The Impact of Fines, Enforcement Actions, and Inspections on Environmental Compliance

A Statistical Analysis of the Pulp & Paper Industry

Jay P. Shimshack

Department of Economics

Tufts University

Michael B. Ward

Department of Agricultural and Resource Economics

University of California, Berkeley

Running Head: Fines, Enforcement Actions, and Inspections

This paper explores empirically the impact of monitoring and enforcement efforts on environmental compliance for conventional water pollutants. We consider three key questions about the impact of environmental enforcement strategies. First, how effective are fines at inducing environmental compliance? Second, how effective are less severe intermediate enforcement actions? And third, how much do inspections contribute to compliance on the margin? We find the following answers: Fines significantly reduce effluent violations. Violations by both the fined firm and other firms in the same regulatory jurisdiction are reduced. The impact of less severe intermediate enforcement actions is substantially less than that of fines. Finally, inspections weakly induce additional compliance at the margin.

Key Words: Fines, Inspections, Pollution, Compliance, Enforcement

JEL Classifications: K32 Environmental, Health, and Safety Law

D83 Learning and Information

Q25 Environmental Government Policy

1.0 INTRODUCTION

The Environmental Protection Agency and state regulatory authorities devote considerable resources to ensuring compliance with environmental standards.[1] This paper explores empirically the impact of their regulatory efforts on firms' compliance decisions. To motivate the analysis, we begin by briefly summarizing the current regulatory context. Enforcement of environmental standards for conventional water pollutants is largely based upon self-reported emissions. Frequent on-site inspections verify these self-reported values. Violations may result in one or more enforcement actions, including the levying of fines. Fines for detected violations, however, are rarely imposed.

We proceed by asking three key questions about the impact of environmental enforcement strategies. First, how effective are fines at inducing environmental compliance? Second, how effective are less severe intermediate enforcement actions? And third, how much do inspections contribute to compliance on the margin? We find the following answers: Fines significantly reduce effluent violations. Violations by both the fined firm and other firms in the same regulatory jurisdiction are reduced. The impact of less severe intermediate enforcement actions is substantially less than that of fines. Finally, inspections weakly induce additional compliance at the margin.

This paper builds upon the work of previous empirical studies. Magat & Viscusi (1990) demonstrated that, in the early 1980s, inspections caused American pulp and paper firms to more frequently comply with EPA water pollution standards. Deily & Gray (1991) demonstrated that the EPA is responsive to local economic and political conditions when regulating steel mills, and subsequently conducts fewer regulatory actions against firms with higher probabilities of closing. See also Deily & Gray (1996) and Dion et al (1998). Laplante & Rilstone (1996) then showed that inspections and the threat of inspections negatively impacted

[1] The EPA alone budgeted over $300 million to this task in 1998. Congressional Research Service. "Environmental Protection Agency: An Analysis of Key FY1999 Budget Issues II," 1999.

discharges from Canadian pulp and paper mills in the late 1980's. Nadeau (1997) found that increasing monitoring and enforcement of air pollution standards results in a shorter duration of noncompliance. Finally, Helland (1998) examined the role of targeting[2] in the enforcement of pollution control standards. His empirical evidence suggests that targeting is present, but only for a small subset of firms.

The credible threat of fines, or some other sanctions, is required to deter violations in standard economic models of enforcement.[3] Other authors, including Magat & Viscusi (1990), Laplante & Rilstone (1996), and Helland (1998), have recognized the need to consider effluent fines but data limitations have hindered their empirical investigations. Using a larger and more recent panel dataset, we are able directly address the impact of sanctions on U.S. compliance.[4] We examine the impact of sanctions on both the sanctioned firm and on other firms in the jurisdiction. This is appropriate because the regulator may enhance its credibility with all firms by making an example of a single firm.

Section 2.0 focuses on the regulatory background of our case study industry. Section 3.0 discusses the data, its sources, and the assumptions involved in its collection. Section 4.0 more closely examines firms' compliance decisions, and section 5.0 presents the econometric models. A discussion of both aggregate and plant-level analyses is included. Section 6.0 presents results, interpretations, and conclusions.

2.0 BACKGROUND

Conventional water pollutants for the U.S. pulp and paper industry are the focus of our case study. We choose this industry for several reasons. The pulp and

---

[2] Helland's work is an empirical test of Harrington (1988)'s leverage model.
[3] See for example those works summarized in Russell, Harrington, and Vaughan (1986).
[4] Dasgupta *et al* (1999) examine the impact of pollution fines on environmental performance in China.

paper industry is the largest discharger of both biochemical oxygen demand (BOD) and total suspended solids (TSS) into U.S. waterways, releasing over 16 million cubic meters of wastewater daily. Additionally, pulp, paper, and paperboard mills exist in a wide range of states and fall under the jurisdiction of many different permitting authorities. Major production areas are located where raw materials (fiber-furnish) are most plentiful: the southeast, the northwest, the northeast, and the north central region. Finally, given functioning abatement equipment, the amount of conventional pollutants emitted is directly related to the amount of output produced.[5]

Permitting, inspection, and enforcement activities are conducted by a variety of regulatory authorities. These authorities can either be regional EPA offices or the state in which the mill is located.[6] Compliance information is gathered by on-site inspections and monthly self-monitoring reports. Self-reporting, however, is the primary source of effluent information. The inspection program is a secondary source of compliance evaluation, a source of information for future permitting, and an avenue to gather evidence to support enforcement actions.[7] These actions range from the levying of fines to a warning telephone call.

Prior to 1997, each permitting authority was required by law to inspect major dischargers at least once a year. Five types of inspections directly apply to the non-toxic discharges of a standard industry. The first type is a reconnaissance inspection. This brief type typically lasts less than one day, and simply involves a visual inspection of the facility, its effluent, and its receiving waters. Compliance evaluation inspections and performance audits involve a more in-depth analysis of a plant's compliance. These inspections include the visual monitoring of a firm's self-reporting records to determine accuracy and quality. Regulators check

---

[5] United States EPA, Office of Water. "Development Document for Proposed Effluent Limitations Guidelines and Standards for the Pulp, Paper, and Paperboard Point Source Category," October 1993. This desirable characteristic is, of course, also true of several other industries.

