When Threats Become Credible: A Natural Experiment of Environmental Enforcement from Florida

Wesley Blundell*

December 10, 2019

Abstract

Environmental regulators often use dynamic enforcement, which bases penalties and enforcement effort on plants' past compliance history, to improve compliance and decrease emissions when enforcement resources are limited. Using plant-level data from the Environmental Protection Agency (EPA), I examine an unexpected shift in the use of traditional enforcement by environmental regulators in Florida, showing that all of the state's plants decreased emissions and improved compliance following an increase in penalties for those with Priority Violations. The largest improvements were observed among plants with the highest expected costs of compliance, which is consistent with the theory of dynamic enforcement. These results are robust to the use of control plants from nearby southern states, as well as control plants selected via a matching algorithm. The paper's findings (1) provide quasi-experimental evidence on the effectiveness of traditional enforcement actions, and (2) suggest that dynamic incentives may matter for plant compliance decisions.

Keywords: Air pollution; Compliance; Enforcement; Clean Air Act; Manufacturing; Dynamic regulation.

---

*Assistant Professor, Department of Economics, California State University, East Bay. Contact at wesley.blundell@csueastbay.edu or at (510) 885 - 4685. No financial support was received for this research. Conflicts of interest: none. I thank Dietrich Earnhart, Mary Evans, Jay Shimshack, Ashley Langer, Tiemen Woutersen, Gautam Gowrisankaran, Derek Lemoine, Kyle Wilson, Anatolii Kokoza, Ivan Rudik, Jed DeVaro, Ryan Lampe, Gavin Conner, and Alex Hollingsworth for helpful comments. I am also grateful for suggestions from seminar participants at the University of Arizona, the Arizona Environmental Economics Group, the Western Economic Association International 2016 Summer Conference, the International Industrial Organization 2019 conference, and, the Association of Environmental and Resource Economists 2016 and 2019 Summer Conferences. All mistakes in this paper are my own.

When Threats Become Credible: A Natural Experiment of Environmental Enforcement from Florida

Abstract

Environmental regulators often use dynamic enforcement, which bases penalties and enforcement effort on plants' past compliance history, to improve compliance and decrease emissions when enforcement resources are limited. Using plant-level data from the Environmental Protection Agency (EPA), I examine an unexpected shift in the use of traditional enforcement by environmental regulators in Florida, showing that all of the state's plants decreased emissions and improved compliance following an increase in penalties for those with Priority Violations. The largest improvements were observed among plants with the highest expected costs of compliance, which is consistent with the theory of dynamic enforcement. These results are robust to the use of control plants from nearby southern states, as well as control plants selected via a matching algorithm. The paper's findings (1) provide quasi-experimental evidence on the effectiveness of traditional enforcement actions, and (2) suggest that dynamic incentives may matter for plant compliance decisions.
Keywords: Air pollution; Compliance; Enforcement; Clean Air Act; Manufacturing; Dynamic regulation.

1 Introduction

The economic theory of enforcement and criminal punishment is well known: in order to deter crime, the expected punishment for a crime must exceed its expected benefit (Becker, 1968). However, there is a significant challenge in designing policies that minimize the sum of criminal and enforcement costs. "Dynamic enforcement," where penalties and other enforcement actions are based on previous noncompliance, is one policy that can help reduce these costs. In the general economics literature, dynamic enforcement, state-dependent enforcement, graduated punishments, and otherwise similar practices have been shown to achieve cost minimization even when monitoring is imperfect or penalties are restricted (Ostrom, 1990; Polinsky and Rubinfeld, 1991; Dixit, 2009).

In environmental economics, dynamic enforcement is consequential because, despite low penalties for first-time offenders, regulators can deter noncompliance through the threat of increased penalties and monitoring for repeat offenders. This means that a plant must consider the current penalty for a violation, as well as the increased cost for any future violations. The effect is bolstered if there are associated reputational costs (Hamilton, 1995; Konar and Cohen, 2001). Although the combination of these factors makes dynamic enforcement theoretically appealing, it poses significant challenges in terms of empirically estimating the effectiveness of traditional enforcement measures. These challenges include the potential for omitted variable bias, reverse causality, and the improper specification of firms' regulatory perceptions (List, 2007; Gray and Shimshack, 2011; Shimshack, 2014).

