# AN EXPERIMENTAL STUDY OF COMPLIANCE AND LEVERAGE IN AUDITING AND REGULATORY ENFORCEMENT

TIMOTHY N. CASON and LATA GANGADHARAN*

*Evidence suggests that individuals often comply with regulations even though the frequency of inspections and audits is low. We report a laboratory experiment based on the dynamic model suggested by Harrington (1988) to explain this puzzle in which participants move between two inspection groups that differ in the probability of inspection and severity of fine. Enforcement leverage arises in the Harrington model from movement between the groups based on previous observed compliance and noncompliance. We find that compliance behavior does not change as sharply as the model predicts. A simple model of bounded rationality explains these deviations from optimal behavior.* (*JEL* C91, Q20, Q28)

## I. INTRODUCTION

Regulatory policymakers have observed that many firms and individuals comply with regulations even when both the frequency of audits and the penalty for violations are low. This is seen in areas as diverse as income tax collection, customs, antitrust laws, health and safety, and environmental regulation. This phenomenon is difficult to explain using static enforcement models—for example, see Linder and McBride (1984), Storey and McCabe (1980), and Harford (1978)—in which the penalty facing the firm depends only on the firm's performance in the current period and not on its previous compliance record.

Economists in recent years have proposed dynamic repeated game models to reconcile the low expected penalties and high observed compliance rates. In these models, the regulated firm and the enforcement agency can react to previous actions by the other—see Landsberger and Meilijson (1982), Greenberg (1984), and Harrington (1988). The enforcement agency alters the expected penalty and the inspection frequency based on the firm's past performance. Harrington finds that a firm could have an incentive to comply with regulations even though the costs of compliance in individual periods exceed the expected penalty for violation. This is important in practice because political or practical considerations often limit the size of the fine that can be imposed on a firm. For example, in many states there is a restriction on the size of penalties that can be levied for violating an environmental regulation (e.g., \$5,000 per day).[1]

The strategy that the enforcement agency uses to achieve this result divides the firms into two groups: the firms in one group face a more severe enforcement regime than do the firms in

*We thank Ashraf Al Zaman for valuable research assistance. For helpful comments we also thank, without implicating, seminar and conference participants at the Public Choice Meetings/ESA Meetings, and McMaster, Purdue and Lakehead Universities, as well as a co-editor and an anonymous referee. This research has been supported by a grant from the U.S. Environmental Protection Agency's National Center for Environmental Research (NCER) Science to Achieve Results (STAR) program. Although the research described in the article has been funded in part through EPA grant number R829609, it has not been subjected to any EPA review and therefore does not necessarily reflect the views of the Agency, and no official endorsement should be inferred.

*Cason:* Department of Economics, Krannert School of Management, 100 S. Grant Street, Purdue University, West Lafayette, IN 47907-2076. Phone 1-765-494-1737, Fax 1-765-494 9658. E-mail cason@mgmt.purdue.edu

*Gangadharan:* Department of Economics, University of Melbourne, Vic 3010, Australia. Phone +61 3 8344 5408, Fax +61 3 8344 6899. E-mail latag@unimelb.edu.au

1. Compliance can occur for other reasons, of course. Firms may comply with regulations to guide regulatory authorities to set higher standards for the whole industry, thereby increasing the costs of their rivals; see Salop and Scheffman (1983). Firms may also comply to obtain a reputation of being an environmentally conscious organization. Arora and Gangopadhyay (1995) show that public recognition plays an important role in the success of voluntary environmental programs. Individuals and firms may also comply because they are honest and get disutility from violating regulations.



Economic Inquiry
(ISSN 0095-2583)
Vol. 44, No. 2, April 2006, 352–366

doi:10.1093/ei/cbj019
Advance Access publication November 16, 2005

the other group. A firm's compliance status determines which group it is in. Each firm can move from one group to the other depending on its performance. Violations discovered in the rarely inspected "good" group are punished by transfer into the more frequently inspected "bad" group, and compliance discovered in the more frequently inspected group is rewarded with the chance of a return to the rarely inspected good group. This enforcement scheme poses a Markov decision problem from the firm's perspective. The firm moves from group to group according to transition probabilities that depend not only on the inspection probabilities and the current state of the system but also on the action taken (comply or not) during that period. Harrington shows that firms' optimal policies in this scheme depend upon their individual costs of compliance. Low-cost firms are always in compliance; high-cost firms are never in compliance; and medium-cost firms move in and out of compliance depending on the results of recent inspections.

This article reports laboratory evidence on compliance behavior of decision makers when faced with enforcement conditions consistent with the Harrington model framework. We examine treatments in which the compliance costs are low, medium, or high. In these within-session treatments, we also change the probability of the firm's switching from the frequently inspected group to the rarely inspected group if inspected and found compliant from 10% to 90%. Our results indicate that, consistent with theoretical predictions, violation rates increase when compliance costs become higher and as the probability of switching groups becomes lower. Behavior does not change as sharply as the model predicts, however, given that violation rates do not jump from 0 to 1, as parameters vary across critical thresholds. A simple model of bounded rationality in which agents choose more profitable strategies with higher probability, but not with probability equal to one, can explain these deviations from optimal behavior.

Although these conditional audit rules have received significant attention in the theoretical literature, direct empirical evidence on their performance is scarce. Empirical research using field data exists—for example, Helland (1998), Oljaca et al. (1998), and Eckert (2004)—but it is hampered by the absence of reliable information regarding individual reporting behavior and unknown compliance decisions for uninspected firms.[2] Laboratory experiments, however, are well suited to study the different features of compliance schemes and individual behavior within these schemes. Most of the existing experimental literature on compliance and auditing has focused on static models, where different policy changes—such as an increase in tax rate, a change in penalty rates, tax amnesties, or changes in audit probabilities—are introduced to determine the impact on compliance behavior. Alm and McKee (1998) provide a survey of this literature. Torgler (2002) surveys the experimental findings on the tax compliance literature with a focus on social norms and institutional factors, which are seen to encourage compliance.

Alm et al. (1993) examine dynamic audit rules and compare these to a 5% inspection probability random audit rule. The auditor's discovery of noncompliant behavior in a random audit scheme could lead to audits of previous or future years with certainty. The authors find that the forward-looking rules achieve lower compliance rates because, in this scheme, an person can cheat until audited in the current period and can then avoid any additional penalties by reporting honestly for the next two periods. On the other hand, under the backward-looking audit policy, an individual found to be noncompliant in the current period has no chance of avoiding penalties on previous periods' records. This increases the incentive for people to comply under backward-looking policies and might be more attractive from the viewpoint of regulators, particularly in the area of tax reporting. Backward-looking schemes, however, would not typically be feasible in others kinds of regulatory areas, such as environmental and natural resource management when the data (e.g., for actual emissions rates) from previous periods cannot be checked. Therefore, forward-looking conditional audit rules such

2. Helland (1998) uses data from the American pulp and paper industry to test whether environmental regulators audit and fine according to the model described in Harrington (1988). He finds that firms who are discovered in violation experience a one- or two-quarter penalty period, during which they are inspected more frequently. Eckert's data (2004) on Canadian petroleum storage facilities are also consistent with the Harrington framework. The author finds that inspections deter future violations, although the effect is small.

