}{q(s|a)} - \frac{p'(a)\alpha}{\theta_{N|N} - p(a)\alpha}\right\} + \frac{\lambda(s)}{m[\theta_{N|N} - p(a)\alpha]q(s|a)} \geq 0, \quad \Phi_f \cdot f(s) = 0 \ \ \forall s, \tag{17}$$

where, as before, μ is the Lagrange multiplier for the incentive compatibility constraint and $\lambda(s) \geq 0$ is the multiplier for the limited liability constraint. It can be verified quite easily that μ must still be strictly positive.

From (17), we can see that forgiveness now requires

$$\frac{c_f}{\mu} + \frac{q_a(s|a)}{q(s|a)} > \frac{p'(a)\alpha}{\theta_{N|N} - p(a)\alpha}. \tag{18}$$

The presence of the fine cost term implies that this condition is less stringent than the earlier condition in (11). Therefore, (18) invariably holds for all signals that satisfy (11), $s \in S_0$, and it may hold for some $s \notin S_0$. Thus, the set of signals for which noncompliance is forgiven is weakly larger when fines are costly: $n(S_0^f) \geq n(S_0)$ with $S_0^f \supset S_0$, where S_0^f denotes the set of signals that satisfy (18).

This result is consistent with common intuition. We would expect the regulator to be more inclined to forgive noncompliance when fines are costly. Because forgiveness may now be motivated by the desire to reduce fine costs, the optimal monitoring probability may now be larger than when fines are costless, $m_f^* \geq m^*$, to compensate for the reduced frequency with which fines are imposed.

This can be established by inspecting the denominator of the expression in (9) for the optimal monitoring probability. Suppose there is some signal s' for which it is optimal to impose a fine $f(s') > 0$ when fines are costless but not when they are costly. Then the expression for m_f^* differs from that for m^* by the *absence* of the term

$$\left\{-[\theta_{N|N} - p(a)\alpha]q_a(s'|a) + p'(a)\alpha q(s'|a)\right\} f(s') \tag{19}$$

in the denominator of m_f^*. From (11), we can determine that this expression must have a positive sign given the premise that $f(s') > 0$ when fines are costless. Thus, we can conclude that if $n(S_0^f) > n(S_0)$, m_f^* must be larger than m^*; however, if $S_0^f = S_0$, $m_f^* = m^*$.[16]

These results are summarized as follows:

Proposition 3 *When fines are costly to impose the set of signals for which the regulator forgives noncompliance, S_0^f, is (weakly) larger than when fines are costless, $S_0^f \supset S_0$. In addition, the optimal monitoring probability may be larger, $m_f^* \geq m^*$. For signals $s \notin S_0^f$, it is optimal to impose the maximum fine F for perceived noncompliance.*

Because forgiveness may now be motivated by the desire to reduce fine costs rather than inspection costs, it is no longer possible to state unequivocally that forgiveness lowers the optimal monitoring probability ($m_f^* < \overline{m}^*$). As a result, we cannot rule out the possibility that the regulator's expected inspection costs are now higher when it engages in forgiveness. The regulator's expected fine costs must be lower, however, as can be verified by examining the regulator's objective function in (16).[17]

A comparison of (16) and (4) shows that the regulator's expected fine costs are linearly related to the expected fine faced by the firm. The above observation therefore implies that the firm's

[16] Qualitatively identical conclusions hold if fine costs are fixed. For each value of s for which $f(s) > 0$ when fines are costless, the regulator would now have to compare the benefit of a positive fine, as captured by the reduction in expected inspection costs due to a lower monitoring probability, to the fixed fine cost. Positive fines would be retained only for those values of s for which this benefit exceeds the fixed fine cost.

