

An Investigation of Voluntary Discovery and Disclosure of Environmental Violations Using Laboratory Experiments

JAMES J. MURPHY*

Department of Resource Economics &
Center for Public Policy and Administration
University of Massachusetts-Amherst

JOHN K. STRANLUND

Department of Resource Economics
University of Massachusetts-Amherst

February 2006

Forthcoming in:

Experimental Methods in Environmental Economics . T. Cherry, editor. Oxford: Routledge.

Abstract: This paper uses laboratory experiments to test individual responses to policies that seek to encourage firms to voluntarily discover and disclose violations of environmental standards. We find that it is possible to motivate a significant number of voluntary disclosures without adversely affecting environmental quality. In doing so, government monitoring effort can be reduced because there is no need to monitor the subset of firms that report their violations. In addition, disclosure policies will result in more violations being sanctioned, but fewer of these sanctions are for violations that are uncovered by the government. In general, the theoretical literature on the role of self-reporting in law enforcement and our experimental tests of this theory suggest that the question of whether disclosure policies are a worthwhile complement to regular environmental enforcement will have to be answered on a case-by-case basis.

Keywords: enforcement, compliance, environmental standards, self-reporting, self-auditing
voluntary disclosure, experiments

JEL Codes: C91, L51, Q58

* Correspondence to James J. Murphy, Department of Resource Economics, 80 Campus Center Way, 219A Stockbridge Hall, University of Massachusetts-Amherst, Amherst, MA 01003, USA. Phone: (413) 545-5716, Fax: (413) 545-5853, E-mail: murphy@resecon.umass.edu.

An investigation of voluntary discovery and disclosure of environmental violations using laboratory experiments

James J. Murphy
John K. Stranlund

INTRODUCTION

State and federal self-discovery and disclosure rules seek to encourage greater compliance with environmental regulations by reducing penalties for violations that are voluntarily discovered and reported to authorities. For example, the EPA's Audit Policy reduces penalties "for regulated entities that voluntarily discover, promptly disclose, and expeditiously correct noncompliance."¹ Concurrent with the implementation of rules for voluntary discovery and disclosure of environmental violations over the last decade or so, a significant body of literature emerged that examines the conceptual properties of these rules (e.g., Malik, 1993; Kaplow and Shavell, 1994; Innes, 1999, 2001a, 2001b; Pfaff and Sanchirico, 2000). Taken as a whole this literature is noncommittal on the question of whether voluntary disclosure policies are worthwhile complements to conventional enforcement strategies. In fact, provided that the predictions about the performance of voluntary disclosure policies hold up under empirical scrutiny, it is clear that whether these schemes are worthwhile will depend upon the specifics of particular regulatory settings.

Unfortunately, empirical analyses of the performance of voluntary disclosure policies are limited to just a few examinations of the effects of existing state and federal discovery and disclosure rules. For example, Stafford (2005) finds evidence that the EPA's Audit Policy and state audit policies have had a positive effect on compliance among hazardous waste facilities. Pfaff and Sanchirico (2004) examine the effects of the Audit Policy on the number and form of self-

disclosed violations, and find that the policy has encouraged self-discovery and disclosure of violations, but these reported violations are minor in comparison with the violations uncovered by conventional EPA audits.

While econometric studies with field data are critical for understanding the effectiveness of existing policies, data limitations and the inability to vary these policies in a controlled setting can preclude direct tests of theoretical predictions. Moreover, experiments provide direct control over the parameters of interest, which allows researchers to perform sensitivity analyses that may not be possible outside of the lab. Therefore, in this paper we report the results of a series of experiments designed to test fundamental hypotheses about the performance of voluntary discovery and disclosure policies. In particular, we address the following questions: How well do these policies perform in terms of motivating firms to voluntarily investigate whether they are in violation of an environmental standard and to disclose any violations they discover? How do voluntary discovery and disclosure policies affect the care that firms exercise to prevent environmental violations? Relative to conventional enforcement strategies, what are the effects of these policies on enforcement effort and environmental quality?

We designed and conducted a series of experiments with seven treatments. Using a within-subject design, each subject participated once in each of the seven treatments. All experiments began with a conventional enforcement model as the baseline treatment. In this treatment subjects were responsible for making a costly decision about the level of care taken to reduce the likelihood of a violation occurring. Subjects did not incur any costs if a violation occurred; however, they were audited with a known, exogenous probability, and they were penalized if a

violation was discovered. The elements of the conventional enforcement treatment were contained in the other six treatments. Each of the other treatments gave subjects the opportunity to voluntarily disclose their violations under different conditions. These treatments varied according to the penalty for voluntarily disclosed violations and whether it was costly for the subjects to determine their compliance status.

Subjects responded strongly to the disclosure incentive. In each of our treatments involving an opportunity for voluntary disclosure, a significant number of subjects chose to disclose. As expected, the number disclosers tended to fall as the automatic penalty for reported violations was increased. The policy significance of inducing a significant number of voluntary disclosures is well known—relative to a conventional enforcement strategy, the government can reduce the effort it expends to detect violations because it can focus these efforts on the subset of firms that do not disclose a violation (Malik, 1993; Kaplow and Shavell, 1994).

However, reducing the penalty for disclosed violations to motivate more self-reporting also reduced the care that the subjects took to avoid these violations. Thus, we find strong evidence of a tradeoff between increased violation disclosures and reduced environmental quality. This does not mean, however, that every disclosure policy results in lower environmental quality. In fact, under the condition that subjects did not have to pay to discover their compliance status, we find that it is possible to induce a significant number of violation disclosures without affecting the deterrence of a conventional enforcement strategy (Kaplow and Shavell, 1994). However, attempting to induce increasing numbers of voluntary disclosures will at some point result in less

deterrence relative to a conventional enforcement strategy, and hence, will eventually lead to reduced environmental quality.