**TABLE 1**
Payoff Parameters for Enforcement Game

| | Group 1 | | Group 2 | |
|---|---|---|---|---|
| | **Comply** | **Violate** | **Comply** | **Violate** |
| Inspection probability | $p_1 = 0.2$ | | $p_2 = 0.5$ | |
| No inspection | $c = 100, 200, 375$ (baseline $c = 7$) | 0 | $c = 100, 200, 375$ (baseline $c = 7$) | 0 |
| Inspection | $c = 100, 200, 375$ | $F_1 = 50$ Moved to $G_2$ with prob = 1 | $c = 100, 200, 375$ Prob (moved back to $G_1$) = $u = 0.1, 0.9$ | $F_2 = 300$ |

as those studied here are practical for a wider range of applications.

The previous research most closely related to the present study is that of Clark et al. (2004), which compares two dynamic audit rules: Harrington's scheme (1988) and one proposed by Friesen (2003) that is designed to minimize the inspections that regulators must make to achieve a target rate of compliance. Both the rules use the current audit record of the firm to assign them to different audit groups in future periods, but in Friesen's scheme all of the transitions between audit groups can be probabilistic, whereas in Harrington's scheme all transitions are deterministic except for the movement of an inspected compliant firm from the bad group to the good group. In Friesen's optimal targeting scheme the firms face a fixed probability of moving from the good group to the bad group, which is independent of compliance status in the current period. There are no inspections conducted of firms in the first group. Clark et al. find an enforcement possibility frontier between compliance and minimizing inspections, with the Friesen rule requiring slightly lower inspection rates. Their experiment focuses on a comparison of the two conditional audit rules against simple random auditing for a single compliance cost and one set of enforcement parameters in each rule. Harrington's rule performs well on certain measures, such as for the rate of compliance per inspection. This suggests that further exploration of the performance of this enforcement policy is warranted, and in the present study we consider seven different enforcement parameter and compliance cost combinations to more fully examine its empirical properties. These multiple treatments allow us to study how compliance choices respond to different enforcement rules, and we estimate a boundedly rational choice model to characterize behavioral responses for this type of probabilistic enforcement.

## II. THEORETICAL FRAMEWORK

We are interested in the relationship between the firm's compliance cost, its compliance decisions, and the conditional audit scheme chosen by the regulator. Our experiment is structured by Harrington's model (1988), which determines for a two-state model the level of compliance that can be achieved when enforcement budgets and the maximum feasible penalty are limited. Let $G_1$ and $G_2$ denote the two inspection groups of firms and denote the inspection probability in $G_i$ as $p_i$ and the penalty for violation as $F_i$, with $p_1 < p_2$ and $F_1 < F_2$. Firms can avoid a violation by incurring the compliance cost $c$. If a firm is inspected, its compliance status is observed perfectly.[3] Firms found to be in violation in $G_1$ are punished by a transfer into $G_2$, and firms found to be in compliance in $G_2$ are rewarded with a chance of a return to $G_1$. The probability that a firm found in compliance in $G_2$ is returned to $G_1$ is denoted by $u$. Table 1 presents the payoffs to the firm in this game.

3. This is not a critical assumption for Harrington's two-state model, but it is for other models. Greenberg (1984) shows that a three-state model in which transitions out of a third, "habitual offender" group were impossible can dramatically reduce the rate of violations when compared to a two-state model. However, if false positives are possible (i.e, situations where compliant firms are wrongly classified as violators with some positive probability), every firm eventually moves into the third group and all firms are inspected every period.

**TABLE 2**
Expected Payoff of Alternative Policies

| Policy | Expected Payoff if in Group 1 | Expected Payoff in Group 2 |
|---|---|---|
| Always comply: $f_{00}$ | $\frac{R-c}{1-\beta}$ | $\frac{R-c}{1-\beta}$ |
| Comply only in Group 1: $f_{10}$ | $\frac{R}{1-\beta}-\frac{cp_1\beta+p_1F_1(1-\beta+p_2\mu\beta)}{(1-\beta)(1-\beta+p_1\beta+p_2\mu\beta)}$ | $\frac{R}{1-\beta}-\frac{c(1-(1-p_1)\beta)+\beta p_2\mu p_1F_1}{(1-\beta)(1-\beta+p_1\beta+p_2\mu\beta)}$ |
| Never comply: $f_{11}$ | $\frac{R}{1-\beta}-\frac{p_1F_1(1-\beta)+\beta p_1p_2F_2}{(1-\beta)(1-(1-p_1)\beta)}$ | $\frac{R-p_2F_2}{1-\beta}$ |

For future reference, this table also includes the parameters chosen for the experiment.

A policy for the firm is a map $f$: $\{1, 2\} \rightarrow (0, 1)$ of states 1 and 2 into decisions to comply with (0) or violate (1) the regulations. The firm's goal is to choose the policy that minimizes the present value of its expected costs over an infinite horizon. The firm has four available policies: $f_{00}$, $f_{01}$, $f_{10}$, and $f_{11}$, where $f_{00}$ is the policy that the firm would comply in states 1 and 2, and $f_{01}$ is the policy of complying when in $G_1$ and that of violating when in $G_2$ and so on. The expected present value of the policy would be the cost this period plus the expected present value discounted one period. This leads to four sets of simultaneous equations that can be solved to obtain the present values of the four policies. For example, the expected cost of policy $f_{10}$ in state 1 is

$$[cp_1\beta+p_1F_1(1-\beta+p_2u\beta)] \, / \, [(1-\beta)(1-\beta+p_1\beta+p_2u\beta)] \tag{1}$$

and the expected cost of $f_{10}$ in state 2 is

$$[c\,(1-(1-p_1)\beta)+\beta p_2up_1F_1] \, / \, [(1-\beta)(1-\beta+p_1\beta+p_2u\beta)], \tag{2}$$

where $\beta$ is the discount factor.

Harrington shows (his lemma 1) that in this framework, $f_{01}$ is never an optimal policy, as it is dominated by $f_{00}$ when the cost of compliance $c < p_2F_2$ and by $f_{11}$ when $c \geq p_2F_2$. Hence, the firm chooses between three policies—$f_{00}$, $f_{10}$, $f_{11}$—and the optimal policy depends on the compliance costs facing the firm and the enforcement parameters chosen by the regulatory agency. Table 2 presents the expected payoff for each policy, based on an exogenous per-period revenue of $R$. Firms with compliance costs below a particular threshold ($p_1F_1$) always comply, and those with costs above a higher threshold never comply. For an intermediate range of costs, the firm chooses policy $f_{10}$, and it cheats when in $G_1$ and complies in $G_2$. Ironically, for these intermediate compliance costs, the "good guys" in $G_1$ can afford to cheat, whereas the "bad guys" in $G_2$ comply until they are moved back into $G_1$. Compared to a static model, in this dynamic model compliance is achieved in $G_2$ even though the expected penalty is not large, because firms in $G_2$ may be allowed to return to $G_1$ depending on their compliance record.

The enforcement agency in this model wants to minimize the resources spent on monitoring and enforcement subject to achieving a target compliance rate $Z$. The agency has five parameters that can be changed to achieve desired compliance rates: the probability of inspections, $p_1$ and $p_2$; the two penalties, $F_1$ and $F_2$; and the probability $u$ of the firm's moving back into $G_1$ if found compliant. We manipulate $u$ as well as the compliance cost $c$ as exogenous treatment variables in the experiment. For certain parameters—specifically, the $u = 0.9$, compliance cost $=$ 200 treatment described later—firms have an incentive to comply even though the expected penalty ($p_2F_2 = 0.5 \times 300 = 150$) is less than the single-period compliance cost. This property is termed *leverage* in the literature.