[17] If the optimal monitoring probability is higher when the regulator engages in forgiveness, then forgiveness must reduce the expected fine costs incurred by the regulator (captured by the third term in (16)). There would, otherwise, be no reason to engage in forgiveness. If the optimal monitoring probability is lower, then expected fine costs must be lower, since forgiveness implies that perceived noncompliance is not always punished.

expected costs continue to be lower when the regulator engages in forgiveness. Thus, we have:

Corollary 2 *When fines are costly, forgiveness may not result in a lower monitoring probability. However, it invariably lowers the regulator's expected fine costs, and the expected costs faced by the firm.*

5. Self-Reporting

A common feature of regulations is the requirement that firms self-report their compliance status (Malik 1993, Kaplow and Shavell 1994, Innes 1999). In this section, I incorporate self-reporting in the model and show that forgiveness can be optimal in this setting as well. Forgiveness can now occur in two scenarios: when the firm reports noncompliance and when it is perceived to have dishonestly reported compliance. For the sake of brevity, I only allow for forgiveness in the first scenario. Although forgiveness may be desirable in the second scenario, it hinges on $\theta_{N|C}$, the probability of the regulator incorrectly concluding that the firm is noncompliant, being sufficiently large. To identify the effects of self-reporting more easily, I ignore fine costs.

The firm's compliance status is now assessed only if it reports itself in compliance. Let m_C denote the probability that it is monitored following such a report and an assessment of its compliance status conducted. If the assessment indicates the firm is in noncompliance, it is fined an amount f_{NC}, where NC denotes the event "noncompliance perceived when compliance reported."

If the firm reports noncompliance, there is presumably no need for an assessment of the firm's compliance status. However, the regulator does conduct an investigation and acquires the signal s, so as to be able to selectively forgive noncompliance. In order to conduct an investigation, the regulator must visit the firm, incurring the monitoring cost c_m, in addition to the investigation cost c_S.

For the firm to report its compliance status truthfully, the expected fine it faces when honest must be lower than the expected fine it faces when dishonest.[18] When the firm is in noncompliance,

[18] The revelation principle allows us to restrict attention to incentive compatible policies that induce truthful reporting. I assume, as usual, that if the firm is indifferent between being honest and dishonest, it chooses to be honest.

truthful reporting requires

$$\sum_{s\in S} q(s|a)f_N(s) - m_C\theta_{N|N}\, f_{NC} \leq 0, \tag{20}$$

where $f_N(s)$ denotes the fine imposed when the firm self-reports noncompliance. When the firm is in compliance, truthful reporting requires

$$m_C\theta_{N|C}f_{NC} - \sum_{s\in S} q(s|a)\, f_N(s) \leq 0. \tag{21}$$

The first term in (21) reflects the possibility of the firm being erroneously fined for noncompliance when it reports compliance. Intuition suggests that (21) should not be binding at an optimum. This is, in fact, correct. Comparing (21) and (20), we can see that if (20) is binding, then (21) will invariably hold as a strict inequality, since $\theta_{N|C} < \theta_{N|N}$ (see (1)). To simplify the exposition, I will proceed by solving a relaxed problem in which (21) is ignored and show that (20) is binding at the solution to this relaxed problem.

Assuming the truth-telling constraint in (20) holds, the firm always reports noncompliance honestly, and its problem takes the form[19]

$$\min_{a\geq 0} \; C(a) + [1-p(a)]\sum_{s\in S} q(s|a)f_N(s) + p(a)m_C\theta_{N|C}f_{NC}. \tag{22}$$

Given $a > 0$, the first-order condition for the firm's problem, and the incentive compatibility

[19]The firm's objective function reflects a typical paradox of principal-agent models: the regulator can mistakenly punish the firm for dishonestly reporting compliance even though it "knows," given the structure of the game, that the firm must have submitted an honest report. Ignoring perceived dishonest reporting would, of course, eliminate the firm's incentive to report truthfully.

constraint faced by the regulator, is[20]

$$C'(a) + [1 - p(a)] \sum_{s \in S} q_a(s|a) f_N(s) - p'(a) \sum_{s \in S} q(s|a) f_N(s) + p'(a) m_C \theta_{N|C} f_{NC} = 0. \qquad (23)$$

Absent fine costs, the regulator's problem is

$$\min_{m_C \geq 0,\, f_N(s) \geq 0,\, f_{NC} \geq 0} \quad p(a) m_C [c_m + c_A] + [1 - p(a)][c_m + c_S], \qquad (24)$$

subject to (23), the truth-telling constraint (20), and two limited liability constraints,[21]

$$f_N(s) \leq F \quad \forall s; \qquad (25)$$

$$f_{NC} \leq F. \qquad (26)$$

The first term in the objective function captures expected inspection costs, and the second, expected investigation costs. The former are incurred only when the firm reports compliance and the latter only when it reports noncompliance. Note that this problem is also linear in the decision variables.