We also find strong support for a hypothesis of Malik's (1993) that, relative to conventional enforcement, disclosure policies will result in more violations being sanctioned, but fewer of these sanctions are for violations that are uncovered by the government. Sanctioning violations is likely to be costly. If the costs of sanctioning voluntarily disclosed violations are roughly equal to the costs of sanctioning violations that the government uncovers, then voluntary disclosure policies will tend to increase sanctioning costs. However, because fewer sanctions are applied to violations that the government uncovers, Malik (1993) argues that a voluntary disclosure policy may decrease sanctioning costs, in spite of the increase in the number of sanctions applied, because punishing disclosed violations is probably less expensive than punishing violations that the government uncovers. A firm that voluntarily discloses a violation is essentially admitting liability for being noncompliant. This admission can reduce the burden on the government to produce sufficient evidence for a finding of liability. Moreover, a firm that voluntarily admits liability is less likely to engage in costly efforts to challenge or otherwise avoid the imposition of a penalty.²

Although our results are largely consistent with the qualitative predictions of the existing theory regarding the role of voluntary disclosure in regulatory enforcement, we do observe one unanticipated effect. For each disclosure policy we examined, subjects who chose not to disclose their violations when given the opportunity to do so tended to exercise more care to avoid violations than under a conventional enforcement strategy. This is unexpected because a subject

who chooses to not disclose a violation opts to face the identical random monitoring and penalty as under conventional enforcement. That the addition of a voluntary disclosure policy to a conventional enforcement strategy tended to induce non-reporters to exercise more care likely suggests a framing effect associated with the opportunity to voluntarily disclose violations. It is important to ask whether this effect is simply an artifact of the laboratory setting, or if there is some reason to believe that regulated firms would likely behave in this way. Since we see no reason to expect that this framing effect would hold in non-laboratory regulatory settings, our view is that it is probably limited to the lab.

Despite this framing effect, our work provides strong empirical evidence of the fundamental tradeoffs inherent in voluntary discovery and disclosure policies. Thus, the policy significance of our work is clear. Both the theoretical underpinnings of this work and our experimental tests make it clear that any conclusions about the relative benefits and costs of voluntary disclosure policies will require detailed knowledge of monitoring costs, sanctioning costs, the harm caused by environmental violations, and firms' costs of internal audits to determine their compliance status. Therefore, it is likely that the question of whether disclosure policies are a worthwhile complement to regular environmental enforcement will have to be answered on a case-by-case basis.

THEORY AND HYPOTHESES

The theoretical underpinnings of our study are drawn from a simple model of an industry composed of n identical risk neutral firms.³ Each firm chooses a level of care to reduce the

probability, p , of a violation of an environmental standard. Each has a profit function $v(p)$, with $v'(p) > 0$, and $v''(p) < 0$. Under conventional enforcement of the standard, firms do not have an opportunity to disclose their violations. The government randomly audits a subset of firms so that the probability that any firm will be audited is π . Uncovered violations are punished with a monetary penalty ϕ . A risk neutral firm chooses the probability that a violation occurs to maximize its expected profit, $V(p, \pi\phi) = v(p) - p\pi\phi$. The interior choice of the probability of a violation is $p^*(\pi\phi)$, which is the implicit solution to $v'(p) - \pi\phi = 0$.

The main hypotheses of our work are devoted to examining the behavior of firms when voluntary disclosure rules are added to an existing conventional enforcement strategy, and the policy implications that flow from these behavioral hypotheses.⁴ All of the behavioral hypotheses are focused on the reporting and care decisions of firms under disclosure rules that would leave risk neutral firms indifferent between voluntarily disclosing their violations and choosing instead to face the random monitoring and penalty of conventional enforcement. These disclosure rules provide useful benchmark policies from which we can derive the relative merits of voluntary disclosure policies. Moreover, we conducted additional experiments to provide sensitivity analyses around these benchmarks.

An obvious starting point is to ask whether a disclosure policy can motivate noncompliant firms to voluntarily report their violations to the government, and whether this has any affect on deterrence. Suppose that firms are given the opportunity to voluntarily disclose their violations to the government, and those that do so are penalized $\phi_d < \phi$ automatically (the subscript d indicates voluntary disclosure). Under the common theoretical assumption that firms always

disclose their violations when they are indifferent about doing so, a risk neutral firm will disclose a violation if and only if the automatic penalty for a disclosed violation does not exceed the expected penalty it faces if it fails to report the violation; that is, disclosure occurs if and only if $\phi_d \leq \pi\phi$. From a theoretical perspective, setting $\phi_d = \pi\phi$ so that risk neutral firms are indifferent between disclosing their violations and not doing so implies that all violations will be disclosed. Of course, in a laboratory setting we do not expect that indifferent subjects will always choose to disclose, nor do we expect that all subjects are risk neutral. Nevertheless, we test the following hypothesis:

HYPOTHESIS 1: Any voluntary disclosure policy that leaves risk neutral firms indifferent between disclosing their violations and not disclosing them will motivate a significant number of voluntary disclosures.

When risk neutral firms have costless and perfect information about their compliance status, setting $\phi_d = \pi\phi$ yields the care choice $p^*(\phi_d) = p^*(\pi\phi)$, which implies no effect on deterrence.

Therefore, we have:

HYPOTHESIS 2: Suppose that firms know their compliance status without a costly self-audit. A voluntary disclosure policy that leaves risk neutral firms indifferent between disclosing their violations and not disclosing them will not change the care that disclosers take to avoid violations.

However, complex regulations and production technologies may make it difficult for firms, particularly large firms, to determine whether they are in compliance with environmental standards without undertaking a costly self-audit of their operations (Pfaff and Sanchirico, 2000). With costly discovery, Innes (2001b) has shown that inducing voluntary discovery and disclosure of violations requires that the certain penalty for disclosed violations must be reduced to compensate firms for their discovery costs. However, doing so weakens deterrence. To demonstrate this result, suppose that a firm incurs a cost c to discover whether it has violated the standard. If the firm does not invest in discovery, then its expected payoff is the same as under conventional enforcement, that is, $V(\pi\phi) = v(p^*(\pi\phi)) - p^*(\pi\phi)\pi\phi$. However, if the firm has invested in discovery and has discovered a violation, then the firm will report the violation if $\phi_d \leq \pi\phi$. Assuming that this holds, a firm's choice of violation probability is $p^*(\phi_d)$ if it invests in self-discovery and its expected payoff is $V(\phi_d) - c$, where $V(\phi_d) = v(p^*(\phi_d)) - p^*(\phi_d)\phi_d$. Clearly, the firm is indifferent to discovery and disclosure if $V(\phi_d) - c = V(\pi\phi)$. Since this requires $V(\phi_d) > V(\pi\phi)$, the penalty for disclosed violations must be strictly lower than the expected penalty under conventional enforcement, thereby weakening deterrence. Consequently, $p^*(\phi_d) < p^*(\pi\phi)$, and we have:

HYPOTHESIS 3: If firms must incur a cost to discover their compliance status, then a voluntary disclosure policy that leaves risk neutral firms indifferent between voluntarily discovering and disclosing their violations and facing a conventional enforcement strategy will motivate disclosers to decrease the care they take to avoid violations.