[6] As discussed in Chapter 1, states have the option to oversee compliance. The EPA steps in for states which decline this option.

that equipment required by the permit is in place and being properly operated. Additionally, performance audits involve an inspector observing a plant's sample collection. No regulator sampling is conducted, and these inspections are likely to last between two and 12 days. The final monitoring methods, compliance sampling and bio-monitoring inspections, require approximately thirty days to complete and involve all of the actions and observations of the other types, in addition to regulator sampling.

Inspections are, to some degree, predictable. Before any regulatory monitoring can occur, the inspector must conduct a pre-inspection discussion with the owner / operator of the plant, outlining the inspection's plan. Also, specific inspections must be conducted based upon administrative factors or specific evidence of an existing violation. Historically, the vast majority of resources have been devoted to inspections motivated by administrative factors. In fact, a Supreme Court ruling requires that the EPA base its monitoring activities on "neutral selection," wherein the choice of plants to be inspected is based upon geographic factors and the length of time since the last inspection.[8] Purely random inspections are prohibited.

3.0 DATA

The EPA's Permit Compliance System (PCS) serves as our data source. Established in conjunction with the Clean Water Act and its subsequent amendments, the PCS tracks plant-level self-reported emissions, permitted effluent limitations, inspections, and enforcement actions. Although the EPA administers the PCS, state agencies contribute much of its information. Our sample consists of data generated by 23 separate regulatory jurisdictions. 15 of

[7] United States EPA, Enforcement Division, Office of Water Enforcement and Permits. "NPDES Compliance Monitoring Inspector Training: Overview," August 1990.

[8] United States EPA, Enforcement Division, Office of Water Enforcement and Permits. "NPDES Compliance Monitoring Inspector Training: Overview," August 1990

these jurisdictions contain plants directly regulated by the states in which the they are located. The other eight jurisdictions contain plants regulated by one of four EPA regional authorities. We disaggregate the EPA regional offices into the eight states they represent because we are concerned that regulatory information may not necessarily flow freely between states.[9]

We analyze data for the sample period 1988–1996. Since 1988, data on both effluent levels and enforcement actions are significantly more complete than in prior years. The sample period ends in 1996 because, late that year, the EPA instituted a major regulatory change in inspection procedures.

Our sample of PCS data contains the relevant information for biochemical oxygen demand (BOD) and total suspended solids (TSS) emissions from the 217 "major" pulp, paper, and paperboard mills in our sample states. We consider the conventional pollutants BOD and TSS because all pulp and paper mills produce wastewater with significant amounts of these discharges. We only consider "major" plants because these facilities are required to report their own emissions levels for operating pipes each and every month. The EPA identifies plants as "major" if they have a flow of one million gallons or more per day or pose a significant impact to water quality.[10] We consider all states with four or more major pulp, paper, or paperboard mills. Our 217 examined plants emit from a total of 253 distinct pipes.

Although self-reporting for major firms is mandatory and our dataset is quite complete, pipe closures sometimes result in missing data in the PCS. A probit analysis of missing reports, however, yielded no evidence of strategic non-reporting. In particular, lagged effluent levels do not predict missing data. Additionally, lagged inspections have no systematic impact on missingness.

[9] MA, ME, NH, and TX contain plants regulated solely by the EPA regional offices. AR, LA, NC, and PA contain some plants that are regulated by state permitting authorities and some plants regulated by EPA regional offices.

[10] United States EPA, Center for Environmental Information and Statistics. "Major Findings from the CEIS Review of EPA's Permit Compliance System Database," 1999.

As Laplante and Rilstone (1996) note, the natural question that arises with self-reported data is whether firms strategically misreport effluent discharges. The ideal test of self-reporting would be a secret and random check of effluent concentrations by the regulator. Unfortunately, given the available data, only imperfect tests of self-reporting exist.

Suppose that firms tend to under-report emissions when there is not an inspection. This strategic behavior would result in a positive correlation between inspections and contemporaneously reported emissions levels. On the other hand, suppose firms emit less during an inspection. This behavior would result in a negative correlation between inspections and reported emissions. Any residual correlation (after accounting for exogenous covariates) between inspections and reported emissions indicates strategic firm behavior.[11]

It is possible, though less plausible, that the two effects could exactly offset. Consider, for instance, a plant that always violates the standard when no inspector is present, yet reports no violation and complies when an inspector is present. This possibility, however, requires that the plant can adjust its emissions very quickly and entirely secretly.

The absence of any anomalous residual correlation is consistent with accurate self-reporting. Laplante & Rilstone (1996) test for such a correlation by running a comparison of means test on emissions when an inspector is present versus emissions when no inspector is present. They find no statistical difference. We test whether current inspections, after correcting for possible inspection targeting, have explanatory power for reported emissions. We find a rather small and statistically insignificant correlation, suggesting the plausibility that most firms do not respond strategically to the presence of an inspector. We therefore fail to

[11] Additionally, it may be possible that the regulator is aware of plants' discharges and conducts inspections during periods of high emissions. Such targeting would tend to produce a positive correlation. However, in the empirical section we correct for this possibility.

reject the accuracy of self-reporting. It seems quite likely that this result is due, at least in part, to the potentially large criminal penalties associated with intentional misreporting.

The analysis considers both fines and intermediate enforcement actions (IEAs). We include only those sanctions attributable to BOD or TSS non-compliance. This excludes penalties for other types of violations such as paperwork errors, toxics emissions, or poor equipment maintenance. Although enforcement actions in the PCS are not explicitly linked to particular violations, we are able to distinguish the type of violation. We therefore included all effluent sanctions preceded by one or more BOD or TSS violations in the previous year.[12]

Table 3.1 Summary Statistics by Permitting Authority

Variable	States	EPA Regions	Total
Authorities	15	8	23
Plants	172	45	217
Violations	299	124	423
Fines	22	2	24
Fine Average	$43,500	$97,500	$48,000
Fine Maximum	$600,000	$100,000	$600,000
Fine Minimum	$500	$95,000	$500
IEAs	28	16	44
Inspections	1,718	414	2,132

The reader may find the summary statistics presented in Table 3.1, broken down by EPA and state jurisdictions, useful. Additional details are presented below. The average number of inspections per year is approximately 1.5. Eight permitting authorities levied BOD/TSS fines, and eighteen authorities levied

BOD/TSS formal intermediate enforcement actions (IEAs). To check the dataset's completeness, we confirmed that all 23 authorities record enforcement actions of some sort, including sanctions for non-effluent violations.[13]

Just under half of our sample plants, 99 out of 217, violated their effluent limitations at least once in the sample period. Violations occurred in all 23 jurisdictions. In an average month, over two percent of plants are discovered to be in violation. Violations also appear to be seasonal: one-third as many violations occur in September than occur in January. These numbers indicate that although compliance is generally high in a given month,[14] over time the number of violations is significant.