As before, I first characterize the optimal enforcement policy when the regulator does not conduct investigations and never forgives noncompliance. The truth-telling constraint (20) then simplifies to $f_N - m_C \theta_{N|N} f_{NC} \leq 0$, and the incentive compatibility constraint (23) reduces to $C'(a) - p'(a) f_N + p'(a) m_C \theta_{N|C} f_{NC} = 0$, where f_N is a constant fine. Since the regulator does not conduct investigations, the second term in the regulator's objective function (24) drops out.

[20] The condition is both necessary and sufficient for a unique optimum if the firm's objective function is strictly convex,

$$C'' + [1 - p] \sum_{s \in S} q_{aa} f_N - 2p' \sum_{s \in S} q_a f_N + p'' \left\{ m_C \theta_{N|C} f_{NC} - \sum_{s \in S} q f_N \right\} > 0.$$

The first term clearly has the correct sign, as does the third term using an argument similar to that in footnote 8. The last term also has the correct sign, because, as argued above, (21) invariably holds as a strict inequality at the solution to the regulator's problem. The sign of the second term is ambiguous. The inequality will hold even if this term is negative, provided it is not too large.

[21] Once again, to reduce clutter, I omit the constraint for the monitoring probability and assume that the parameters of the problem ensure $m_C \in (0, 1]$.

Following Malik (1993), it can be determined quite easily that the optimal enforcement policy takes the following form:

Result 4 (Malik 1993) *When the firm is required to self-report its compliance status and the regulator never forgives noncompliance, the firm is fined an amount* $\overline{f}_N^* < F$ *when it reports noncompliance. When it reports compliance and is perceived to have done so falsely, it is fined an amount* $\overline{f}_{NC}^* = F$.

The fine for reported noncompliance, $\overline{f}_N^*$ must be set below F, because the truth telling constraint would not hold otherwise. The precise value of $\overline{f}_N^*$ is determined by the incentive compatibility constraint. Setting $\overline{f}_{NC}^* = F$ is optimal since it lowers the monitoring probability needed to satisfy the truth-telling constraint.

Let us now characterize the optimal enforcement policy with forgiveness. It is easily verified that setting $f_{NC} = F$ is still optimal. The optimal values of the fines for reported noncompliance, $f_N(s)$, are determined by the Kuhn-Tucker conditions:

$$\Phi_{f_N}^{SR} \equiv \frac{\tau}{1-p(a)} + \mu \left\{ \frac{q_a(s|a)}{q(s|a)} - \frac{p'(a)}{1-p(a)} \right\} + \frac{\lambda_N(s)}{[1-p(a)]q(s|a)} \geq 0, \quad \Phi_{f_N}^{SR} \cdot f_N(s) = 0 \quad \forall s, \tag{27}$$

where $\tau \geq 0$ is the multiplier associated with the truth-telling constraint in (20), $\lambda_N(s) \geq 0$ is the multiplier associated with the limited liability constraint in (25), and μ is the multiplier associated with the incentive compatibility constraint, (23). I show in the Appendix that $\mu > 0$, as before, and that $\tau > 0$, so the truth-telling constraint is binding (as claimed earlier).

From (27), we can see that reported noncompliance is forgiven when the signal s is observed if, and only if,

$$\frac{\tau}{\mu[1-p(a)]} + \frac{q_a(s|a)}{q(s|a)} > \frac{p'(a)}{1-p(a)}. \tag{28}$$

Comparing this condition to (11), the condition for forgiveness to be optimal in the absence of self-reporting, we see that the LHS of (28) is larger given the presence of the additional term associated with the truth-telling constraint. The RHS of (28) is also larger than the RHS of (11) if $\theta_{N|C} > 0$; but if $\theta_{N|C} = 0$, it is of the same magnitude. Therefore, in general, we cannot specify whether reported noncompliance is more, or less, likely to be forgiven than perceived noncompliance absent self-reporting. But if the assessment technology is such that the firm is never mistakenly found to be in noncompliance ($\theta_{N|C} = 0$), then reported noncompliance is more likely to be forgiven, since (28) is then less stringent than (11). By continuity, this conclusion must also hold for values of $\theta_{N|C}$ sufficiently close to zero. Note that this requirement on the value of $\theta_{N|C}$ is weaker than the requirement that the assessment technology be perfect ($\alpha = 1$), or nearly so. This would additionally require that $\theta_{N|N}$ equal one or be sufficiently close to one.