Our final behavioral hypothesis deals with firms that choose not to disclose their violations. Obviously, firms that choose not to disclose their violations are simply choosing to face the conventional enforcement strategy. Therefore, the addition of a voluntary disclosure rule to a conventional enforcement strategy should have no effect on the care that non-disclosers take to avoid violations. Of course, this is true of any voluntary disclosure rule, leading to:

HYPOTHESIS 4: Add any voluntary disclosure rule to a conventional enforcement strategy.

Those firms that choose not to report their violations will not change the care they take to avoid violations.

From Hypotheses 1 through 4 follow several important policy implications that reveal the relative merits of voluntary disclosure policies. First, if Hypotheses 2 and 4 hold, Kaplow and Shavell (1993) have shown:

POLICY IMPLICATION 1: If firms know their compliance status without a costly self-audit, then a voluntary disclosure policy that leaves risk neutral firms indifferent between disclosing their violations and not disclosing them will not change the expected number of violations.

But if discovery is costly and a disclosure rule motivates a significant number of firms to report their violations (Hypothesis 1), and if Hypotheses 3 and 4 hold, then we have the following result due to Innes (2001b):

POLICY IMPLICATION 2: If firms must incur a cost to discover their compliance status, then a voluntary disclosure policy that leaves risk neutral firms indifferent to voluntary discovery and disclosure will increase the expected number of violations.

While a disclosure rule may or may not reduce deterrence depending on whether firms must undertake a costly self-audit to determine their compliance status, these rules will always allow the government to reduce its monitoring effort if they motivate a significant number of voluntary disclosures (Malik, 1993; Kaplow and Shavell, 1994). Since the government does not need to audit those that disclose their violations, it can focus its monitoring effort on the subset of firms that do not report a violation. Clearly, maintaining the same level of deterrence for those who do not report a violation requires fewer audits. This result holds regardless of whether firms must conduct a costly self-audit. Thus, if a significant number of firms are motivated to disclose their violations (Hypothesis 1), then we have:

POLICY IMPLICATION 3: A voluntary disclosure policy that leaves risk neutral firms indifferent to disclosing their violations will reduce the number of audits that are required to maintain the same level of deterrence for those that do not disclose their violations.

While we expect that adding a voluntary disclosure policy to a conventional enforcement strategy will allow the government to reduce its monitoring effort, we also expect that more violations will be sanctioned. If firms do not have to pay to determine their compliance status, then Policy Implication 1 asserts that the expected number of violations will be unchanged. However, the voluntarily disclosed violations are sanctioned with certainty, whereas without the

disclosure opportunity these violations would only be sanctioned with the probability of an audit. Thus, if a significant number of violations are disclosed (Hypothesis 1) when firms know their compliance status without cost, then a voluntary disclosure policy that leaves risk neutral firms indifferent to disclosing their violations and not doing so will increase the total number of sanctioned violations. This move toward more sanctions is reinforced when firms must audit themselves to determine their compliance status, simply because deterrence is weaker and the expected number of violations increases (Policy Implication 2). Thus, we have:

POLICY IMPLICATION 4: Any voluntary disclosure policy that leaves risk neutral firms indifferent about disclosing their violations and not doing so will increase the expected number of total sanctions.

Since sanctioning noncompliant firms is likely to be costly, Kaplow and Shavell (1994) note that it is possible that a voluntary disclosure rule could increase enforcement costs if the additional costs of sanctioning a higher number of violations outweigh the reduction in monitoring costs. However, the effect on sanctioning costs is complicated by the fact that voluntary disclosure policies induce a shift from penalizing violations that are uncovered by the government to penalizing voluntarily disclosed violations. Malik (1993) argues that punishing violations that are voluntarily disclosed is probably cheaper than punishing violations that the government uncovers, because punishments for disclosed violations require less evidence and are less likely to be challenged. Although more violations are punished, fewer sanctions are levied for violations that are uncovered by the government. This is a simple consequence of Hypotheses 1 and 4—firms that do not disclose their violations choose the same level of care to prevent their

violations as under conventional enforcement (Hypothesis 4), but there are fewer violations that the government uncovers because only a subset of firms choose not to disclose their violations (Hypothesis 1), hence:

POLICY IMPLICATION 5: A voluntary disclosure policy that leaves risk neutral firms indifferent to disclosing their violations will reduce the expected number of sanctions that are levied on undisclosed violations.

Before we move on to our experiments it is worth saying a few words about risk preferences. The theory we present, the hypotheses we test, and the implications that follow from these hypotheses are all based on the assumption that agents are risk neutral. We refrain from developing and testing a theory with more flexible risk preferences because the main motivation of this work is to test hypotheses from the existing theory of self-reporting in law enforcement. To our knowledge there is no theoretical work in this area that allows for flexible risk preferences. Thus, developing new theory to account for non-neutral risk preferences is beyond the scope of this paper. Despite its limitations, the existing theory of voluntary disclosure policies with risk-neutral firms provides a useful benchmark from which to judge the qualitative effects of disclosure policies on the variables of real policy significance, that is, the effects of these policies on overall deterrence and government enforcement efforts. Nevertheless, we do believe that all experimental studies that examine compliance behavior in various settings could benefit from information about subjects' risk preferences.⁵

EXPERIMENTAL DESIGN

Our experiments were designed to test the hypotheses and policy implications presented in the previous section, and were conducted in a computer lab using software specifically developed for this research. In all treatments, subjects were responsible for making a production decision that yielded earnings, v . When they produced, there was a probability, p , that a violation would occur.⁶ Subjects could reduce the likelihood of a violation, but this was costly in terms of foregone production earnings, in particular, $v(p) = 3.60 - [0.55/(0.30 + p)]$. The computer screen presented each subject with a table that displayed all the possible violation probability/production earning combinations in 0.05 increments between 0.05 and 0.95.