Currently there is only limited empirical evidence on the effect of traditional enforcement actions within a dynamic enforcement setting (Helland, 1998; Eckert, 2004; Shinkuma and Managi, 2012; Blundell, Gowrisankaran, and Langer, 2018). This study employs a unique quasi-experiment with detailed panel data on emissions, regulatory activity, and compliance in order to evaluate the response of manufacturing plants in the United States to the use of traditional enforcement, in a dynamic setting. This experiment follows the federal EPA's review of the Florida Department of Environmental Protection (DEP) in 2012. The EPA found that penalties issued previously by the Florida DEP for "High-Priority Violations" were inadequate. Following the review, there was a substantial increase in penalties for priority plants in Florida, with little change in the enforcement for other plants. The limitation of a penalty increase for priority plants, without changing the criteria for classification, allows for clean identification of the effects of traditional enforcement actions on plant behavior.

Therefore, this study makes two contributions to the literature on enforcement. First, I use quasi-experimental variation to quantify the extent to which an increase in penalties improves compliance and other environmental outcomes. Second, I provide empirical evidence that firms' responses from this quasi-experiment are consistent with the theory of dynamic enforcement, indicating that dynamic incentives may matter for the design of enforcement policy.

Using a difference-in-differences approach, I find that all Florida plants had an increased rate of compliance following the 2012 review. These effects range from 0.15% to 43.1%

with the largest estimated effects for plants that were a priority as well as plants that were noncompliant but not classified as a priority during the pre-treatment period. In addition, I examine the mechanisms behind the improved compliance of Florida plants. I find that the emission of harmful pollutants from all Florida plants decreased roughly 31% and evaluation failures by Florida plants with a history of noncompliance decreased by 2.1%. I show that these results are consistent with the underlying theory of dynamic enforcement by using a Difference-in-Difference-in-Difference framework. Florida plants with the highest expected costs of compliance had the largest improvements in compliance and other environmental outcomes.

These findings are robust and not based on sample selection. In particular, corresponding results are found using control plants from Alabama, South Carolina, and Georgia, as well as with control plants selected by subindustry and plant characteristics via a matching algorithm. Using supplementary data, it is evident that the improvements in Florida plant compliance continued into 2017. Furthermore, I find that the changes in compliance found in Florida cannot be explained by macroeconomic shocks. Overall, these findings indicate that plant responses to increased enforcement are consistent with the theory of dynamic enforcement, but do not rule out other possible theories.

This study's findings should extend to other situations, including environmental settings, where dynamic enforcement has been documented. This includes enforcement of the Clean Water Act (Shimshack and Ward, 2008; Earnhart, 2004), regulations related to petroleum storage (Eckert, 2004), waste disposal (Stafford, 2002; Shinkuma and Managi, 2012), and oil spills.[1] Non-environmental situations include tax compliance and worker safety (Landsberger and Meljinson, 1982; Ko, et al., 2010). Evans (2016) also finds significant effects with non-traditional enforcement via a watch list, in conjunction with the dynamic enforcement of the Clean Air Act (CAA).

[1]"2016 Rule Changes" Page 32 (State of North Dakota, 2016).

2 Institutional Details

Prior to 1970, the federal government did not play a significant role in the regulation of air pollution. In the absence of federal regulation, some states enacted their own regulations for polluters within their jurisdictions. However, amid concerns about the adverse effects on health posed by high concentrations of total suspended particulate matter, Congress passed the Clean Air Act amendments of 1970, and created the Environmental Protection Agency (EPA).

A series of amendments saw CAA enforcement encompass multiple regulatory programs to reduce harmful pollutants, including the Acid Precipitation Program, the New Source Review program, Maximum Allowable Control Technology standards, and others. To ensure that plants are compliant with every program for which they qualify, the EPA has created a system of monitoring and enforcement, primarily implemented by the environmental protection agencies within each state.[2]

Plants generally require a full inspection to verify their compliance at least once every five years. Larger facilities, or those emitting hazardous pollutants, are subject to more regular compliance evaluations.[3] Other measures for monitoring compliance include partial inspections, stack tests conducted by the agency, stack tests submitted by the facility, and remote evaluations on submitted facility diagnostics. These measures are also used for general enforcement, evidence collection for pending litigation, and information collection for the revision of future regulatory standards. The criteria for noncompliance are generally consistent across states, with standards varying in cases where the state adopts stricter standards than the federal EPA. Once a facility is found in noncompliance by committing at least one violation in a regulatory program, the facility can be issued an enforcement action

[2]Some of these state agencies are called EPA as well, e.g. "California Environmental Protection Agency" but still operate independently from the federal EPA.

[3]To quote the EPA's AFS Business Rules Compendium (2011) "Major Sources: FCE every 3 Years."

such as a warning or penalty.