4.0 COMPLIANCE DECISIONS

Economic logic suggests that inspections and self-reporting alone will not deter violations; economic sanctions are also required. Fines are the obvious sanction. Intermediate enforcement actions are also potentially important because fines themselves are often the culmination of a sequence of less severe sanctions. It is therefore surprising that this is the first paper to examine explicitly how fines and IEAs deter U.S. conventional pollutant violations. Data limitations, combined with the infrequency of fines, have made such an analysis difficult in the past. On the other hand, perhaps sanctions have been overlooked because they are commonly perceived as "only a slap on the wrist."[15]

Our point of departure is traditional law and economic models beginning with Becker's (1968) seminal work on crime. Other relevant citations include Stigler

[12] We were able to track down legal records for several of these sanctions. In each case, the sanctions were in fact for BOD and TSS violations.
[13] The Center for Environmental Information and Statistics has also performed an independent analysis of the reliability of PCS data. They conclude that emissions, limit, inspection, and enforcement data are accurate.
[14] In fact, many plants significantly over-comply with the standards.

(1970), Posner (1977), and Viscusi (1979). The firm is a rational profit-maximizing agent that will exceed effluent standards if it is profitable to do so. In our case, the relevant marginal benefit is the gain from exceeding the permitted average effluent limitation over the course of a month. The marginal cost is the expected sanction for noncompliance.

It is worth noting that under current abatement technologies, effluent typically varies directly with production, and so is a direct choice variable of the firm. Conventional pollutant violations are rarely the result of catastrophic equipment failure, and violations are interpreted as intentional. Specifically, given the relative maturity of BOD and TSS regulatory regimes, adequate investment and maintenance could prevent violations.

In typical enforcement models, the expected sanction is a function of the probability of detection and the nature of sanction if detected. However, our working hypothesis is that self-reported effluent values are accurate, so the probability of detection is one.[16] In the view of the EPA, inspections are intended to ensure the accuracy of self-reported data, rather than to detect and deter effluent violations. Of course, inspectors may also help ensure compliance by identifying potential problems.

The relevant probability in our study is the chance of suffering a sanction for a violation, rather than the chance of detection. Both this probability and the nature of the potential sanction are uncertain to the firm, because the firm has limited information on which to base its expectations. However, the firm learns through experience. The main credible source of information is the actual enforcement history of the regulator.

---

[15] See, for example, Goodstein (1999): pgs 276.

[16] As previously discussed, our analysis includes a test of the accuracy of self-reported data, and fails to reject the hypothesis. Accurate self-reporting is economically plausible because intentional misreporting is always punishable by criminal sanctions. See Cohen (1992).

To model learning, we follow Sah's (1991) work on social osmosis in crime. The firm observes the regulator's past responses to its own violations and to violations of other firms under the same regulator. Here, sanctions levied on one firm in Oregon should increase Oregon's enforcement credibility with other firms in that state, but have no impact on Texas firms.

Within an overall regulatory strategy, actual enforcement is random from the firm's perspective. As the firm updates its beliefs about expected sanctions, it may change its compliance behavior. So, by combining random enforcement data with compliance data, we are able to estimate statistically an analog of Sah's model even if the overall regulatory strategy is unchanging. Of course one may expect that the regime changes over time for a variety of reasons, including political economy and other local factors. It is also possible that the regulator adjusts the regime over time in response to new information about firms. We consider and correct for this possibility in the econometric section. Because of the possibility of changing enforcement regimes, the value of enforcement history information may decline over time.

The direct pecuniary costs of fines alone may not be the only true economic penalties. In our context, one must interpret the role of any sanctions with caution. Specifically, the total expected cost of a sanction may be greater than any direct cost. As Polinsky and Shavell (1998) point out, it may be rational for a regulator to increase sanctions for repeat offenders. Therefore, the expected cost of noncompliance reflects both immediate costs and the possibility of increased future sanctions, including Department of Justice (DOJ) litigation.[17] Even a "slap on the wrist" may serve as a deterrent if it suggests larger future sanctions.

[17] While expected future EPA penalties may significantly alter firms' compliance decisions, it is unlikely that threats of DOJ suits play a large role in our analysis. The overwhelming majority of sanctions imposed by the EPA are internal, and, as Magat and Viscusi (1990) emphasize, "court action is a lengthy process that is started only as a last resort." Further, nearly all water cases submitted to the DOJ

5.0 ECONOMETRICS

We examine the data at two levels: Aggregated to the permitting authority level and disaggregated to the plant level. The aggregate estimation requires fewer assumptions on the error structure of the data and allows for simple interpretation. The plant level estimation, on the other hand, allows for a more detailed look at our three key questions. Additionally, we are better able to control for potential endogeneity of the monitoring process using the plant-level analysis. Both estimation approaches should yield similar results for enforcement parameters, providing an informal check of consistency.

The overall econometric strategy is to identify the effects on firm compliance of variation in monitoring and enforcement intensity over time. As previous empirical researchers have done, we omit a structural theoretical model, primarily because the fundamental incentive process is well understood.[18]

5.1 AGGREGATE ANALYSIS

Simplicity is the primary virtue of analyzing data pooled at the state level.[19] This aggregate analysis allows us to examine the average firm impact of enforcement actions by the permitting authority. The most important of these actions are conducting inspections, imposing intermediate sanctions, and levying fines. Using a fixed-effects panel model, we are able to identify the plants' short-run responses to changes in a permitting authority's enforcement strategy. Since this

regard toxic chemical, metal, or pathogenic discharges, rather than BOD or TSS. See U.S. E.P.A. (2001).