Table 1 summarizes the experimental design. The Conventional Enforcement treatment (CE) formed the baseline; the remaining six treatments built upon CE such that all features of CE were common throughout the experiment. Under the CE treatment, each subject knew that they would be audited with probability $\pi = 0.6$. If a violation occurred and was uncovered by an audit, then the subject incurred a fine of $\phi = \$2.50$. Subjects in this treatment did not have an opportunity to voluntarily disclose their violations.

<INSERT TABLE 1 ABOUT HERE>

The middle column of Table 1 contains our Voluntary Disclosure Only treatments: D-H, D-I, and D-L. In these treatments, subjects knew automatically and without cost whether a violation occurred. These treatments were identical to the CE treatment, except that subjects had the option to voluntarily disclose whether a violation occurred. If a subject chose not to disclose a

violation, then she faced the identical enforcement strategy as CE (0.6 audit probability, \$2.50 fine). However, if she chose to voluntarily disclose a violation, then she automatically paid a reduced fine, ϕ_d . The level of this fine is the distinguishing factor among the D-H, D-I and D-L treatments and is shown in parentheses next to the treatment labels in Table 1.

In treatment D-I, the automatic penalty for a voluntarily disclosed violation ($\phi_d = \$1.50$) was set such that a risk-neutral subject would be indifferent between disclosing a violation if one occurred and facing the uncertainty of the conventional enforcement strategy. Note that this penalty equals the expected penalty under conventional enforcement; that is, $\pi\phi = \phi_d = \$1.50$. To examine the responsiveness of the subjects to the voluntary disclosure incentive, we chose a higher disclosure penalty of \$2.35 for the D-H treatment and a lower disclosure penalty of \$0.97 for the D-L treatment.

The final column of Table 1 contains our Voluntary Disclosure with Costly Discovery treatments: CD-H, CD-I, and CD-L. These treatments were the same as the Voluntary Disclosure Only treatments, except that subjects did not know whether a violation occurred unless they paid \$0.20 to find out. Those who chose not to pay the cost of self-discovery could not voluntarily disclose a violation and therefore faced the identical enforcement strategy as CE. Like the Voluntary Disclosure Only treatments, the Costly Discovery treatments varied according to the penalty for disclosed violations. In treatment CD-I, the automatic penalty for a disclosed violation was $\phi_d = \$0.97$. In theory, this disclosure penalty makes a risk neutral subject indifferent between discovering and disclosing a violation, or facing the conventional enforcement strategy. Notice that this automatic penalty for a disclosed violation in the CD-I

treatment is lower than the expected penalty under conventional enforcement ($\pi\phi = \$1.50$). This is necessary to motivate subjects to invest in self-discovery. As with the Voluntary Disclosure Only treatments, we chose a higher disclosure penalty for CD-H and a lower penalty for CD-L to examine the responsiveness of subjects to the disclosure incentive.

For the six treatments that included the option to voluntarily disclose a violation, we used the strategy method to ensure that we had an observation for each subject's disclosure decision regardless of whether a violation occurred. Before it was revealed whether a violation occurred, subjects had to decide whether they would commit to voluntarily disclosing a violation if one occurred, or face the uncertainty of random audits and potential penalties under conventional enforcement. Conceptually, since the disclosure decision was not costly, forcing subjects to commit to this decision at the outset should not affect their behavior.

A total of 180 students were recruited from the student population at the University of Massachusetts, Amherst. Subjects were paid \$5 for agreeing to participate and showing up on time, and were then given an opportunity to earn additional money in the experiment. These additional earnings ranged between \$10.55 and \$18.27, with a mean of \$14.88 ($\sigma = 1.49$).

Earnings were paid in cash at the end of each experiment. Each experiment lasted about an hour and a half. Subjects were given a copy of the instructions that the experimenter then read aloud.⁷ The experimenter used an overhead projector to demonstrate the software while the subjects performed the same tasks on their individual computers. It took about 30 minutes to complete the instructions and answer any questions. The same experimenter conducted all sessions.

Every subject participated in all seven treatments, starting with Conventional Enforcement as the baseline. The remaining six treatments were presented in one of six sequences using a Latin Square design to control for possible order effects; 30 subjects participated in each of the six sequences. The sequences of treatments are provided in Table 2. Within a sequence (*i.e.*, a row in Table 2) there were seven stages, one for each treatment. A stage consisted of three practice rounds, followed by one “real money” round. The parameters for the practice and real rounds were the same; data from the practice rounds were discarded. Thus, each subject generated seven observations, one for each treatment. The Latin Square was constructed such that each treatment appears once in each sequence, once in each stage (*i.e.*, a column in Table 2), and each treatment precedes and follows every other treatment one time.

<INSERT TABLE 2 ABOUT HERE>

RESULTS

The tests of the behavioral hypotheses and their policy implications that we specified in Section 2, as well as sensitivity analyses with respect to the disclosure incentive, are conducted with the data in Table 3. The second and third columns of this table contain the mean violation probabilities by disclosure decisions for each treatment, as well as the numbers of individuals who did and did not commit to voluntary disclosure. To show how we calculated the remaining values in Table 3, define these variables in the following way: for a particular treatment, \bar{p}_d and \bar{p}_{nd} are the mean violation probabilities for those who did and did not commit to disclosing their violations, respectively; n_d is the number of subjects who committed to disclosure, and, given 180 subjects, $180 - n_d$ is the number of subjects who chose not to commit to disclosure.

<INSERT TABLE 3 ABOUT HERE >

Using the mean violation probabilities and the numbers of individuals who did and did not commit to disclosing their violations, we calculated the expected number of violations for those who committed to disclosure, $\bar{p}_d n_d$, and the expected number violations for those who did not, $\bar{p}_{nd} (180 - n_d)$. These values are reported in the fourth column of Table 3. Note that the expected number of violations under CE is simply the mean violation probability for this treatment times the number of subjects.