In the optimal combination of enforcement parameters characterized by Harrington, marginal firms that adopt policy $f_{10}$ just slightly prefer to comply rather than violate in $G_2$. In our choice of parameters described in the next section, we avoid these optimal parameter cases in which individuals are nearly indifferent between two strategies. This design choice is guided by experience with previous experiments that demonstrate that more than marginal incentives are necessary for participants to learn optimal behavior. This is confirmed by the noisy choice model results reported in section IV.

## III. EXPERIMENTAL DESIGN

We conducted 13 sessions with 8 or 9 participants in each session, all 114 of whom were undergraduate students at Purdue University and inexperienced in the sense that they had not participated in a similar experiment. The University of Zurich's *z*-tree program was employed to conduct all sessions; see Fischbacher (1999). Each session lasted about 45 minutes, including instruction time. Payoffs in the experiment were converted using an exchange rate of 1,500 experimental dollars = 1 U.S. dollar, and participant earnings ranged from \$6.75 to \$15.25, with median earnings of \$12.75. These sessions constituted the first half of a longer session that trained participants to make compliance choices in a study of emissions permit trading with imperfect enforcement—see Cason and Gangadharan (forthcoming). Each participant made 61 separate compliance choices over seven different period sequences, one for each treatment variable combination.[4]

At the start of each period sequence, participants were randomly assigned into inspection group 1 or 2, which differ in the probability of inspections and severity of fine. Participants had a binary choice: comply or violate in each period. If they decided to comply, they paid a compliance cost, which remained unchanged within a period sequence but varied across period sequences. Participants were inspected with a certain probability that depended on which group they were in. Group 1 participants were inspected with a probability of 20%, and those of group 2 were inspected with a probability of 50%. Participants were required to pay a fine if they did not comply in a particular period, and they were inspected. The fine for violation was 50 experimental dollars in group 1 and 300 experimental dollars in group 2. In addition, participants in group 1 were moved to group 2 when they were caught violating. If participants were in group 2 and were observed to comply on inspection, they were then moved back into group 1 with a low or high probability. The instructions were framed using the terminology of this paragraph (i.e., *comply*, *violate*, *inspection*, *fine*, and so on). Comparison of our results with the more neutrally framed terminology employed in Clark et al. (2004) suggests that framing does not have a substantial impact on the results.[5]

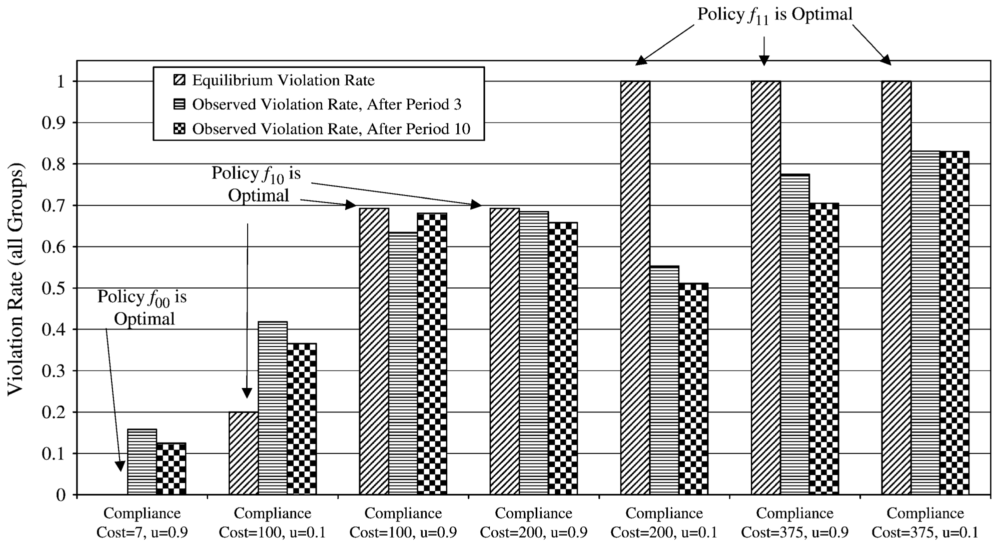
Each participant participated in a random number of periods in seven separate period sequences. The number of periods in each period sequence was determined before the session and was unknown to the participants. Participants in the same session faced the different treatments in different orders, which implies that our treatment comparisons control for sequencing effects. The random ordering also leads to an approximately equal number of decisions in each treatment. As explained in the instructions, for each period there was a 90% chance that the same period sequence continued for an additional period. This implements a discount factor $\beta = 0.9$. Participants were told only at the end of the last period in a sequence that a new period sequence would begin.

The period sequences were a combination of two treatment variables, both varied within sessions, in a three-by-two factorial design. For one treatment variable we varied the compliance costs ($c$) across three levels from low to medium to high to determine whether participants changed their compliance decisions in the presence of different levels of compliance costs. The compliance costs were 100 in the low-cost scenario, 200 in the medium cost, and 375 in the high cost. For the other treatment variable, we manipulated at two levels the probability ($u$) of participants moving from group 2 to group 1 in order to determine whether they complied more when the probability of switching groups was higher. Participants faced a switching probability of 0.1 in some period sequences and 0.9 in others. As noted, these enforcement parameters do not represent the "optimal" parameters derived in the Harrington model; instead, they reflect our design goal to explore

4. Experiment instructions are available at http://www.krannert.purdue.edu/faculty/cason/papers/leverage-instr.pdf.

5. For example, Clark et al. (2004) use *Option A* and *Option B* instead of *comply* and *violate*. Our leverage treatment with compliance cost = 200 and $u = 0.9$ is most similar to the one treatment Clark et al. study, in that violation is optimal in group 1 and compliance is optimal in group 2, even though the compliance cost exceeds the expected fine in group 2. Clark et al. observe overall compliance rates of 12% in group 1 and 75% in group 2, whereas in our similar treatment we observe overall compliance rates of 11% in group 1 and 63% in group 2. Though obviously not identical, these rates are similar—especially when considering the many other procedural, training, and payment design differences between our experiment and that of Clark et al.

**FIGURE 1**
Predicted and Observed Overall Violation Rates, by Treatment, for Two Sets of Later Periods




a variety of compliance conditions with strong and weak incentives to comply or violate in the different inspection groups. We also employed a seventh period sequence that served as a baseline with very low compliance costs (7) and $u = 0.9$, for which compliance is always optimal. All participants made compliance decisions in all treatment variable combinations.

Although quite stylized, several features of the experimental design increased its parallelism with the field or were chosen specifically to explore the range of possible behavioral responses to a variety of enforcement conditions. As already noted, we employed natural, nonneutral terminology in which participants chose to *violate* or *comply*. Second, participants made individual rather than group compliance decisions. This may have limited the range of application of the behavioral results, because some decisions in response to regulations are made by groups; however, a large proportion are made by individuals in the field, including many individual decisions when reporting personal taxable income. Third, participants were exposed to different compliance costs in different period sequences, helping us determine how individual behavior changes with changes in the costs. Fourth, participants were moved from the low-intensity audit group to the high-intensity group with different switching probabilities. This dynamic audit rule is similar to what often happens in income tax auditing and health and environmental auditing. Moreover, as already noted, the forward-looking conditional audit rule that we study can apply in cases where past compliance cannot be assessed, which makes it relevant and applicable for a more general and broader class of regulatory issues.