[18] See, for example, Magat and Viscusi and Laplante and Rilstone.

[19] More accurately, we mean permitting authority level, but we refer to the state level to reduce verbiage.

is a fixed-effects regression, identification comes from within-group variation (the time-series), rather than between-group variation (the cross-section).[20]

Although desirable, it would be difficult to attribute cross-state differences in compliance to cross-state variation in inspection and fine frequency. The difficulty is potential endogeneity: The long-run strategy of the permitting authority may be a reaction to the peculiarities of the plants in that state. For example, plants in some states may have an unusually high tendency to violate standards, causing the permitting authority to adopt an unusually vigorous enforcement strategy. Another possibility is that such correlation could be the result of some permitting authorities being more strict than others in both standards and enforcement. Both might well produce a positive cross-state correlation between enforcement activities and violations. In fact, the cross-state correlation between violation rate and inspection rate is about 30%; similarly, the correlation between violation rate and IEA rate is over 60%. Of course, the fixed-effects in the panel regression control for such variation in regulatory strategy across states.

5.1.1 EXPLANATORY VARIABLES

The violation rate for a permitting authority in a given month provides our measure of the effectiveness of enforcement activities. We examine how three classes of regulatory variables impact this measure.

First, we include a dummy (denoted FINED1) indicating whether an effluent fine was levied by that permitting authority in the previous 12 months. If fining one firm establishes the credible threat of fines with other firms under the same authority, we would expect to see a negative coefficient on this variable.

[20] Note that this specification precludes any difficulties arising from the possibility that a plant's permitting authority is endogenous. The fixed-effects model removes any potential bias introduced if there was a specific regulatory reason that

Similarly, we include these variables lagged one year (FINED2) to account for fines levied 13 to 24 months ago. If the credibility of a threat established by a fine decays over time, we would expect to see a smaller coefficient on this variable than on the dummy for more recent fines. Since a fine's magnitude as well as its existence may matter, in a parallel regression we include the log of the sum of the fines (FINEM1 and FINEM2).

The number of intermediate enforcement actions, such as formal notices of noncompliance, comprises our second set of variables. We include intermediate enforcement actions for each of the last two years (IEA1 and IEA2). Again, the impact of intermediate actions might be expected to decay over time. Including the second lag allows us to measure this potential effect.

The final set of regulatory variables measures the rate of inspection. The first two variables are the rate of inspections in the previous year (INS1) and the rate two years ago (INS2). Even without the threat of a sanction, inspections may prevent some violations. For example, an inspector may notice an easily correctable problem. Additionally, inspections are often a necessary precursor to the levying of sanctions. We also include the rate of statewide inspections in the current month (INS0). As previously discussed, this variable primarily serves to provide a weak test of self-reporting accuracy. It is possible that inspections are simultaneously determined with the observed violation rate. This suggests the use of instrumental variables. Unfortunately, we have no instruments at the aggregate level other than lagged inspections, which we consider to be explanatory variables in their own right. However, we later identify valid instruments for the plant-level analysis.

Finally, we capture cross-state variation with state-specific dummies (STATE), seasonality with monthly dummies (MONTH), and a time trend with annual dummies (YEAR).

---

some plants are regulated by the EPA and some are regulated by state

5.1.2 REGRESSION MODEL

Regressing the violation rate on these explanatory variables can be interpreted as a linear probability model. Unfortunately, this model has the undesirable property that the predicted violation rate can be negative. We therefore apply the logit transformation to the violation rate prior to running the regression. Since about 80% of our state-month observations contain no violations, we employ Cox's (1970) correction for sample size. The left-hand side of the regression becomes log $\{(V_{it} + (2n_{it})^{-1}) / (1 - V_{it} + (2n_{it})^{-1})\}$, where V_{it} is the violation rate for state i at time t, and n is the number of pipes. The right-hand side of the regression includes the variables discussed above. For reference, the regression equation is:

$$\log \frac{V + (2n)^{-1}}{1 - V + (2n)^{-1}} = X\beta + \varepsilon.$$

The columns of the matrix X are: CONSTANT, FINED1, FINED2, IEA1, IEA2, INS0, INS1, INS2, $YEAR_j$ (j=1..6), $MONTH_j$ (j=1..11), and $STATE_j$ (j=1..22).[21]

jurisdictions.

[21] For parsimony, we omit time and state subscripts. For each set of the year, month, and state variables, we omit one dummy to achieve full rank. The equation including fine magnitudes is identical after replacing FINED with FINEM.

Table 5.1 Important Coefficients & t-stats for the Aggregate Regression

Variable	Description	Fine Dummies	Fine Logged
FINED1	Fine 1-12 months ago (dummy)	-0.212	
		(-3.43)	
FINED2	Fine 13-24 months ago (dummy)	-0.152	
		(-2.41)	
FINEM1	Fines 1-12 months ago (logged)		-0.021
			(-3.51)
FINEM2	Fines 13-24 months ago (logged)		-0.015
			(-2.58)
IEA1	IEAs 1-12 months ago	-0.225	-0.227
		(-0.88)	(-0.89)
IEA2	IEAs 13-24 months ago	-0.156	-0.151
		(-0.64)	(-0.62)
INS0	Inspection rate this month	0.073	0.073
		(0.88)	(0.88)
INS1	Inspection rate over last 12 months	0.048	0.050
		(1.00)	(1.02)
INS2	Inspection rate over 13-24 months prior	0.037	0.038
		(0.76)	(0.77)

Table 5.1 lists the most important coefficients for the aggregate regressions. We defer interpretation to section 6.0, where we also discuss results from the plant-level analysis. The sample included in these regressions consists of 1932 observations from 23 permitting authorities, each reporting over the 84 months between 1990 and 1996. Analysis of residuals reveals serial correlation of less than 6 months. Consequently, our final estimates are from a GLS regression. The standard errors are robust, using White's heteroskedastic consistent correction and a correction for serial correlation. To conserve space, we do not report the incidental coefficients, such as time and state dummies. However, most of these are significant at conventional levels. There is a clear downward trend in violation rates over time. Not surprisingly, the month dummies show seasonality in compliance, with highest compliance in the summer months.