Table 3 also includes the expected number of audits necessary to maintain the $\pi = 0.6$ audit probability for those who did not report a violation. To calculate these values, note first that individuals who committed to disclosure might have been subject to an audit if they did not experience a violation, and hence, did not submit a violation report. For a particular treatment, the expected number of individuals who committed to disclosure but were subject to an audit because a violation did not occur is $(1 - \bar{p}_d) n_d$. Obviously, all who did not commit to disclosure were subject to an audit. Therefore, the expected number of audits required to maintain the $\pi = 0.6$ probability of a random audit for those who did not disclose a violation is $\pi[(1 - \bar{p}_d) n_d + (180 - n_d)] = (0.6)(180 - \bar{p}_d n_d)$. These values are reported in the fifth column of Table 3. The required number of audits under CE is simply $\pi N = (0.6)(180) = 108$.

Finally, we calculated the expected numbers of fines levied on disclosed and undisclosed violations. Since fines are levied on all reported violations, the expected number of fines for disclosed violations in a particular treatment equals the expected number of violations of those

who committed to disclosure, $\bar{p}_d n_d$. On the other hand, fines for undisclosed violations are levied with the probability of an audit. Thus, the expected number of fines for undisclosed violations is the expected number of these violations times the audit probability; that is, $(0.6)(\bar{p}_{nd})(180 - n_d)$. In the final column of Table 3 we report the expected numbers of fines for disclosed and undisclosed violations for each treatment, as well as their sums. The expected number of fines under CE is the expected number of violations for this treatment times the audit probability.

Before discussing the results contained in Table 3, it is worth noting that, as expected, the observed outcomes do not match specific predictions about violation probabilities and reporting choices based on a model of risk neutral, expected payoff-maximizing agents. Violation probabilities tended to be higher than what a risk-neutral subject would be expected to choose. Furthermore, no risk neutral subject would choose to disclose a violation in the D-H and CD-H treatments, yet a non-trivial minority of subjects did so. Similarly, every risk neutral subject would choose to disclose their violations in the D-L and CD-L treatments, but a non-trivial minority of subjects chose not to do so.

All of our hypotheses and their policy implications entail pairwise comparisons of the treatment effects of each disclosure treatment relative to conventional enforcement. Because each subject participated first in the conventional enforcement treatment and then once in each of the six disclosure treatments, we preserve the within-subject comparison by using the nonparametric Wilcoxon signed rank test for matched pairs.⁸

Behavioral Hypotheses

Hypothesis 1 holds as expected. In both the D-I and CD-I treatments, there is a roughly even split between the number of people who committed and who did not commit to disclosing their violations (79:101 for D-I and 94:86 for CD-I). This confirms our expectation that a significant number of subjects would choose to disclose their violations under disclosure policies that make risk neutral individuals indifferent between disclosure and non-disclosure.

Now consider the impacts of voluntary disclosure on the care taken by individuals to prevent violations in the D-I and CD-I treatments relative to their choice under CE (conventional enforcement). When self-discovery is costless, as in D-I, there should be no change in the choice of violation probabilities by those who committed to disclosing their violations (Hypothesis 2). Our results are consistent with this hypothesis: for those who committed to disclosure in D-I, their mean violation probability (0.534) is not statistically different from their mean violation probability under CE (0.518, $p = 0.61$).⁹ However, when it is costly for subjects to discover whether a violation occurred, as in CD-I, the disclosure penalty must be reduced below the expected penalty under conventional enforcement in order to induce discovery. Since deterrence is weaker for those who choose to discover and disclose, we should observe higher violation probabilities for these individuals relative to their choices under conventional enforcement (Hypothesis 3). As predicted, the mean violation probability for those subjects who committed to discovery and disclosure under CD-I (0.616) is significantly higher than the mean violation probability made by these same subjects under CE (0.506, $p = 0.01$).

Although those subjects who committed to disclosing their violations under the D-I and CD-I treatments behaved as theory predicts, those who did not commit to disclosing their violations behaved unexpectedly. Hypothesis 4 asserts that these individuals should not change their violation probabilities when a disclosure policy is added to a conventional enforcement strategy, simply because choosing to not disclose their violations means they are choosing to face the unchanged conventional enforcement strategy. Instead, in all six voluntary disclosure treatments, those subjects who did not commit to disclosure tended to choose lower violation probabilities than under CE. Let Δ_t be the mean change in violation probability from CE for those who did not commit to disclosure in voluntary disclosure treatment t , and let p be the result of a Wilcoxon signed rank test of the hypothesis that there is no change. We observe: $\Delta_{D-H} = -0.053$ with $p = 0.00$; $\Delta_{D-I} = -0.035$ with $p = 0.02$; $\Delta_{D-L} = -0.02$ with $p = 0.64$; $\Delta_{CD-H} = -0.079$ with $p = 0.00$; $\Delta_{CD-I} = -0.065$ with $p = 0.01$, and $\Delta_{CD-L} = -0.047$ with $p = 0.24$. Clearly we reject the hypothesis of equal violation probabilities in four of the six voluntary disclosure treatments (including D-I and CD-I), which is inconsistent with Hypothesis 4. These reductions are surprising, because the incentives for exercising care to prevent violations are unchanged by the introduction of a voluntary disclosure policy if one does not intend to disclose a violation. Hence, these reductions suggest a framing effect that is due to the introduction of a voluntary disclosure option that, for some reason, motivated non-disclosers to choose lower violation probabilities.

Policy Implications

In the absence of any framing effects, all of the policy implications follow directly from the behavioral hypotheses. However, the framing effect (Hypothesis 4) is a potentially complicating factor, so we need to determine whether the policy implications hold despite this effect.

Policy Implication 1 asserts that the expected number of violations under treatment D-I should not be different from that number under CE. This is precisely what we observe despite the framing effect for those who chose not to disclose. Table 3 shows that the mean violation probability (and therefore the expected number of violations) under CE (0.509) is about the same as the mean violation probability for both disclosers and non-disclosers under D-I (0.497), and this difference is not statistically significant ($p = 0.23$). This implies that the expected numbers of violations in these treatments are not significantly different (91.6 under CE vs. 89.5 under D-I). Thus, it appears that it is possible to add a disclosure opportunity to a conventional enforcement strategy without affecting deterrence—at least as long as there are no discovery costs.