Letting S_0^{SR} denote the set of signals for which reported noncompliance is forgiven, we can therefore write:

Proposition 4 *Forgiveness may be optimal when the firm self-reports noncompliance. When* $\theta_{N|C} = 0$ *or is sufficiently close to zero, the set of signals for which self-reported noncompliance is forgiven,* S_0^{SR}*, is (weakly) larger than the set of signals for which noncompliance is forgiven in the absence of self-reporting,* $S_0^{SR} \supset S_0$*.*

An examination of the regulator's objective function in (24) indicates that the benefit of forgiveness must be a reduction in the monitoring probability m_C. There is no other avenue by which forgiveness can be beneficial to the regulator. This implies that the expected fine for falsely reporting compliance, $m_C \theta_{N|N} f_{NC}$, which is the second term in the truth-telling constraint (20), must be smaller when the regulator engages in forgiveness (recall that setting $f_{NC} = F$ is optimal whether or not the regulator engages in forgiveness). Since (20) is always binding, this, in turn, implies that the first term in (20), $\sum q(s|a) f_N(s)$, the expected fine for reported noncompliance, must also be lower when the regulator engages in forgiveness. These observations, together with the form of the firm's objective function in (22), yield

Corollary 3 *When the firm is required to self-report its compliance status, forgiveness lowers the optimal monitoring probability and the firm's expected costs.*

Thus, as before, the firm is better off when the regulator conducts investigations and selectively forgives noncompliance.

6. Concluding Remarks

By extending a simple deterrence-based model of regulation to incorporate an additional, possibly multi-dimensional, signal of the regulated firm's compliance effort, I have shown that selectively forgiving noncompliance can be socially desirable. The additional signal of effort corresponds to information commonly collected by regulatory authorities on firms' compliance activities. The model thus explains selective forgiveness of noncompliance and the collection of information on a firm's compliance activities and not just its compliance status.

If fines are costless transfers, selective forgiveness allows the regulator to reduce the probability with which it needs to conduct costly inspections of the firm. If fines are costly, another benefit of selective forgiveness is a reduction in the fine costs incurred by the regulator. The benefits of selective forgiveness must be weighed against the costs of conducting the investigations that yield the additional signal of the firm's effort. In the basic model examined, with costless fines and no self-reporting of compliance status, I identify a simple sufficient condition under which investigations are socially desirable. The condition indicates that investigations are desirable under a range of circumstances, especially if the regulator is only able to assess the firm's compliance status imperfectly. From the firm's perspective, forgiveness is always desirable, since it reduces the firm's expected costs.

One of the assumptions underlying the analysis is that the regulator is benevolent and minimizes social costs. The results obtained would be much the same for a budget-constrained regulator that maximizes the probability of compliance or minimizes the expected social losses from noncompliance (as in Nyborg and Telle, and Heyes and Rickman). Regulators with either of these objectives would want the firm to choose the highest effort level possible, given the regulatory budget constraint. Hence, the regulator would still want to minimize the regulatory costs of achieving a given effort level, as is true of the benevolent regulator. Thus, regulators with these objectives would view forgiveness no less favorably than a benevolent regulator. The same would not be true, however, of a regulator that valued expected fine revenues, so that some weighted value of these revenues were subtracted from its objective function (as in Grieson and Singh 1990, for example). Such a regulator would view forgiveness less favorably, because, as we have seen, forgiveness reduces the expected fines paid by the firm. The degree to which such a regulator would engage in forgiveness would depend on the precise form of its objective function and the weight attached to fine revenues.