## IV. RESULTS

### *Overall Violation Rates*

Figure 1 presents the average violation rate for later periods in the period sequences, along with the steady state predicted violation rates, for each of the seven treatments. The predicted violation rate is 0 when compliance policy $f_{00}$ is optimal and 1 when compliance policy $f_{11}$ is optimal. When policy $f_{10}$ is optimal, the predicted violation rate is the stationary probability of being in inspection group 1, $p_2u/(p_1 + p_2u)$. The figure shows that violations usually increase when they are predicted to increase but that they do not reach the corner solution rate of 0 or 1 when policies $f_{00}$ or $f_{11}$ are optimal.

Tables 3 and 4 present the overall violation rates separately for the compliance cost ($c$), switching probability ($u$), and inspection group combinations. The model predicts that

**TABLE 3**
Predicted and Observed Violation Rates for Inspection Group 1

| Probability an Inspected, Compliant Firm Exits Group 2 | | Compliance Cost = 7 | Compliance Cost = 100 | Compliance Cost = 200 | Compliance Cost = 375 |
|---|---|---|---|---|---|
| $u = 0.1$ | Observed violation rate | | 371/511 = 73% (158/206 = 77%) | 210/243 = 86% (59/68 = 87%) | 198/221 = 90% (47/50 = 94%) |
| | Predicted violation rate | | 1 | 1 | 1 |
| $u = 0.9$ | Observed violation rate | 136/795 = 17% (69/395 = 17%) | 506/607 = 83% (151/177 = 85%) | 538/603 = 89% (210/218 = 96%) | 502/539 = 93% (277/276 = 97%) |
| | Predicted violation rate | 0 | 1 | 1 | 1 |

*Note:* Data for late sequences (5–7) are shown in parentheses.

**TABLE 4**
Predicted and Observed Violation Rates for Inspection Group 2

| Probability an Inspected, Compliant Firm Exits Group 2 | | Compliance Cost = 7 | Compliance Cost = 100 | Compliance Cost = 200 | Compliance Cost = 375 |
|---|---|---|---|---|---|
| $u = 0.1$ | Observed violation rate | | 116/526 = 22% (54/217 = 25%) | 359/732 = 49% (115/244 = 47%) | 529/655 = 81% (229/262 = 87%) |
| | Predicted violation rate | | 0 | 1 | 1 |
| $u = 0.9$ | Observed violation rate | 28/188 = 15% (21/130 = 16%) | 65/359 = 18% (20/135 = 15%) | 138/369 = 37% (45/154 = 29%) | 255/399 = 64% (128/204 = 63%) |
| | Predicted violation rate | 0 | 0 | 0 | 1 |

*Note:* Data for late sequences (5–7) are shown in parentheses.

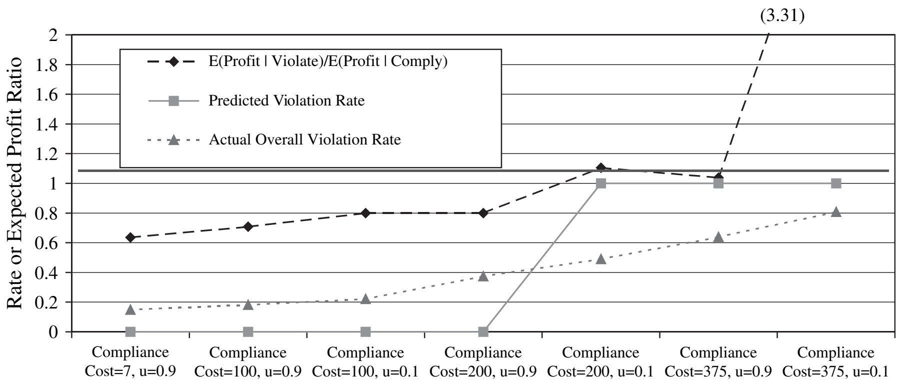
for our experimental parameters, participants will violate whenever they are in inspection group 1 except for the baseline treatment with a very low compliance cost of 7. Table 3 shows that this prediction is broadly supported, with observed violation rates of participants in group 1 between 73% and 93% when violation is predicted. These rates typically increase for the later sequences (5–7) when participants have more experience across treatments, as shown in parentheses in the table. The violation rate is 17% in the baseline treatment, for which violation is not predicted.

For the parameters employed in the experiment, the model predicts violation in only three of the seven treatment cells when participants are in inspection group 2. Participants should not violate in the low-compliance cost cases (7 and 100) and should violate in the high-compliance cost case (375), irrespective of the value of the switching probability $u$. In the medium-compliance cost case (200), they should violate only when they are unlikely to escape from inspection group 2 ($u = 0.1$). Table 4 indicates that violations are more common when they are predicted, but that in all seven cases the violation rates differ from the predicted rates by at least 13 percentage points, even when considering only the late sequences, 5–7. Violations also rise when moving to the right or upward in Table 4.[6]

Table 4 clearly shows that participants do not dramatically switch from never violating to always violating when the expected return

6. These deviations from the optimal compliance choices are not explained by participants who were "chronic" violators or who may have derived utility from being "honest" and complied constantly even when violation is more profitable. Indeed, we find no evidence that such extreme behaviors were present, given our analysis of participants' individual play. All participants violated at least one-third of the time and complied at least 15% of the time. As shown, most individual participants' compliance choices changed in response to the different incentives generated by the different enforcement treatments.

**FIGURE 2**
Predicted and Actual Violation Rates when in Inspection Group 2




from violating exceeds the expected return from compliance. Figure 2 illustrates the contrast between the sharp never/always violate prediction of the model and the smoothly monotonically increasing violation rate observed in the experiment. This figure is based on choices in inspection group 2 only, and it displays the ratio of expected profits from violating to the expected profits from compliance. These expected profits are based on the discounted, infinitely repeated compliance choice problem with the optimal compliance policy followed in all subsequent periods. The model predicts a violation rate of 1 if and only if this ratio exceeds 1. The observed violation rate, however, is merely higher whenever this ratio indicates a higher return to violation. (An exception to this occurs for one transition: compliance cost $=$ 200 and $u = 0.1$ to compliance cost $=$ 375 and $u = 0.9$.) In other words, participants' choices appear to be sensitive to the relative payoffs from violation and compliance, but the overall averages do not switch from one corner solution prediction to the other at the sharp threshold when the ratio passes through 1. We return to this issue later in this section, where we explain this behavior using a simple model of boundedly rational, or "noisy," decision making.

The data do not support the point predictions of the model, but they are consistent with many of the comparative static predictions about how the compliance rates differ in the various treatment cells. For formal tests we did not use the overall averages displayed in Tables 3 and 4, because participants made multiple compliance choices; therefore, the data points in this table are not statistically independent. Fortunately, we could conduct rather powerful tests based on statistically independent observations of participants' compliance rates and compliance rate differences across treatment cells. Recall that 114 people participated in this study and that they did not interact at all so that each provides statistically independent observations. For instance, to test whether the violation rate in inspection group 2 for $u = 0.9$ is significantly higher when compliance cost $=$ 375 than when compliance cost $=$ 200, we first calculated the violation rate for each participant within those two treatment cells. We then calculated the difference in these rates for the 70 participants who made choices in both treatment cells and employed a nonparametric Wilcoxon signed-rank test to determine whether these differences were significantly different from zero. The test statistics for all the comparisons are presented in Tables A-1, A-2, and A-3, in the appendix.

This statistically conservative and yet powerful procedure (due to our sample size) yields the following conclusions. All statements are based on a 5% significance threshold.

First, violation rates are significantly higher when in inspection group 1 than when in inspection group 2 for all seven treatment cells. Note that the model predicts a significant difference in only three of the treatment cells (i.e., for both $u = 0.1$ and $u = 0.9$ when compliance cost $=$ 100 and when $u = 0.9$ and compliance cost $=$ 200).