However, when subjects incur a cost to discover whether a violation occurred under CD-I, the expected number of violations should be higher than under CE (Policy Implication 2) because those that disclose their violations increase their violation probabilities (Hypothesis 3) while those that do not disclose should choose the same violation probabilities (Hypothesis 4). We have already seen that the disclosers in CD-I significantly increased their violation probabilities by 0.11, on average, from their violation probabilities under CE. However, the framing effect we have identified led non-disclosers to reduce their violation probabilities by 0.065 on average. Overall, the mean violation probability under CD-I is a bit higher than under CE (+0.027), but

this difference is not statistically significant ($p = 0.93$). Thus our data do not support Policy Implication 2.

Despite our failure to support Policy Implication 2, we have good reasons to continue to expect that disclosure policies when firms must invest in self-discovery would lead to less overall deterrence in non-laboratory settings. First, our view is that it is hard to justify a belief that the framing effect which led to our failure to support Policy Implication 2 would actually motivate firms in the field. Second, our results strongly support the hypothesis that inducing costly discovery and voluntary disclosure will lead to weaker deterrence for those who choose to discover and disclose. Thus, assuming that the framing effect for non-disclosers is unlikely to hold outside the lab, we believe that our results do justify a continued expectation of weaker deterrence when firms are given the opportunity to voluntarily discover and disclose their violations.

Policy Implication 3 asserts that if a voluntary disclosure policy motivates a significant number of violation disclosures, then fewer government audits are required to maintain the same level of deterrence for those that do not disclose. In fact, it is impossible for the expected number of audits to increase: as long as some firms disclose their violations, the remaining subset that are subject to random audits will necessarily be smaller than under CE. This is precisely what we observe; since a significant number of subjects voluntarily disclosed their violations under both D-I and CD-I (consistent with Hypothesis 1), as shown in Table 3, the expected number of audits required to maintain a 0.6 audit probability for those who did not report a violation under D-I (82.7) and CD-I (73.3) is significantly lower than under CE (108).

Although these results indicate that voluntary disclosure policies can lead to reduced monitoring effort, Policy Implication 4 suggests that the expected number of fines, and possibly sanctioning costs, could increase. This result necessarily holds if some violations are voluntarily disclosed (Hypothesis 1) and therefore automatically sanctioned, and if there is no change in violation probabilities of those who choose not to disclose (Hypothesis 4). Despite our observation that individuals who did not commit to disclosure tended to reduce their violation probabilities relative to conventional enforcement, as shown in Table 3, the expected total number of fines under D-I (70.5) and under CD-I (81.0) are significantly greater than the expected number of fines under CE (55.0).

While voluntary disclosure policies are likely to result in a greater number of sanctions, Policy Implication 5 asserts that the number of the potentially more costly fines for undisclosed violations will be smaller. Since a significant number of subjects chose not to disclose their violations in the D-I and CD-I treatments and there is only a small reduction in their mean violation probabilities in these treatments, it is not surprising that the results in Table 3 support this policy implication. Of the 70.5 expected fines under D-I, only 28.4 are for violations that are uncovered by a government audit. Similarly, under CD-I, only 23.1 of the 81.0 expected fines are for undisclosed violations. Both of these are significantly lower than the 55.0 expected fines under CE.

Sensitivity analysis

Our behavioral hypotheses, their policy implications, and hence, the discussion thus far has focused on the two treatments that were parameterized such that risk neutral agents would be indifferent about committing to disclosing their violations (D-I and CD-I). To examine the sensitivity of our results to the incentive for voluntary disclosure, we varied the reduced penalty for voluntarily reported violations.

Our results suggest that the commitment to disclose a violation is accompanied by a potentially substantial decrease in the care that individuals took to prevent violations. Table 3 shows that within each of the six treatments that allow voluntary violation disclosure, the mean violation probability (and corresponding expected number of violations) is higher for those who committed to disclosure than for those who did not (e.g., 0.656 vs. 0.525 for D-L). As one would expect, this difference is not statistically significant for the D-I treatment ($p = 0.14$ using a Mann-Whitney test for unmatched pairs).¹⁰ However, in four of the other five cases, this difference is significant at the one percent level: only for D-H is this difference not significant ($p=0.26$). Moreover, the mean violation probabilities of those who committed to disclosing their violations increases rather rapidly as the penalty for disclosed violations is reduced. For the Disclosure Only treatments, observe that the mean violation probability for those who committed to disclosure increases from 0.497 to 0.534 and then to 0.656 as the penalty for disclosed violations is reduced. A Mann-Whitney test making the pairwise comparison of treatments D-H and D-L indicates that this increasing trend is statistically significant ($p = 0.00$).¹¹ A similar pattern occurs in the Costly Discovery treatments.¹² Since the number of subjects who committed to disclose their violations also increases as the disclosure penalty is reduced, the expected number of violations of those who committed to disclosure also increases quickly. For the

Disclosure Only treatments, these values increase from 18.9 for D-H to 42.2 for D-I, and on to 76.8 for D-L. The same pattern holds for the Costly Discovery treatments.

Largely because the number of subjects who committed to disclosing their violations and their violation probabilities both increase as the disclosure penalty is reduced, the expected total number of violations also increases. From Table 3, the mean violation probability for disclosers and non-disclosers combined increases in the Disclosure Only treatments from 0.448 to 0.497 and then to 0.611 as the disclosure penalty is reduced. Again, this overall trend is highly significant: the p -value for a Mann-Whitney test comparing D-H to D-L is 0.00.¹³ Perhaps more revealing is the resulting increase in the expected number of violations from 80.5 to 89.5 and then to 109.8 for the Disclosure Only treatments. These patterns hold for the Costly Discovery treatments as well.¹⁴

While the expected number of violations increases as the disclosure penalty is reduced, the expected number of required audits falls—from 96.7 to 82.7 and then to 61.9 for the Disclosure Only treatments and similarly for the Costly Discovery treatments. Our results clearly suggest an important trade-off inherent in voluntary discovery and disclosure policies: reducing the penalty that firms automatically pay if they voluntarily disclose a violation effectively induces more of them to report their violations, thereby conserving monitoring effort, but these lower penalties also provide an incentive for firms to exercise less care in avoiding violations.