5.2 PLANT-LEVEL ANALYSIS

Plant-Level data allows for a more detailed examination of the effectiveness of enforcement strategies. Using a probit analysis, we are able to examine a fine's direct and reputation effects. We are also able to predict the probability of inspection for a particular plant, much like a firm itself might do. We then investigate the impact of this threat of inspection on a firm's compliance decision. We similarly examine the impact of lagged inspections on a particular plant. We are also able to control for potential endogeneity of inspections at the plant-level. Finally, the more detailed analysis provides an opportunity to look for the presence of self-reporting anomalies and to better capture the effects of plant and source heterogeneity.

5.2.1 EXPLANATORY VARIABLES

We begin by discussing fine variables. Fines may have a deterrent effect on both the fined plant and on other plants regulated by the same authority. We therefore decompose the deterrent effect of fines into two parts: A direct effect and a reputation effect. The direct effect is a decrease in violations by the particular plant fined for non-compliance. This effect may be particularly strong because sanctions typically increase for frequent violators, thus increasing the expected penalty for future infractions. The more indirect reputation effect is a decrease in violations by plants other than the one fined. Fines on any one plant may increase the regulator's credibility with all plants.

Therefore, our econometric specification includes a dummy variable (FINED1) for fines on a given plant in the last year.[22] This captures the direct deterrent

[22] We use many of the same variable names for the plant level analysis and the aggregate analysis. Variables with the same name in the two analyses reflect the same concept; they are, of course, not literally identical.

effect. In order to capture the reputation effect, we include a variable (OFINED1) indicating whether another plant in the same jurisdiction was fined in the previous year. As with the aggregate analysis, we suspect that the deterrent effects of a fine may decay over time. We therefore also include these variables lagged an additional year (FINED2 and OFINED2).

One might expect that the magnitude of a fine, as well as its mere existence, impacts compliance decisions. A sanction's magnitude goes directly to the expected cost of a violation. So, in a parallel analysis we replace the fine dummies with corresponding magnitude variables. These variables are expressed as the logged sum of fines. The direct effect variables are fines on a given plant in the past year (FINEM1) and lagged an additional year (FINEM2). The reputation effect variables are fines on other plants in the jurisdiction in the past year (OFINEM1) and lagged one additional year (OFINEM2).

As with the aggregate analysis, our second set of enforcement variables is the number of intermediate enforcement actions. We include intermediate enforcement actions in each of the two previous years (IEA1 and IEA2). As with fines, a regulator's reputation may be enhanced with all firms by issuing an IEA against any firm. We therefore also include variables for IEAs levied on other plants within a jurisdiction lagged one and two years (OIEA1 and OIEA2).

We consider the impact of inspections at the plant level, and include both lagged inspections (INS1, INS2) and the probability (PINS), or threat, of an inspection as explanatory variables. Lagged inspections are directly observed. However, the threat of inspection must be inferred from a probit regression of inspection determinants, such as the time since last inspection and lagged compliance. We also include current inspections (INS0) as an explanatory variable. This allows us to examine whether firms respond strategically to the presence of an inspector. This provides the basis of our test of self-reporting anomalies.

We incorporate several other explanatory variables of less direct interest. First, we include a dummy variable (BOD) for the type of pollutant. To capture changes in the plant's technology over time, we use the ratio of actual to permitted emissions lagged by 12 months (ACTPER12). We also include a corresponding dummy (CLOSED12) to allow for pipes closed 12 months prior. Production capacity for the plant (CAP), gathered from an industry directory,[23] is also a covariate since large plants may enjoy economies of scale in abatement or may be more visible targets for enforcement actions. We include a corresponding dummy (MISSCAP) for the few plants where the capacity data was missing from the directory. State-level fixed effects (STATE) correct for unexplained heterogeneity across states. A linear time trend (TIME) captures drift in the average probability of violation. Seasonality terms (MONTH) correct for variability in production rates over the course of a year. We also incorporate dummies for a plant's standard industrial classification (SIC). Finally, the producer price index (PPI) is included to account for variation in output price.

5.2.2 Regression Model

The decision to violate is a dichotomous choice of the type typically estimated using the familiar probit analysis. The latent variable is expected profits conditional on a violation minus expected profits conditional on no violation. This model is sensible even if a random event, such as equipment failure, causes the plant to violate. Such shocks are included in the error process, and can be interpreted as an extremely high cost of compliance for that period.

The basic model is $y_{it}^* = \alpha + X_{it}\beta + \alpha_i + \varepsilon_{it}$; $y_{it} = 1$ if ($y_{it}^* > 0$) and $y_{it} = 0$ otherwise. The term α_i can be thought of as an unobserved fixed effect, i.e., a random effect. The term ε_{it} is the usual idiosyncratic shock, which may be serially correlated over time.

[23] Lockwood Post's Directory of the Pulp, Paper, and Allied Trades.

In order to produce consistent estimates, careful attention must be paid to the model's error structure. There are several potential sources of inconsistency. First, the time-invariant random effect α_i must be uncorrelated with the explanatory variables X_{it}. This concern is the standard motivation for a fixed effects model. The time-varying shock ε_{it} must similarly be uncorrelated with X_{it}. Further, if there were serial correlation in ε_{it}, even lagged inspections may then be correlated with the current error term.

The time-invariant element α_i accounts for plant heterogeneity. Since this random effect partially reflects variation in plants' costs of compliance, it is likely that this term is correlated with the average inspection rate for that plant, as well as other regulatory variables. For example, consider the possibility that regulators frequently inspect plants that generally seem to be more likely to violate. Such correlation would produce an omitted variable bias. In a linear model, one could correct for this problem with fixed effects dummies. Unfortunately, including fixed effects dummies in a panel probit regression yields inconsistent estimates of the slope coefficients for a fixed-length panel.

We correct for this possible bias using Chamberlain's (1980) conditional random effects model. This approach conditions the distribution of the error term's persistent component on the average value of inspections for that plant over the sample period. In practice, this correction is equivalent to including the average inspections for the plant as an explanatory variable. We apply this conditional random effects correction to account for heterogeneity reflected in inspections, fines, IEAs, and emissions. Thus, the average values by pipe of INSP0, FINE1, IEA1, and ACTPER12 are included as conditioning variates. The impact of this conditioning is that time variation, but not cross-sectional variation, in the corresponding variables contributes to identification.