Another assumption underlying the analysis is that acquisition of the signal s is contingent on the firm's perceived compliance status. But s, or some component of it, could be acquired first, and assessment of the firm's compliance status could be conditioned on its realization. This would be desirable when assessing the firm's compliance status is relatively costly, as is true for some environmental regulations (e.g., see Russell et al. 1986, Chapter 2; Hunter and Waterman 1996, 39-46). In such cases, the regulator could first acquire a less costly signal of the firm's compliance effort (e.g., by inspecting the condition of the firm's physical plant), and then assess the firm's compliance status only if this signal suggested that the firm had chosen a lower-than-desired effort level. Russell et al. (Chapter 2) and Nyborg and Telle report that this is not uncommon in practice. This sequence of events could be accommodated by modifying the model so that the signal s is acquired in the first stage and the firm's compliance status is assessed in the second one. In such a model, scope for forgiveness would be eliminated: if the realization of s induced the regulator to assess the firm's compliance status, the regulator would invariably punish the firm if noncompliance were perceived—there would otherwise be no reason to incur the assessment cost c_A. Scope for forgiveness could be restored by adding a third stage and separating out the components of s into those that can be acquired at low cost (relative to c_A) and those that can only be acquired at high cost. Let s^L and s^H denote these two types of signals. The regulator would begin by acquiring s^L and conditioning the decision to assess the firm's compliance status on its realization. If an assessment is conducted and indicated noncompliance, the regulator would then acquire the signal s^H. Fines would be contingent on the realization of s^H in the same way that fines are contingent on the realization of s in the model presented here.

References

[1] Ayres, Ian and John Braithwaite. 1992. *Responsive Regulation: Transcending the Deregulation Debate*. New York: Oxford University Press.

[2] Bardach, Eugene, and Robert A. Kagan. 1982. *Going by the Book: The Problem of Regulatory Unreasonableness*. Philadelphia: Temple University Press.

[3] Becker, Gary S. 1968. "Crime and Punishment: An Economic Approach." *Journal of Political Economy* 76:169-217.

[4] Code of Federal Regulations, Title 40, Part 123.27. Protection of Environment. 2004. Washington, D.C: U.S. Government Printing Office.

[5] Demougin, Dominique and Claude Fluet. 1998. "Mechanism Sufficient Statistic in the Risk-Neutral Agency Problem." *Journal of Institutional and Theoretical Economics* 154(4):622-639.

[6] Dye, Ronald A. 1986. "Optimal Monitoring Policies in Agencies." *Rand Journal of Economics* 17: 339-50.

[7] Garvie, Devon and Andrew Keeler. 1994. "Incomplete Enforcement with Endogenous Regulatory Choice." *Journal of Public Economics* 55: 141–162.

[8] Gormley, Jr., William T. 1998. "Regulatory Enforcement Styles." *Political Research Quarterly* 51(2): 363-383.

[9] Grieson, Ronald D., and Nirvakar Singh. 1990. "Regulating Externalities through Testing." *Journal of Public Economics* 41(3):369-87.

[10] Grossman, Sanford D., and Oliver D. Hart. 1983. "An Analysis of the Principal-Agent Problem." *Econometrica* 51:7-45.

[11] Harford, Jon D. 1991. "Measurement Error and State-dependent Pollution Control Enforcement." *Journal of Environmental Economics and Management* 21: 67–81.

[12] Harrington, Winston. 1988. "Enforcement Leverage when Penalties are Restricted." *Journal of Public Economics* 37:29-53.

[13] Harrison, Kathryn. 1995. "Is Cooperation the Answer? Canadian Environmental Enforcement in Comparative Context." *Journal of Policy Analysis and Management* 14(2): 221-244.

[14] Hawkins, Keith. 1984. *Environment and Enforcement: Regulation and the Social Definition of Pollution*. Oxford: Oxford University Press.

[15] Heyes, Anthony G. and N Rickman, 1999. "A Theory of Regulatory Dealing - Revisiting the Harrington Paradox." *Journal of Public Economics* 72(3): 361-78.

[16] Holmstrom, Bengt. 1979. "Moral Hazard and Observability." *Bell Journal of Economics* 10(1): 74-91.

[17] Hunter, Susan and Richard W. Waterman. 1996. *Enforcing the Law: The Case of the Clean Water Acts*. Armonk, NY: M.E. Sharpe.

[18] Innes, Robert. 1999. "Remediation and Self-reporting in Optimal Law Enforcement." *Journal of Public Economics* 72:379-393.

[19] Jost, Peter-J. 1997. "Monitoring, Appeal, and Investigation: The Enforcement and Legal Process." *Journal of Regulatory Economics* 12:127-146.

[20] Kaplow, Louis and Steven Shavell. 1994. "Optimal Law Enforcement with Self-Reporting of Behavior." *Journal of Political Economy* 102(3): 583-606.

[21] Kitamura, Yoshinobu. 2000. "Regulatory Enforcement in Local Government in Japan." *Law and Policy* 22(3/4): 305-318.