Second, when in inspection group 1 the violation rate increases significantly when the compliance cost increases in three of the five pairwise comparisons: for $u = 0.1$ when moving from compliance cost $= 100$ to 200 and for $u = 0.9$ when moving from compliance cost $= 7$ to 100 and when moving from 100 to 200. When in inspection group 2 the violation rate increases significantly when the compliance cost increases in four of the five pairwise comparisons: all cases except for $u = 0.9$ when moving from compliance cost $= 7$ to 100. Note that the data support all three compliance cost treatment effects predicted by the model (for $u = 0.1$ when moving from compliance cost $= 100$ to 200 in inspection group 2 and for $u = 0.9$ when moving from compliance cost $= 200$ to 375 in inspection group 2 and when moving from compliance cost $= 7$ to 100 in inspection group 1). However, note that four additional differences are also significant (for $u = 0.1$ when moving from compliance cost $= 100$ to 200 in inspection group 1 and from compliance cost $=$ 200 to 375 in inspection group 2 and for $u = 0.9$ when moving from compliance cost $= 100$ to 200 in both inspection groups).

Third, the violation rate is significantly higher when $u = 0.1$ than when $u = 0.9$ for all three pairwise comparisons when participants are in inspection group 2. This is predicted only for the medium compliance cost $= 200$ case, where the leverage of the two inspection groups is greatest. The violation rate is not significantly different for any of the three pairwise $u$ comparisons when participants are in inspection group 1, as predicted by the model.

These statistical conclusions generally hold for alternative subsets of the data, including for compliance choices based on only the initial inspection group that participants were randomly assigned to or for compliance rates based only on participants who have at least three compliance choices for a particular treatment cell. They are also robust to alternative statistical tests such as a simple nonparametric sign test or the standard parametric $t$ test.

### *Classification of Strategies*

The violation rates just analyzed separately for each compliance cost ($c$), switching probability ($u$), and inspection group combination employ a state-by-state perspective of this choice problem that differs from the strategy specification of Harrington's model. Recall that agents in the model adopt an entire compliance policy; for example, if they adopt strategy $f_{10}$, they violate when in inspection group 1 and comply when in inspection group 2. Therefore, in this section we examine the entire sequence of compliance choices within treatment cells to classify individual participants' compliance policies. The main difficulty that we encounter in this classification is that some participants did not make choices in both inspection groups so that their observed choices are consistent with multiple policies. Table 5 presents the classification for only those participants who can be perfectly classified into a specific strategy for a particular treatment, and Table 6 classifies every participant based on his or her "best-fitting" strategy.

Table 5 classifies a participant as choosing compliance policy $f_{11}$ for a particular treatment cell if he or she always violated in that cell, regardless of which inspection group one was in. We classify a participant in compliance policy $f_{10}$ for a cell if one always violated when in inspection group 1 and never violated when in inspection group 2. The classifications for $f_{01}$ and $f_{00}$ are defined analogously. Some participants never made choices in one of the inspection groups for some treatment cells, so we have no data to classify their behavior in that group. These cases are denoted with question marks. For example, $f_{1?}$ indicates that a participant always violated in inspection group 1 but never made decisions in inspection group 2. This person's behavior is consistent with both $f_{10}$ and $f_{11}$. In the summary sections in Table 5 we count observations as being consistent with $f_{10}$—for example, if they are identified in the "frequency (rate)" section as $f_{1?}$, $\mathrm{f}_{?0}$, or $f_{10}$. Likewise, we count observations as being consistent with $f_{11}$ if they are identified in the "frequency (rate)" section as $f_{1?}$, $\mathrm{f}_{?1}$, or $f_{11}$, and we count observations as being consistent with $f_{00}$ if they are identified in the "frequency (rate)" section as $f_{0?}$, $\mathrm{f}_{?0}$, or $f_{00}$. The percentage of individuals who are classifiable as being consistent with each policy does not sum to 100%, because of the "question mark" participants whose choices are consistent with two policies.

Clearly, we have a large number of participants who are not classifiable into any policy, ranging from 37% to 70% of the people, depending on the treatment cell. In Table 6

**TABLE 5**
Compliance Strategy Classification Rates, Allowing for 0% Error Classification Threshold

| Probability an Inspected, Compliant Firm Exits Group 2 | Compliance Policy | Compliance Cost = 7 | Compliance Cost = 100 | Compliance Cost = 200 | Compliance Cost = 375 |
|---|---|---|---|---|---|
| $u = 0.1$ | $f_{00}$ frequency (rate) | | 0 (0%) | 0 (0%) | 0 (0%) |
| | $f_{01}$ frequency (rate) | | 0 (0%) | 0 (0%) | 0 (0%) |
| | $f_{0?}$ frequency (rate) | | 4 (4%) | 0 (0%) | 0 (0%) |
| | $f_{1?}$ frequency (rate) | | 6 (5%) | 7 (6%) | 11 (10%) |
| | $f_{10}$ frequency (rate) | | 21 (19%) | 9 (8%) | 2 (2%) |
| | $f_{11}$ frequency (rate) | | 2 (2%) | 3 (3%) | 14 (13%) |
| | $f_{?0}$ frequency (rate) | | 5 (5%) | 8 (7%) | 0 (0%) |
| | $f_{?1}$ frequency (rate) | | 2 (2%) | 6 (5%) | 30 (28%) |
| | Other freq. (rate) | | 71 (64%) | 77 (70%) | 52 (48%) |
| | *Total subjects* | | 111 | 110 | 109 |
| $u = 0.1$ Summary | $f_{00}$ consistent (rate) | | 9 (23%) | 8 (24%) | 0 (0%) |
| | $f_{10}$ consistent (rate) | | **32 (80%)** | 24 (73%) | 13 (23%) |
| | $f_{01}$ consistent (rate) | | 6 (15%) | 6 (18%) | 30 (53%) |
| | $f_{11}$ consistent (rate) | | 10 (25%) | **16 (48%)** | **55 (96%)** |
| | Classifiable subjects | | 40 | 33 | 57 |
| | Optimal policy | | $f_{10}$ | $f_{11}$ | $f_{11}$ |
| $u = 0.9$ | $f_{00}$ frequency (rate) | 30 (27%) | 2 (2%) | 0 (0%) | 0 (0%) |
| | $f_{01}$ frequency (rate) | 0 (0%) | 0 (0%) | 0 (0%) | 0 (%) |
| | $f_{0?}$ frequency (rate) | 31 (28%) | 1 (1%) | 0 (0%) | 0 (%) |
| | $f_{1?}$ frequency (rate) | 1 (1%) | 3 (3%) | 14 (13%) | 13 (12%) |
| | $f_{10}$ frequency (rate) | 5 (5%) | 45 (41%) | 35 (32%) | 18 (16%) |
| | $f_{11}$ frequency (rate) | 2 (2%) | 0 (0%) | 2 (2%) | 19 (17%) |
| | $f_{?0}$ frequency (rate) | 1 (1%) | 1 (1%) | 0 (0%) | 0 (0%) |
| | $f_{?1}$ frequency (rate) | 0 (0%) | 0 (0%) | 1 (0.9%) | 5 (5%) |
| | Other freq. (rate) | 41 (37%) | 57 (52%) | 58 (53%) | 56 (50%) |
| | *Total subjects* | 111 | 109 | 110 | 111 |
| $u = 0.9$ Summary | $f_{00}$ consistent (rate) | **62 (89%)** | 4 (8%) | 0 (0%) | 0 (0%) |
| | $f_{10}$ consistent (rate) | 7 (10%) | **49 (94%)** | **49 (94%)** | 31 (56%) |
| | $f_{01}$ consistent (rate) | 31 (44%) | 1 (2%) | 1 (2%) | 5 (9%) |
| | $f_{11}$ consistent (rate) | 3 (4%) | 3 (6%) | 17 (33%) | **37 (67%)** |
| | Classifiable subjects | 70 | 52 | 52 | 55 |
| | Optimal policy | $f_{00}$ | $f_{10}$ | $f_{10}$ | $f_{11}$ |