Our second source of potential inconsistency is correlation between the time-variate error term ε_{it} and the explanatory variables X_{it}. Again, our concern is the inspection process. It is possible that a regulator may inspect a given plant more frequently when that plant is more likely than usual, given the other explanatory variables, to be out of compliance. While the conditional random effects approach corrects for general inspection targeting of a plant, it does not correct for variation in idiosyncratic targeting over time.

Instrumental variables estimation is the standard approach to correct for this type of correlation. The obvious difficulty is identifying valid instruments. That is, we need variables correlated with inspections, but not associated with idiosyncratic targeting. Our chosen instrument is the rate of inspections on other plants in the same jurisdiction for that month.[24] Changes in the inspection rate on other plants partially reflect changes in the overall inspection rate within a jurisdiction. So, inspections on a given plant should be positively correlated with the corresponding instruments. We believe that our instrument is not affected by idiosyncratic targeting because the pulp and paper industry is only one component of the various regulators' monitoring responsibilities. Therefore, an additional targeting inspection at a given plant does not necessarily imply one fewer inspection at other plants.

An alternative argument is that there is a predetermined number of inspections per period. If this were the case, our instrument would be weakly correlated with the targeting component, because a targeting inspection at a given plant would imply one fewer inspection at other plants. The data, however, suggests that there seems to be no fixed number of inspections per period for any state. This makes sense because regulators have many other industries to inspect. However, if there were a predetermined number of inspections, then the total number of inspections within a jurisdiction, including those on the plant of interest, would be a valid instrument.

We believe that our chosen instrument, the inspection rate on other plants, is appropriate. However, there is no way to test for the exogeneity of an instrument;

[24] Two of the 23 jurisdictions contain only one plant, so we drop these 2 from the plant-level analysis.

it is simply a maintained hypothesis of the model.[25] We therefore also check the results using the total number of inspections as an instrument. Reassuringly, this does not alter our results substantively in the plant-level analysis. The reader may wish to note that no instruments are necessary for the aggregate regression if the assumption of a predetermined total number of inspections holds.

We also use instruments for the first year of lagged inspections. In the aggregate analysis, we find evidence for serial correlation of less than six months in the error process. The correlation between current and lagged residuals rapidly declines from about 0.3 in the first lagged month to about 0.02 by the sixth lag. Therefore, it seems possible that lagged inspections are correlated with the current error process. For example, if $INS0_{it}$ is correlated with ε_{it} and ε_{it} is correlated with ε_{it+1}, then $INS1_{it+1}$ may be correlated with ε_{it+n}. To give some intuition, suppose that the regulator conducts a targeting inspection when it suspects that a firm is unusually likely to be in violation. Since we find positive serial correlation for six months, the firm may also be unusually likely to violate a few months later.[26] Therefore, the first year of lagged inspections may be correlated with the current error term.

The consistency corrections discussed above are crucial; however, efficiency should also be a consideration. So long as we make these consistency corrections, even an independent probit specification is a valid, but inefficient, method of moments estimator. Explicitly modeling the persistent shocks and serial correlation increases efficiency. Further, it is plausible that there is cross-equation correlation in shocks. For example, shocks to different pollutants out of a given pipe may be correlated. In this case, efficiency can also be gained by explicitly modeling these sources of correlation.

[25] For example, a Hausman test cannot be conducted unless at least one instrument is known to be valid *a priori*. This point is made in most good econometrics textbooks. See, for example, Ruud (2000).

[26] By "unusually likely," we mean unusual after taking into account all of our other explanatory variables.

We therefore need to directly account for possible sources of correlation in the estimation. Unfortunately, a probit of this complexity is computationally problematic using direct integration methods. We therefore apply the method of simulated likelihood (McFadden (1989)) using the idea of Stern's (1992) factor analytic decomposition.

Such simulators are based on decomposing the error process into the sum of several components. Here, we have a persistent random effects part α_i and an idiosyncratic part ε_{it}. Note that the α_i shock appears in all observations for a pipe, inducing correlation. In addition, serial correlation in ε_{it} can be modeled by breaking it smaller parts: $\varepsilon_{it} = \alpha_{it} + \sum_{r=0}^{6} \sigma_r \phi_{it-r}$. Here, σ are scale parameters. Although ϕ and α are independent over time, serial correlation is induced because the variables appear in multiple time periods. For example, the covariance induced by this effect for observations six months apart is $\sigma_6\sigma_0$. Finally, contemporaneous correlation in ε_{it} can be similarly modeled by decomposing the α_{it} term in an analogous fashion.

Again, the basic model is $y_{it}^* = \alpha + X_{it}\beta + \alpha_i + \varepsilon_{it}$; $y_{it} = 1$ if $(y_{it}^* > 0)$ and $y_{it} = 0$ otherwise. The matrix of explanatory variables X consists of the following: CONSTANT, FINED1, FINED2, OFINED1, OFINED2, IEA1, IEA2, OIEA1, OIEA2, PINS, INS0, INS1, INS2, INS2, BOD, ACTPER12, CLOSED12, CAP, MISSCAP, SIC, PPI, TIME, MONTH_j (j=1..11), and STATE_j (j=1..20). We also implement the conditional random effects correction by including the means, by plant, of current inspections, fines, IEAs, and the lagged ratio of actual to permitted emissions. We refer to these variables as CRE.INS, CRE.FINE, CRE.IEA, and CRE.ACTPER. As previously discussed, we also include an instrumental variables correction, following Nelson and Olson (1978).

If $y^* = \alpha + X\beta + \alpha_i + \varepsilon_{it} > 0$, a violation is predicted. Equivalently, if $\alpha + X\beta + \alpha_i + \sum_{r=0}^{6} \sigma_r \phi_{it-r} > -\alpha_{it}$, a violation is predicted. If α_{it} is *iid* normal across time and plants, this yields an independent probit conditional on the other error terms.[27] However, since the other error terms are unobservable, we must take the expectation of the probit likelihood with respect to them. We simulate this by taking the average of 20 independent draws over these terms.