[22] Livernois, John, and Chris J. McKenna. 1999. "Truth or Consequences: Enforcing Pollution Standards." *Journal of Public Economics* 71(3): 415–440.

[23] Lofgren, Don J. 1989. *Dangerous Premises: An Insider's Perspective of OSHA Enforcement.* Ithaca, NY: ILR Press.

[24] Malik, Arun S. 2006. "Optimal Environmental Regulation Based on More Than Just Emissions." *Journal of Regulatory Economics* forthcoming.

[25] Malik, Arun S. 1993. "Self-Reporting and the Design of Policies for Regulation Stochastic Pollution."*Journal of Environmental Economics and Management* 24(3): 241-257.

[26] Nielsen, Viebeke L. 2006. "Are Regulators Responsive?" *Law and Policy* 28(3): 395-416.

[27] Nyborg, Karine and Kjetil Telle. 2004. "The Role of Warnings in Regulation: Keeping Control with Less Punishment." *Journal of Public Economics* 88: 2801-2816.

[28] Russell, Clifford S., Winston Harrington, and William J. Vaughan. 1986. *Enforcing Pollution Control Laws*. Washington, DC: Resources for the Future.

[29] Scholz, John T. 1984. "Cooperation, Deterrence and the Ecology of Regulatory Enforcement." *Law and Society Review* 18(2): 179-224.

[30] Sinclair-Desgagne, Bernard. 1994. "The First-Order Approach to Multi-Signal Principal-Agent Problems." *Econometrica* 62(2): 459-465.

[31] Stigler, George J. 1970. "The Optimum Enforcement of Laws." *Journal of Political Economy* 78:526-536.

[32] Weil, David. 1996. "If OSHA Is So Bad, Why Is Compliance So Good?" *Rand Journal of Economics* 27(3): 618-640.

Appendix

A.1 Proof of Result 2

This result is most easily established by working with the reciprocals of the optimal monitoring probabilities with and without forgiveness, and examining the difference $(1/m^* - 1/\overline{m}^*)$. The larger is this difference, the larger is the reduction in the monitoring probability achieved by forgiveness. Using Result 1, along with (3) and (9), the difference can be written

$$\frac{1}{m^*} - \frac{1}{\overline{m}^*} \equiv \frac{F}{C'(a)} \left\{ [\theta_{N|N} - p(a)\alpha] \sum_{s \in S_0} q_a(s|a) - p'(a)\alpha \sum_{s \in S_0} q(s|a) \right\}. \qquad \text{(A-1)}$$

We can establish that the expression in braces is positive as follows. By definition, for $s \in S_0$, noncompliance is forgiven and the expression in braces in the first condition in (10) is positive. Summing the expression in braces in (10) for all $s \in S_0$ yields the expression in braces in (A-1), which, in turn, must be positive.

Now consider the effect on (A-1) of a reduction in α from α' to α'', due either to a reduction in $\theta_{N|N}$ or an increase in $\theta_{N|C}$. We know from Proposition 1 that S_0 is weakly larger the smaller is α. Thus we can write $n(S_0'') \geq n(S_0')$ with $S_0' \subset S_0''$. First consider the case where S_0 is unchanged: $S_0' = S_0''$. A reduction in α due to an increase in $\theta_{N|C}$ unambiguously increases the difference $(1/m^* - 1/\overline{m}^*)$, since it increases the magnitude of the first term in braces in (A-1) while reducing the magnitude of the second one. The effect of a reduction in $\theta_{N|N}$ is less transparent, since it reduces the magnitude of both terms. However, we know that the first sum in braces is positive, as is the second sum. As is easily verified, the relative weight attached to the first sum, $[\theta_{N|N} - p(a)\alpha]/p'(a)\alpha$, is decreasing in $\theta_{N|N}$. Therefore, a *reduction* in $\theta_{N|N}$ increases the difference $(1/m^* - 1/\overline{m}^*)$.

Now consider the case where $n(S_0'') > n(S_0')$. Let $D \equiv S_0'' \backslash S_0'$ denote the set of additional signals for which noncompliance is forgiven as a result of the reduction in α. The expression in braces in

(A-1) will now be augmented by the term

$$[\theta_{N|N} - p(a)\alpha] \sum_{s \in D} q_a(s|a) - p'(a)\alpha \sum_{s \in D} q(s|a). \tag{A-2}$$

This term must also have a positive sign given (10). Thus, the magnitude of the difference $(1/m^* - 1/\overline{m}^*)$ is further increased.