*Note:* The percentage rates shown in the *frequency* section of the table are percentages of the total number of subjects making choices in that treatment condition. The percentage rates shown in the *consistent* section of the table are percentages of the classifiable subjects in that treatment condition. These latter percentages sum to greater than 100 percent because some subjects' observed choices are consistent with multiple compliance policies. The numbers in **bold** are the number of classifiable subjects consistent with the optimal policy.

we present an alternative classification based on the policy that provides a best fit to each participant's choices. This simple procedure counts the number of "errors" assuming that participants followed a particular strategy and yields a strategy classification for every participant that minimizes the number of errors in classification. For some participants, two policies are equally best fitting. This occurs most frequently when participants did not make choices in both inspection groups.

The results are largely consistent across the two classification methods in the two tables. Both indicate that more participants were consistent with the optimal policy (shown in bold on the tables) than with any other policy for six treatment cells, with the exception being the cell where compliance cost is medium and

**TABLE 6**

Best-Fitting Compliance Strategy for Each Subject in Each Treatment

| Probability an Inspected, Compliant Firm Exits Group 2 | Compliance Policy | Compliance Cost = 7 | Compliance Cost = 100 | Compliance Cost = 200 | Compliance Cost = 375 |
|---|---|---|---|---|---|
| $u = 0.1$ | $f_{00}$ frequency (rate) | | 30 (27%) | 31 (28%) | 8 (7%) |
| | $f_{10}$ frequency (rate) | | **85 (77%)** | 66 (60%) | 32 (29%) |
| | $f_{01}$ frequency (rate) | | 20 (18%) | 36 (33%) | 55 (50%) |
| | $f_{11}$ frequency (rate) | | 35 (32%) | **59 (54%)** | **96 (88%)** |
| | *Total subjects* | | 111 | 110 | 109 |
| | Optimal policy | | $f_{10}$ | $f_{11}$ | $f_{11}$ |
| $u = 0.9$ | $f_{00}$ frequency (rate) | **89 (80%)** | 14 (13%) | 3 (3%) | 2 (2%) |
| | $f_{10}$ frequency (rate) | 21 (19%) | **93 (85%)** | **88 (80%)** | 67 (60%) |
| | $f_{01}$ frequency (rate) | 43 (39%) | 9 (8%) | 9 (8%) | 10 (9%) |
| | $f_{11}$ frequency (rate) | 8 (7%) | 21 (19%) | 49 (45%) | **75 (68%)** |
| | *Total subjects* | 111 | 109 | 110 | 111 |
| | Optimal policy | $f_{00}$ | $f_{10}$ | $f_{10}$ | $f_{11}$ |

*Note:* The percentage rates shown in the *frequency* section of the table are percentages of the total number of subjects whose choices minimize the number of deviations from the indicated strategy in that treatment condition. The optimal policy is highlighted in **bold**. The percentages sum to greater than 100 percent because some subjects' observed choices are best fit by two different compliance policies, particularly when they do not make compliance choices in one of the inspection groups.

$u = 0.1$. Table 6 shows that in this cell 60% were consistent with policy $f_{10}$, and 54% were consistent with the optimal policy $f_{11}$. For all other conditions, at least two-thirds of the participants' strategies are consistent with the optimal policy. The switching probability $u$ can be an important determinant in the person's decision making, particularly so when the compliance costs are high. When the compliance costs = 375, more participants are consistent with $f_{10}$ when $u = 0.9$ than when $u = 0.1$, although the optimal policy $f_{11}$ is still played by a larger percentage of the participants. For these high compliance costs—which are more than double the single-period expected penalty—some people apparently increased their compliance rates because of the greater opportunity of moving back to the good group 1 as $u$ increased. This suggests that leverage works to some degree even when it is not predicted to work by the model.

Taken together, these points concerning the classification rates suggest that (1) some participants' behavior was either confused or consistent with some alternative model that we have yet to consider and (2) a large portion of participants chose the compliance policy predicted by the Harrington model. In the next section, we present an alternative choice model in an attempt to make sense of some of the systematic deviations from our model.

### *A Noisy Choice Model*

The Harrington model predicts that participants choose the optimal compliance policy with probability 1, regardless of whether this policy provides a return that is, for example, 331% higher or 10% higher than the next best alternative. Figure 2, however, shows that although participants are more likely to make choices that provide greater expected profits, their likelihood of making the optimal choice increases when its return is greater relative to its alternatives. This suggests that a model that permits errors in decision making might be useful to us in order to understand our experimental outcomes. In what follows we employ a "quantal choice" model that accounts for boundedly rational decision making. This model allows participants to make errors, but it accounts, in an intuitive way, for the fact that participants are less likely to make errors that are more costly.[7] In particular, it provides some structure for the distribution of errors

7. Figure 2 clearly shows how deviations from the optimal choice depend on the relative profitability of the different choices and thus rejects alternative choice error models that do not account for relative payoffs, such as the noisy Nash model. In the noisy Nash model, agents make their optimal choice with probability $\gamma$ and randomize (uniformly) over all choices, independent of their relative payoffs, with probability $1 - \gamma$; see McKelvey and Palfrey (1998).

that agents make, by relating the errors to their expected payoff consequences.

We use the logit form of the quantal choice model first introduced by Luce (1959) and popularized more recently by McKelvey and Palfrey (1995) in a game-theoretic context as a quantal response equilibrium. In our study participants were not playing a strategic game—just a game against nature because the inspector is not strategic. The idea is therefore quite simple: if strategy $i$ has expected utility $U_i$, it is played with probability

$$(3) \qquad q_i = [\exp(U_i/m)] \Big/ \left[\sum_{all\ j} \exp(U_j/\mu)\right]$$

The parameter $\mu$ is estimated from the data and scales the sensitivity that participants have to the relative payoffs (in terms of utility) of the various choices. As $\mu$ decreases, the participants put less probability weight on choices that yield suboptimal payoffs, and the probability that they make the optimal choice approaches 1 as $\mu$ approaches 0. As $\mu$ approaches infinity, subjects choose their available strategies with equal probability, independent of the relative expected payoffs.

This framework also allows us to determine if risk aversion, either as a competing or a complementary explanation to this type of boundedly rational decision making, might also explain the deviations from optimal choices. Risk aversion is sometimes argued to lead to higher compliance rates than what is predicted, as risk-averse participants could be sensitive to the probability of being caught—see Alm et al. (1992). The greater risk of a fine increases the cost of violating while leaving unchanged the returns from complying. To introduce risk aversion in a simple way, we posit a constant relative risk-averse utility function for each subject of the form $U(\pi) = \pi^{1-\alpha}/(1-\alpha)$, where $\pi$ is the dollar payoff for the choice and $\alpha$ is the index of relative risk aversion. We can estimate both $\mu$ and $\alpha$ by maximum likelihood techniques within the same model. If $\alpha$ is significantly positive and $\mu$ is near zero, this would suggest that risk aversion rather than bounded rationality is a primary cause of the deviations from the optimal choice. We obtain the opposite result, however. In all of our estimates, whether looking at only late periods, only early periods, or all decisions, we find the maximum likelihood estimate of $\alpha$ to be 0 but $\mu$ to be positive and highly significant. Therefore, we reject risk aversion as a main explanation of our results and focus on the bounded rationality term $\mu$.