Of course, whether the reduction in monitoring effort results in a reduction in total enforcement costs also depends upon the impact on sanctioning costs. The results in Table 3 show that

reducing the penalty for disclosed violations increases the total number of fines in the Disclosure Only treatments from 55.9 to 70.5 and then to 96.6. The same pattern holds for the Costly Discovery treatments. However, the number of fines levied on undisclosed violations falls as the disclosure penalty is reduced: 37 to 28.4 and then to 19.8 in the Disclosure Only treatments, with a similar pattern holding for the Costly Discovery treatments. Thus, although lower disclosure penalties led to more penalties being levied, fewer penalties were levied on undisclosed violations. Clearly, how total enforcement costs change with a greater incentive for voluntary discovery and disclosure depends on the relative costs of monitoring and sanctioning disclosed and undisclosed violations.

CONCLUSIONS

A key conclusion of our study is that, when firms know their compliance status without cost, it is possible to motivate a significant number of voluntary violation disclosures without adversely affecting environmental quality. In this case, whether voluntary disclosure policies are worthwhile depends solely on their impact on government enforcement costs. While voluntary disclosure policies can reduce government efforts to monitor the compliance behavior of firms, their impact on the costs of sanctioning noncompliant firms depends on the relative costs of sanctioning voluntarily disclosed violations and sanctioning violations that the government uncovers.

However, when firms must undertake costly self-audits to determine their compliance status, we doubt that adding a voluntary disclosure policy to an existing conventional enforcement strategy

will leave deterrence unaffected. Although our results in this case fail to provide unequivocal support for the hypothesis that disclosure policies will lead to more violations when firms' self-audits are costly, this failure is due solely to a framing effect that we doubt would persist in field settings of environmental enforcement.

Generally, our work highlights some of the essential tradeoffs inherent in voluntary discovery and disclosure policies. Namely, motivating an increasing number of violation disclosures is associated with increasing incidences of noncompliance, worsening environmental quality, decreasing government monitoring effort, and more sanctions, fewer of which are for violations that the government uncovers. Consequently, it appears that there is little theoretical or empirical justification to warrant general support for, or opposition to, voluntary discovery and disclosure policies. Moreover, any conclusion about the benefits and costs of voluntary disclosure policies will require detailed knowledge of the harm caused by environmental violations, the costs of monitoring firms and sanctioning violations, as well as firms' costs of auditing themselves to determine their compliance status. Clearly, whether voluntary discovery and disclosure policies are an efficiency-enhancing complement to conventional environmental enforcement will have to be determined on a case-by-case basis.

While we have examined many of the essential aspects of voluntary disclosure policies, there are others that we have not considered, but that can be addressed with straightforward modifications of our experimental designs. For example, our framework can easily be adapted to examine Innes' (1999 and 2001a) claims that there are additional benefits to voluntary discovery and disclosure policies when firms are required to undertake costly remediation (*e.g.*, clean-up of

spills), or when they are able to engage in costly efforts to avoid government detection and punishment of their violations. Likewise, the claims of several authors that self-discovery and disclosure rules might not be as effective as hoped because firms fear that the information they discover might improve the government's own monitoring efforts (Pfaff and Sanchirico, 2000; Mishra *et al.*, 1997) can and should be examined within our framework.

Finally, while our study was motivated by voluntary discovery and disclosure policies to support compliance with environmental regulations, our results apply more broadly. Many environmental regulations require that firms report their compliance status to regulators. Although our experiments, and most of the literature on self-reporting in law enforcement, have focused on voluntary reporting, the tradeoffs that we highlight will also manifest themselves when reporting is mandatory. In addition, the use of disclosure policies extends well beyond environmental policies to regulations concerning occupational health and safety, product safety, and federal sentencing guidelines (Kaplow and Shavell, 1994; Innes, 2001b). Our results apply to these contexts as well.

ACKNOWLEDGEMENTS

Primary funding for this research was provided by the U.S. EPA – Science to Achieve Results (STAR) Program grant #R829608. Additional support was provided by the Cooperative State Research Extension, Education Service, U. S. Department of Agriculture, Massachusetts Agricultural Experiment Station, and the Department of Resource Economics under Project No. MAS00871, and by the Center for Public Policy and Administration, University of Massachusetts-Amherst. Maria Alejandra Velez and Elizabeth Gonzalez provided outstanding

research assistance. We are also grateful to Jay Shimshack, Jason Shogren, and an anonymous referee for useful comments and suggestions.

ENDNOTES

- ¹ U.S. EPA (2000). We follow the terminology used by the EPA. “Discovery” refers to costly efforts by regulated entities to discover whether they are in violation of an environmental regulation. Some also call these actions self-audits. “Disclosure” means voluntary reporting of violations to the authorities. In the related economics literature this is usually referred to as self-reporting.
- ² Innes (1999) has argued that the fact that voluntary disclosure policies will tend to lead to a greater number of sanctioned violations could also imply improved environmental quality if firms are required to correct the harm caused by their violations.
- ³ Innes (2001b) provides a comprehensive review of the theoretical literature on voluntary discovery and disclosure policies. The model presented here is essentially the same as the one he employs to motivate his review.
- ⁴ Thus, conventional enforcement is held fixed throughout this paper. Moreover, we do not consider the optimal design of voluntary discovery and disclosure rules, choosing instead to focus on the qualitative affects of adding various disclosure rules to an existing, fixed conventional enforcement strategy. For the design of optimal discovery and disclosure rules see Innes (2001b).
- ⁵ Unfortunately there is no consensus about how to elicit these preferences. Instruments such as that presented by Holt and Laury (2002) may be useful, but there is evidence that risk preferences may be domain specific and not stable across institutions (Isaac and James, 2000). Therefore it is unclear whether risk preferences elicited with the Holt/Laury mechanism would be robust in predicting behavior in other settings. We believe that this is an important area for future research for those who investigate compliance behavior in experimental settings.
- ⁶ To avoid the possibility of introducing unwanted biases, we framed the experiments as a production decision in which subjects chose the probability of an unspecified accident, instead of the probability of a violation of an environmental standard. We will continue to speak of violations throughout the paper, even though the subjects’ actions were about preventing and possibly disclosing accidents.
- ⁷ Instructions are available upon request.
- ⁸ For each hypothesis, a matched pair t-test yielded the same conclusions.
- ⁹ Note that in Table 3 the mean violation probability under CE for all 180 subjects is 0.509. The mean violation probability under CE for the 79 subjects who committed to disclosure under D-I is 0.518.
- ¹⁰ This follows simply by combining Hypotheses 2 and 3.
- ¹¹ A pairwise comparison of the violation probabilities under D-H and D-I for disclosers only is not significant ($p = 0.50$), but a comparison of these values under D-I and D-L is significant ($p = 0.00$).
- ¹² The results of Mann-Whitney tests for the violation probabilities of disclosers only are as follows: CD-H vs. CD-L ($p = 0.00$), CD-H vs. CD-I ($p = 0.12$), CD-I vs. CD-L ($p = 0.02$).
- ¹³ A pairwise comparison of the violation probabilities under D-H and D-I for all subjects is significant ($p = 0.05$), as is a comparison of these probabilities under D-I and D-L ($p = 0.00$).
- ¹⁴ The results of Mann-Whitney tests for the violation probabilities of all subjects are as follows: CD-H vs. CD-L ($p = 0.02$), CD-H vs. CD-I ($p = 0.00$), CD-I vs. CD-L ($p = 0.00$).