Finally, using (14), equation (A-1) can be rewritten as

$$\frac{1}{m^*} - \frac{1}{\overline{m}^*} \equiv \frac{\sigma(a)[\theta_{N|N} - p(a)\alpha]F}{C'(a)} \left\{ \frac{\sigma_a(a)}{\sigma(a)} - \frac{p'(a)\alpha}{[\theta_{N|N} - p(a)\alpha]} \right\}. \tag{A-3}$$

The expression in braces is the difference in likelihood ratios, Δ^σ, specified in Result 2. As can be seen from (A-3), an increase in Δ^σ will imply a larger value for $(1/m^* - 1/\overline{m}^*)$, unless the increase in Δ^σ is accompanied by a proportionally larger reduction in the value of $\sigma(a)[\theta_{N|N} - p(a)\alpha]$. The possibility of such a reduction cannot be ruled out.

A.2. Signs of $\partial K/\partial\theta_{N|N}$ and $\partial K/\partial\theta_{N|C}$

Recall that K represents the RHS of (15). Differentiating the RHS of (15) with respect to $\theta_{N|N}$ yields

$$\frac{\partial K}{\partial \theta_{N|N}} \equiv \frac{-q_a(s_M|a)}{p'(a)\alpha^2} + \frac{q(s_M|a)[1 - p(a)]}{[\theta_{N|N} - p(a)\alpha]^2}. \tag{A-4}$$

It can be verified that this expression is negative given the fact that $\alpha[1 - p(a)] < [\theta_{N|N} - p(a)\alpha]$ and the premise that forgiveness is optimal when s_M is observed (so (11) holds).

Differentiating the RHS of (15) with respect to $\theta_{N|C}$ yields

$$\frac{\partial K}{\partial \theta_{N|C}} \equiv \frac{q_a(s_M|a)}{p'(a)\alpha^2} + \frac{p(a)q(s_M|a)}{[\theta_{N|N} - p(a)\alpha]^2}, \tag{A-5}$$

which is positive since $q_a(s_M|a) > 0$ when forgiveness is optimal, as is true by assumption when $s = s_M$.

A.3. Establishing $\mu > 0$ and $\tau > 0$ given Self-Reporting

I first rule out $\mu = 0$ and then rule out $\mu < 0$. If $\tau > 0$, then $\mu = 0$ implies that the first condition in (27) always holds as an inequality, which in turn implies $f_N(s) = 0 \; \forall s$. But then (20) cannot be binding, contradicting the premise that $\tau > 0$. If $\tau = 0$, then the following first-order condition for the monitoring probability, m_C, cannot hold when $\mu = 0$:

$$p(a)[c_m + \theta_{N|C} c_S] = \tau \theta_{N|N} f_{NC} - \mu \theta_{N|C} p'(a) \, f_{NC}. \tag{A-6}$$

Now suppose $\mu < 0$. Then (27) implies $f_N(s) > 0$ only if $[1 - p(a)]q_a(s|a) - p'(a)q(s|a) > 0$, which in turn implies that the sum of the second and third terms in (23) must have a positive sign. But then (23) cannot hold, because the fourth term is non-negative.

We can now rule out $\tau = 0$. Given $\mu > 0$, (A-6) does not hold if $\tau = 0$.

Table 1: Hypothetical Values of K

$q_a(s_M\|a)/p'(a)$	$p(a)$	$q(s_M\|a)$	$\theta_{N\|N}$	$\theta_{N\|C}$	K
4	0.8	0.2	1	0	3
2	0.8	0.2	1	0	1
1	0.8	0.2	1	0	0
6	0.8	0.2	1	0	5
4	**0.7**	0.2	1	0	3.3
4	**0.9**	0.2	1	0	2
4	0.8	**0.1**	1	0	3.5
4	0.8	**0.3**	1	0	2.5
4	0.8	0.2	**0.9**	0	3.3
4	0.8	0.2	**0.9**	**0.1**	4.2
4	0.8	0.2	**0.9**	**0.2**	5.1
4	0.8	0.2	**0.8**	0	3.7
4	0.8	0.2	**0.8**	**0.1**	4.9
4	0.8	0.2	**0.8**	**0.2**	6.0