**TABLE 7**
Quantal Choice Model Maximum Likelihood Estimates

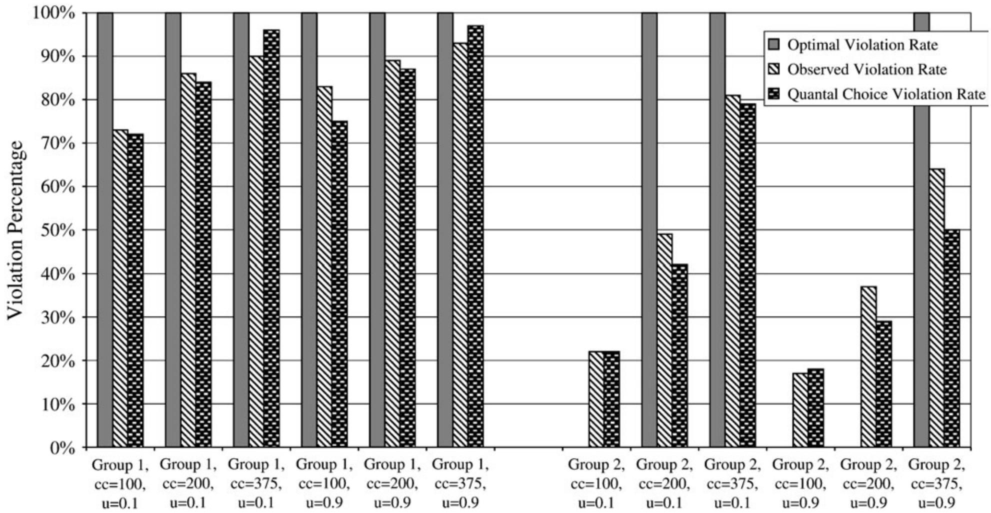
| Data Set | $\mu$ Estimate (standard error) | Log Likelihood | Number of Observations |
|---|---|---|---|
| All periods | 976 (35) | −2927.6 | 5764 |
| Early sequences 1–4 | 1144 (57) | −1915.4 | 3553 |
| Late sequences 5–7 | 747 (39) | −992.6 | 2211 |

To evaluate this model, we look at individual choices within an inspection group, similar to the previous analysis and Table 4. We consider three strategies for the participants: compliance policies $f_{00}$, $f_{10}$, and $f_{11}$ but not policy $f_{01}$ because $f_{01}$ is never optimal and is always played less frequently than other policies (as documented in Tables 5 and 6). Each of these three policies has an expected present value for the agent when in each inspection group for every treatment variable combination employed in the experiment, as shown in Table 2. For example, consider the scenario when the compliance cost is 100 and the transition probability $u = 0.1$: if agents are in inspection group 2, their expected payoff from policy $f_{00}$ (always comply) is 3,000; their expected payoff from policy $f_{11}$ (always violate) is 2,500; and their expected payoff from policy $f_{10}$ (comply only in group 2) is 3,124.62. We use these expected payoffs to calculate the probability of adopting each policy using equation 3 and then translate the rates at which participants choose these compliance policies into observed violation rates for each inspection group. For example, if $\mu = 976$, then the probability of adopting $f_{11}$ based on the expected payoffs just described is

$$(4) \qquad \exp(2500/976) \\ / [\exp(3000/976) + \exp(2500/976) \\ + \exp(3124.62/976)] = 0.219.$$

This is the probability of a violation in inspection group 2 implied by $\mu = 976$ because the other two policies ($f_{00}$ and $f_{10}$) prescribe compliance when in inspection group 2.

**FIGURE 3**
Observed and Predicted Violation Rates (Quantal Choice and Perfectly Optimal Benchmarks) All for $\mu = 976$




We make a similar transformation for all treatment configurations, inspection groups, and all possible μ and then compare the actual compliance decisions to determine the μ most consistent with the data. Table 7 presents the maximum likelihood estimate for μ, pooling across all six main treatment cells (i.e., all treatments except the baseline compliance cost = 7). We present results for all periods pooled as well as results separating the early session treatment sequences from the late session treatment sequences. Consistent with previous research that employs this quantal choice approach—for example, McKelvey and Palfrey (1995)—the choice errors decline as subjects gain experience. This is reflected in the significantly lower μ estimate for the late period sequences.

Figure 3 illustrates the remarkable success that this simple one-parameter model has in explaining the deviations from the optimal choices, based on the pooled estimate for the entire data set.[8] It is important to keep in mind that the noise parameter does not provide freedom to explain any deviations; instead, each particular value of μ is consistent with only one specific combination of deviations across our treatments. Nevertheless, all of the observed violation rates are accurately predicted by the model, with the greatest deviation being only 14%. Moreover, the model accurately captures the qualitative differences across treatments, such as the higher group 1 violation rates when the compliance cost is greater.

## V. DISCUSSION

Enforcement and monitoring of regulatory compliance policies can incur substantial resource costs. Dynamic audit models help us in understanding how individuals and firms might behave when faced with enforcement and compliance rules that are conditional on actions in previous and current periods. Harrington's important model (1988) demonstrates how a regulator could use multiple inspection groups to increase enforcement leverage when political or other practical considerations limit the size of fines. Despite a body of theoretical research in this area, empirical analysis of the compliance strategies of individuals in this dynamic framework is limited by a lack of observability for key variables in the theories.

8. We could obviously fit the observed rates more accurately with treatment cell-specific μ estimates. As Haile et al. (2003) have recently emphasized, however, it is important to leave the estimated parameter unchanged across treatments in order to make comparative statics exercises informative. See Goeree et al. (forthcoming) for further discussion.

Laboratory evidence presented in this article shows that, in a broad sense, participants' behavior is consistent with the theoretical predictions of this dynamic enforcement model. Overall violation rates are significantly higher in group 1 than in group 2. When compliance costs are higher, the violation rates increase significantly. We obtain clear support for the more subtle prediction that compliance increases in the "bad" group 2 if it is more likely to be rewarded with a transition back to the "good" group 1. That is, our results support the general idea of enforcement leverage through transitions across multiple groups.

An examination of the compliance policies chosen by the participants reveals that a large proportion of them chose the strategy predicted by the Harrington model. Participants in our experiments do not, however, follow the sharp predictions of the model. The deviations are more pronounced when the model makes corner solution predictions even though the differences in expected profits are marginal for alternative policies or actions. To account for this, we considered a quantal choice model where subjects were assumed to be boundedly rational. The standard rational choice model assumes that firms and individuals respond to regulatory policies by choosing strategies that increase their payoffs. They might, however, not choose the exact optimal strategy at all times; that is, they may make some mistakes, although it seems sensible that they would tend to make fewer mistakes when the mistakes are more costly. This aspect of bounded rationality is often neglected in a policy setting.

To understand individual and firm behavior and formulate policies that provide incentives for better regulatory enforcement, our results suggest that more attention be paid to models that incorporate noisy decision making. The quantal choice model accurately accounts for the boundedly rational behavior of our participants, and it may also be useful for describing compliance choices of agents in the field. When faced with decisions such as reporting income for taxation purposes and environmental regulation, agents might often be boundedly rational. Though they would choose strategies that increase their earnings, they might be prone to errors at the margin, where the incentives to optimize are not very high. How they act at this margin could in some cases determine the success or failure of the regulatory policy, and the implications of such suboptimal behavior should be examined carefully. In some applications the compliance decisions are made by groups rather than individuals, and future research should study whether groups' rationality is also bounded similarly—see, for example, Cason and Mui (1997) and Blinder and Morgan (forthcoming).