[27] The argument is similar for the contemporaneous correlation case we actually estimate. We omit the additional details for clarity of exposition.

Table 5.2 Important Coefficients & t-stats for the Plant-Level Regression

Variable	Description	Fine Dummies	Fine Logged
FINED1	Fine 1-12 months ago on self (dummy)	-0.627	
		(-3.18)	
FINED2	Fine 13-24 months ago on self (dummy)	-0.349	
		(-1.70)	
OFINED1	Fine 1-12 months ago on other (dummy)	-0.695	
		(-4.64)	
OFINED2	Fine 13-24 months ago on other (dummy)	-0.281	
		(-2.00)	
FINEM1	Fines 1-12 months ago on self (logged)		-0.074
			(-3.26)
FINEM2	Fines 13-24 months ago on self (logged)		-0.036
			(-1.59)
OFINEM1	Fines 1-12 months ago on other (logged)		-0.069
			(-4.47)
OFINEM2	Fines 13-24 months ago on other (logged)		-0.031
			(-2.04)
IEA1	IEAs 1-12 months ago on self	0.144	0.152
		(1.50)	(1.43)
IEA2	IEAs 13-24 months ago on self	-0.162	-0.165
		(-1.34)	(-1.26)
OIEA1	IEAs 1-12 months ago on other	0.045	0.063
		(1.08)	(1.39)
OIEA2	IEAs 13-24 months ago on other	-0.075	-0.083
		(-1.56)	(-1.53)
PINS	Predicted inspection probability	0.530	0.635
		(1.66)	(1.76)
INS0	Inspection this month	0.155	0.142
		(0.77)	(0.63)
INS1	Inspections 1-12 months ago	-0.186	-0.234
		(-2.97)	(-3.36)
INS2	Inspections 13-24 months ago	-0.028	-0.030
		(-0.95)	(-0.90)

Table 5.2 (cont.)

Variable	Description	Fine Dummies	Fine Logged
TIME	Linear time trend	-0.007	-0.008
		(-5.23)	(-5.27)
ACTPER12	Emissions ratio 12 months ago	0.444	0.505
		(6.92)	(7.12)
CLOSED12	Pipe closure 12 months ago	0.362	0.389
		(4.20)	(3.96)
CAP	Plant capacity (kilotons)	-0.222	-0.222
		(-4.57)	(-3.81)
MISSCAP	Plant capacity unknown	-0.133	-0.144
		(-1.61)	(-1.53)
BOD	Pollutant type (BOD or TSS)	0.801	0.911
		(4.58)	(4.76)
CRE.INS	Mean of INS0	2.014	2.500
		(2.86)	(3.25)
CRE.FINE	Mean of DFINE	1.280	0.100
		(2.80)	(1.97)
CRE.IEA	Mean of IEA	1.316	1.631
		(4.54)	(5.10)
CRE.ACTPER12	Mean of ACTPER12	1.180	1.377
		(6.30)	(6.46)

The most important coefficients for this regression are in Table 5.2. The sample included in these regressions consists of 32,953 observations from 253 distinct effluent pipes. The regression predicts infractions reasonably well. For those observations with violations, the predicted probability of violation is over 4 times the average violation rate. The pseudo-R^2 is approximately 0.08, calculated with state-level fixed effects in the restricted regression.

The specification of the error structure turns out to be important. Table 5.2 shows the conditional random effects terms are all positive and significant. For example, examining CRE.INS indicates that, at least in the long term, permitting authorities more frequently inspect those plants that are more likely to be out of compliance. Failure to condition on this fixed-effect would have produced an omitted variable bias in the remaining coefficients. It is relevant for efficiency, but not consistency, that the remainder of the correlation structure is mostly significant. The persistent effects are large and positively correlated across the BOD and TSS equations; likelihood ratio tests for the persistent effects and their correlations are significant at greater than 99%. However, the contemporaneous correlation appears less important, and comes in about zero.

6.0 RESULTS AND CONCLUSIONS

Fines strongly deter violations. Examining the aggregate analysis results in Table 5.1 reveals that the coefficient on a recent fine is negative and strongly significant in both the fine dummy and fine magnitude regressions. For the analysis that includes the fine dummy variables, evaluating the mean of the marginal impacts translates numerically into an average 18% reduction in the statewide violation rate for the year following a fine. In other words, on average, a fine on a single plant within a state translates into an 18 percent reduction in the following year's probability of violation for all other firms regulated by the same authority. In the parallel fine magnitude regression, we find that a one percent increase in a state's total penalty amount translates into an approximately 0.35 percent reduction in the average state-level violation rate for the year following the fine.

A fine's state-level deterrent effects also decay over time. In Table 5.1, the coefficients on fines lagged two years are still negative and significant, but they are noticeably smaller than the corresponding coefficients for the one-year lags discussed in the preceding paragraph. Translating numerically, and evaluating the average of the marginal impacts, suggests that the statewide violation rate decreases by 13% in the second year following the imposition of a fine on a single plant. The coefficients from the parallel fine magnitude regression indicate that a one percent increase in a state's total fine magnitude results in an approximately 0.26 percent reduction in the average state-level violation rate for the 13 - 24 months after the fine.

The plant-level outcomes in Table 5.2 are consistent with the aggregate results discussed above. The coefficients on the reputation-effect fine variables are both negative and strongly significant in both the fine dummy and the fine magnitude regressions. For the year following a fine on a given firm, there is an average 13 percent reduction in the statewide probability of a violation. This effect decays by approximately 50 percent for the second year after the fine. The fine magnitude regressions also suggest that a one percent increase in a fined plant's penalty amount translates into an approximately 0.18 percent reduction in the statewide probability of a violation for the year following the fine. The effect decays to 0.08 percent for the second year.

The plant-level analysis also allows us to examine a sanction's direct impacts on the fined firm. On average, a fined plant will exhibit a 10 percent reduction in the probability of a violation for the year following the sanction. This reduction can again be expected to decay by about 50 percent for the second year. The parallel regression also suggests that a one percent increase in the magnitude of a given plant's fine will result in a 0.20 percent decrease in that plant's probability of violation in the following year. For the 13-24 months following the penalty, a one percent larger fine translates into a .10 percent decrease in the likelihood of a plant-specific violation.