REFERENCES

- Holt, C.A. and Laury, S.K., 2002. Risk Aversion and Incentive Effects. *American Economic Review*, 92 (5), 1644-55.
- Innes, R., 1999. Remediation and Self-Reporting in Optimal Law Enforcement. *Journal of Public Economics*, 72 (3), 379-393.
- Innes, R., 2001a. Violator Avoidance Activities and Self-Reporting in Optimal Law Enforcement. *Journal of Law, Economics, and Organization*, 17 (1), 239-256.
- Innes, R., 2001b. Self Enforcement of Environmental Law. In: A. Heyes, ed. *The Law & Economics of the Environment*. Cheltenham, England: Edward Elgar.
- Isaac, R.M. and James, D., 2000. Just Who Are You Calling Risk Averse? *Journal of Risk and Uncertainty*, 20 (2), 177-187.
- Kaplow, L. and Shavell, S., 1994. Optimal Law Enforcement with Self-Reporting of Behavior. *Journal of Political Economy*, 103 (3), 583-606.
- Malik, A., 1993. Self-Reporting and the Design of Policies for Regulating Stochastic Pollution. *Journal of Environmental Economics and Management*, 24 (3), 241-257.
- Mishra, B.K., Newman, D.P. and Stinson, C.H., 1997. Environmental Regulations and Incentives for Compliance Audits. *Journal of Accounting and Public Policy*, 16 (2), 187-214.
- Pfaff, A. and Sanchirico, C.W., 2000. Environmental Self-Auditing: Setting the Proper Incentives for Discovery and Correction of Environmental Harm. *Journal of Law, Economics & Organization*, 16 (1), 189-208.
- Pfaff, A. and Sanchirico, C.W., 2004. Big Field, Small Potatoes: An Empirical Assessment of EPA's Self-Audit Policy. *Journal of Policy Analysis & Management*, 23 (3), 415-432.
- Stafford, S.L., 2005. Does Self-Policing Help the Environment? EPA's Audit Policy and Hazardous Waste Compliance. *Vermont Journal of Environmental Law*, Volume 6. Available at, <http://www.vjel.org/articles/articles/Stafford11FIN.htm>.
- U.S. Environmental Protection Agency, 2000. Incentive for Self-Policing, Discovery, Disclosure, Correction and Prevention of Violations. *Federal Register*, 65 (70), 19618-19627. Available at <http://www.epa.gov/compliance/incentives/auditing/auditpolicy.html>.

Table 1: Experimental Design

Conventional Enforcement	Voluntary Disclosure Only	Voluntary Disclosure with Costly Discovery
	D-H (\$2.35)	CD-H (\$1.50)
CE	D-I (\$1.50)	CD-I (\$0.97)
	D-L (\$0.97)	CD-L (\$0.60)

The conventional enforcement penalty in all treatments is \$2.50. The reduced penalty, ϕ_d , for voluntary disclosure is shown in parentheses.

Table 2: Sequence of Treatments Using a Latin Square

Sequence ID	Stage						
	1	2	3	4	5	6	7
A	CE	D-H	D-I	CD-L	D-L	CD-I	CD-H
B	CE	D-I	D-L	D-H	CD-H	CD-L	CD-I
C	CE	D-L	CD-H	D-I	CD-I	D-H	CD-L
D	CE	CD-H	CD-I	D-L	CD-L	D-I	D-H
E	CE	CD-I	CD-L	CD-H	D-H	D-L	D-I
F	CE	CD-L	D-H	CD-I	D-I	CD-H	D-L

Table 3: Mean Violation Probabilities, Expected Numbers of Violations, and Expected Numbers of Enforcement Actions

Treatment (Disclosure Penalty)	Mean Violation Probability	N	Expected Number of Violations	Expected Number of Audits	Expected Number of Fines
CE	0.509	180	91.6	108.0	55.0
D-H (\$2.35)	0.448	180	80.5	96.7	55.9
Disclose	0.497	38	18.9		18.9
Not Disclose	0.434	142	61.6		37.0
D-I (\$1.50)	0.497	180	89.5	82.7	70.5
Disclose	0.534	79	42.2		42.2
Not Disclose	0.468	101	47.3		28.4
D-L (\$0.97)	0.611	180	109.8	61.9	96.6
Disclose	0.656	117	76.8		76.8
Not Disclose	0.525	63	33.1		19.8
CD-H (\$1.50)	0.467	180	84.1	86.3	64.9
Disclose	0.547	66	36.1		36.1
Not Disclose	0.421	114	48.0		28.8
CD-I (\$0.97)	0.536	180	96.4	73.3	81.0
Disclose	0.616	94	57.9		57.9
Not Disclose	0.448	86	38.5		23.1
CD-L (\$0.60)	0.650	180	117.0	51.6	107.8
Disclose	0.701	134	93.9		93.9
Not Disclose	0.502	46	23.1		13.9