If bounded rationality in this context displays a robust influence on behavior, then enforcement models themselves should be more accurate if they incorporate bounded rationality explicitly. For example, the Harrington model implies optimal endogenous enforcement parameters to maximize efficiency (for each particular compliance cost) in which the firm only slightly prefers to comply rather than violate in the high-intensity inspection group. Because at this margin the firm is nearly indifferent between the two strategies, the alternative behavioral prediction from the quantal choice model instead predicts that the firm would comply only half the time. This obviously has important implications for the choice of the optimal enforcement rule in practice.

## APPENDIX

**TABLE A-1**

Wilcoxon Signed Rank Tests Comparing Violation Rates in Inspection Group 1 to Inspection Group 2

| Probability an Inspected, Compliant Firm Exits Group 2 | Compliance Cost = 7 | Compliance Cost = 100 | Compliance Cost = 200 | Compliance Cost = 375 |
|---|---|---|---|---|
| $u = 0.1$ | | **1072.5**** | 489.5** | 99.5** |
| $u = 0.9$ | 184.5** | **1757**** | **1517.5**** | 960.5** |

*Note:* The numbers reported in each cell are the values of the Wilcoxon signed rank test statistic. Only the **bold** numbers reflect differences that the model predicts to be significant.

**significant at the 1% level.

**TABLE A-2**
Wilcoxon Signed Rank Tests Comparing Violation Rates for Different Compliance Costs

| Probability an Inspected, Compliant Firm Exits Group 2 | | Inspection Group 1 | Inspection Group 2 |
|---|---|---|---|
| $u = 0.1$ | Cost = 200 versus Cost = 100 | 181** | **800**** |
| | Cost = 375 versus Cost = 200 | 39 | 1121** |
| $u = 0.9$ | Cost = 100 versus Cost = 7 | **2084**** | 23 |
| | Cost = 200 versus Cost = 100 | 283** | 264.5** |
| | Cost = 375 versus Cost = 200 | 126 | **576**** |

*Note:* The numbers reported in each cell are the values of the Wilcoxon signed rank test statistic. Only the **bold** numbers reflect differences that the model predicts to be significant.

**significant at the 1% level.

**TABLE A-3**
Wilcoxon Signed Rank Tests Comparing the Violation Rates for Different Probabilities of Exiting from Group 2 ($u$)

| Inspection Group | Compliance Cost = 100 | Compliance Cost = 200 | Compliance Cost = 375 |
|---|---|---|---|
| 1 | −80.5 | −19.5 | −27 |
| 2 | 207.5** | **566**** | 518.5** |

*Note:* The numbers reported in each cell are the values of the Wilcoxon signed rank test statistic. Only the **bold** numbers reflect differences that the model predicts to be significant.

**significant at the 1% level.

## REFERENCES


Alm, J., M. Cronshaw, and M. McKee. "Tax Compliance with Endogenous Audit Selection Rules." *Kyklos*, 46, 1993, 27–45.

Alm, J., B. Jackson, and M. McKee. "Institutional Uncertainty and Taxpayer Compliance." *American Economic Review*, 82(4), 1992, 1018–26.

Alm, J., and M. McKee. "Extending the Lessons of Laboratory Experiments on Tax Compliance to Managerial and Decision Economics." *Managerial and Decision Economics*, 19, 1998, 259–75.

Arora, S., and S. Gangopadhyay. "Toward a Theoretical Model of Voluntary Overcompliance." *Journal of Economic Behavior and Organization*, 28, 1995, 289–309.

Blinder, A., and J. Morgan. "Are Two Heads Better Than One? An Experimental Analysis of Group versus Individual Decision Making." *Journal of Money, Credit, and Banking*, forthcoming.

Cason, T., and L. Gangadharan "Emissions Variability in Tradable Permit Markets with Imperfect Enforcement and Banking." *Journal of Economic Behavior and Organization*, forthcoming.

Cason, T., and V.-L. Mui. "A Laboratory Study of Group Polarisation in the Team Dictator Game." *Economic Journal*, 107, 1997, 1465–83.

Clark, J., L. Friesen, and A. Muller. "The Good, the Bad and the Regulator: An Experimental Test of Two Conditional Audit Schemes." *Economic Inquiry*, 42, 2004, 69–87.

Eckert, H. "Inspections, Warnings and Compliance: The Case of Petroleum Storage Regulation." *Journal of Environmental Economics and Management*, 47, 2004, 232–59.

Fischbacher, U. "z-Tree–Zurich Toolbox for Readymade Economic Experiments." Working Paper No. 21, Institute for Empirical Research in Economics, University of Zurich, 1999.

Friesen, L. "Targeting Enforcement to Improve Compliance with Environmental Regulations." *Journal of Environmental Economics and Management*, 46(1), 2003, 72–86.

Goeree, J., C. Holt, and T. Palfrey. "Regular Quantal Response Equilibrium." *Experimental Economics*, forthcoming.

Greenberg, J. "Avoiding Tax Avoidance: A (Repeated) Game-Theoretic Approach." *Journal of Economic Theory*, 32(1), 1984, 1–13.

Haile, P., A. Hortaçsu, and G. Kosenok. "On the Empirical Content of Quantal Response Equilibrium." Cowles Foundation Discussion Paper No. 1432, Yale University, 2003.

Harford, J. "Firm Behavior under Imperfectly Enforceable Standards and Taxes." *Journal of Environmental Economics and Management*, 5(1), 1978, 26–43.

Harrington, W. "Enforcement Leverage When Penalties Are Restricted." *Journal of Public Economics*, 37, 1988, 29–53.

Helland, E. "The Enforcement of Pollution Control Laws: Inspections, Violations and Self-Reporting." *Review of Economics and Statistics*, 80(1), 1998, 141–53.

Landsberger, M., and I. Meilijson. "Incentive Generating State Dependent Penalty System: The Case of Income Tax Evasion." *Journal of Public Economics*, 19, 1982, 333–52.

Linder, S., and M. McBride "Enforcement Costs and Regulatory Reform: The Agency and Firm Response." *Journal of Environmental Economics and Management*, 11, 1984, 327–46.

Luce, R. D. *Individual Choice Behavior: A Theoretical Analysis.* New York: John Wiley and Sons, 1959.

McKelvey, R., and T. Palfrey. "Quantal Response Equilibria in Normal Form Games." *Games and Economic Behavior*, 7, 1995, 6–38.

———. "Quantal Response Equilibria for Extensive Form Games." *Experimental Economics*, 1, 1998, 9–42.

Oljaca, N., A. G. Keeler, and J. Dorfman. "Penalty Functions for Environmental Violations: Evidence from Water Quality Enforcement." *Journal of Regulatory Economics*, 14, 1998, 255–64.

Salop, S., and D. Scheffman. "Raising Rivals' Costs." *American Economic Review* (*Paper and Proceedings*), 73(2), 1983, 267–71.

Storey, D., and P. McCabe. "The Criminal Waste Discharger." *Scottish Journal of Political Economy*, 27(1), 1980, 30–40.

Torgler, B. "Speaking to Theorists and Searching for Facts: Tax Morale and Tax Compliance in Experiments." *Journal of Economic Surveys*, 16(5), 2002, 657–83.