Our results also provide evidence that additional inspections weakly deter violations. Table 5.1 indicates that both current and lagged inspections are statistically insignificant. However, in Table 5.2, which includes an instrumental variables correction, the coefficient on the first year of lagged inspections is indeed negative and significant. Additionally, the deterrent impact of inspections seems to decay quite quickly: The second year has a statistically insignificant impact. Of course, this does not suggest that firms would still comply if there were *no* inspections. Rather, over the *observed* range of variation in inspection rates for a state, there is only a modest difference in compliance rates.

It is instructive to further explore the relative marginal impact of fines and inspections. The coefficient on inspections lagged one year is just over 1/4 as large as the coefficient on one's own fine. On the margin, and at a single plant, a fine is therefore about four times as effective at inducing compliance as an inspection. However, fines have the additional leverage of reducing violations over all plants in the jurisdiction. Taking this effect into account, the marginal fine deters about 30 times as many violations as does an additional inspection.

We are unable to conclusively quantify the impact of intermediate enforcement actions since the coefficients are statistically insignificant in both the aggregate and plant-level analyses. This is the case for both the own and reputation effects. Since it seems quite likely that IEAs have some impact, we interpret our results as implying that the impact is quite modest, at least relative to fines.

Our results support much of the EPA's current approach to the regulation of conventional water pollutants for "major" plants. Self-reporting seems to be an effective monitoring strategy. Thus, random and frequent inspections are not required in order to induce compliance. Therefore, the EPA's current policy of reducing inspections on plants with a history of compliance is consistent with cost effectiveness. On the other hand, monetary sanctions for effluent violations have large deterrent effects for all plants in a jurisdiction.

REFERENCES

Becker, G.S. "Crime and Punishment: An Economic Analysis." *Journal of Political Economy*. 76 (March/April 1968): 169-217.

Chamberlain, G. "Analysis of covariance with qualitative data." *The Review of Economic Studies*, 47 (January 1980): 225–238.

Cohen, M.A. "Criminal law: Environmental crime and punishment: Legal/economic theory and empirical evidence of enforcement of federal environmental statutes." *Journal of Criminal Law and Criminology*. 82 (1992): 1054–1108.

Cox, D.R. *The Analysis of Binary Data*. London: Methuen, 1970.

Dasgupta, S., B. Laplante, N. Mamingi, and H. Wang. "Inspections, pollution prices, and environmental performance: Evidence from China." Technical report, World Bank, Development Research Group, 1999.

Deily, M. and W. B. Gray. "Compliance and enforcement: Air pollution regulation in the U.S. steel industry." *Journal Environmental Economics and Management* 31 (July 1996): 96-111.

Deily, M. and W.B. Gray. "Enforcement of pollution regulations in a declining industry." *Journal of Environmental Economics and Management*, 21(November 1991): 260–274.

Dion, C., P. Lanoie, and B. Laplante. "Monitoring of pollution regulation: Do local conditions matter?" *Journal of Regulatory Economics*, 1998.

Goodstein, E. "Monitoring and Enforcement." *Economics and the Environment*. John Wiley and Sons, 1999.

Harrington, W. "Enforcement leverage when penalties are restricted." *Journal of Public Economics*. 37 (October 1988): 29–53.

Helland, E. "The enforcement of pollution control laws: Inspections, violations, and self-reporting." *The Review of Economics and Statistics*. 80 (February 1998): 141–153.

Laplante, B. and P. Rilstone. "Environmental inspections and emissions of the pulp and paper industry in Quebec." *Journal of Environmental Economics and Management.* 31 (July 1996): 19-36.

Magat W. and K. Viscusi. "Effectiveness of the EPA's regulatory enforcement: The case of industrial effluent standards." *Journal of Law and Economics*. 33 (1990): 331–360.

McFadden, D. "A method of simulated moments for estimation of discrete response models without numerical integration." *Econometrica*, 57 (September 1989): 995–1026.

Nadeau, L. "EPA effectiveness at reducing the duration of plant-level noncompliance." *Journal of Environmental Economics and Management*. 34 (September 1997): 54–78.

Nelson, F. and L. Olson. "Specification and estimation of a simultaneous equation model with limited dependent variables." *International Economic Review*, 1978.

Polinsky, A.M. and S. Shavell. "On Offense History and the Theory of Deterrence." *International Review of Law and Economics*. 18 (1998): 305-324.

Posner, R. *Economic Analysis of Law*. Boston: Little, Brown, 1977.

Russell, C.S., W. Harrington and W.J. Vaughan. "Economic Models of Monitoring and Enforcement." *Enforcing Pollution Control Laws*. Resources for the Future, 1986.

Ruud, P.A. "Instrumental Variables Estimation." *An Introduction to Classical Econometric Theory*. Oxford University Press, 2000.

Sah, Raaj K. "Social Osmosis and Patterns of Crime." *Journal of Political Economy.* 99 (1991): 1272-1295.

Stigler, G.J. "The Optimum Enforcement of Laws." *Journal of Political Economy.* 78 (1970): 526-536.

Stern, S. "A method for smoothing simulated moments of discrete probabilities in multinomial probit model." *Econometrica*. 60 (July 1992) : 943–952.

Viscusi, K. "The Impact of Occupational Safety and Health Regulation." *Bell Journal of Economics*. 10 (Spring 1979): 117-140.

Lockwood-Post's Directory of the pulp, paper and allied trades. San Francisco: Miller Freeman, 1994.

United States Congressional Research Service. "Environmental Protection Agency: An Analysis of Key FY1999 Budget Issues II," 1999.

United States EPA, "EPA Achieves Significant Compliance and Enforcement Progress in 2001," Headquarters Press Release. Jan 31, 2002.

United States EPA, Center for Environmental Information and Statistics. "Major Findings from the CEIS review of EPA's Permit Compliance System Database," 1999.

United States EPA, Enforcement Division, Office of Water Enforcement and Permits. "NPDES Compliance Monitoring Inspector Training: Overview," August 1990.

United States EPA, Office of Water. "Development Document for Proposed Effluent Limitations Guidelines and Standards for the Pulp, Paper, and Paperboard Point Source Category," October 1993.