In 1998, the EPA issued guidelines for addressing a certain subset of CAA violations, known as "High-Priority Violations" (HPV, priority violations, or priority).[4] The purpose of these guidelines was to encourage individual state agencies, which conduct the majority of enforcement activity, to prioritize noncompliant plants that are the most "environmentally important," as well as "to permit an increased degree of agency flexibility." Effectively, there are several criteria for classifying a violation as priority. Among them are whether the violation has generated a substantial amount of pollution, or whether it is a persistent violation by a recalcitrant plant. Theoretical literature examines the benefits of having enforcement vary according to a plant's violation persistence, or its violation history. In a survey of the enforcement literature, Shimshack (2014) notes that a major implication of this type of dynamic enforcement is "the regulator saves monitoring resources without sacrificing deterrence." However, this has led to problems in the empirical literature in determining the effectiveness of enforcement actions on ensuring compliance "because plants with more frequent noncompliance are often targeted for more frequent inspections and enforcement actions, simple statistical associations often show a negative correlation between enforcement and compliance" (Gray and Shimshack, 2011).

Although the standards for classifying priority violations and noncompliance is consistent across all states, they are granted flexibility in addressing priority violations, or other cases of compliance, whether it be through increased monitoring, penalties, or both. To limit variation in the intensity of state regulator enforcement, the EPA established the State Review Framework (SRF), a national system for periodically reviewing state CAA enforcement. Under this system, each state regulator is issued an obligatory set of guidelines at least once every five years for their enforcement practices, with the first set of guidelines issued in 2004. For state regulators that fail to abide by these guidelines, the EPA responses can range from

[4] "Issuance of Policy on Timely and Appropriate Enforcement Response to High Priority Violations" (EPA, 1998).

retraining the state's regulatory employees, to decertifying the State Implementation Plan. Early on in this study, the EPA implemented a "watch list," to further encourage states to pursue a subset of priority violations (Evans, 2016). To date, 85% of the EPA's first round of recommendations have been met, and the majority of actions against states have been in the lowest response tier.[5] Only a few states have failed to follow the EPA's recommendations after multiple reviews. This paper takes advantage of a state regulator modifying its enforcement of "High Priority Violations" following an EPA review to study the plant-level response to the use of traditional enforcement actions within a dynamic enforcement setting. In the next section, I provide evidence from the data of how enforcement of the Clean Air Act Amendments through the classification of "High Priority Violations" is consistent with dynamic enforcement.

3 The Dynamics of High Priority Violation Enforcement

3.1 Data

The majority of this study's empirical analyses use the Air Facility System (AFS) data for stationary sources, from the database of the EPA's Enforcement and Compliance History Online. These data consist of plant-level information on compliance, inspections, warnings, fines, fine amounts, and stack tests, and are subject to federal minimum data requirements. Because state regulators are incentivized to meet these requirements via the State Review Framework, these data provide a relatively complete regulatory history for large stationary sources of air pollution, from 2007 to 2014.

Compliance information in the AFS is derived from a combination of regulator reports and plant self-reports. Fraudulent self-reporting is considered severe and can result in criminal penalties. Therefore, this study follows the literature, and treats this self-reported data as

[5] "State Oversight Strategy" (EPA, 2014).

accurate (Laplante and Rilstone, 1996).[6][7] The key outcome of the analyses is the binary variable of compliance. This value reflects the minimum compliance status across a facility's air programs within a given quarter. A value of zero indicates noncompliance in at least one air program. The value is one if the plant is fully compliant, or if compliance is unknown and missing.[8] This coding is consistent with EPA performance tracking and other papers in the literature (Evans, 2016). Plants that are in violation, having a compliance value of zero, are classified as having a priority violation according to more stringent criteria. The quarterly nature of the compliance data dictates the aggregation of the other plant regulatory information to the plant-quarter level for my analysis.

The final dataset contains over a million quarterly observations over the period, corresponding to 31,000 different manufacturing facilities. Approximately 37.4% of all plants are in noncompliance for at least one quarter of the period, with approximately one tenth of these in noncompliance over the entire timeframe. To extend the data beyond 2014, when the AFS was retired, I use the EPA's new Integrated Compliance Information System (ICIS). This system contains similar plant-level information on enforcement, but tracks compliance information differently. The use of the ICIS, as well as data from the U.S. Census of Manufacturers, the EPA's Greenbook, the Bureau of Labor Statistics (BLS), the National Organization File, and the AFS Air Emissions Data, is discussed in the appendix. The use of these datasets allows me to consider: (i) alternative environmental outcomes beyond compliance, (ii) information about a plant's parent company, (iii) conduct sensitivity tests of the primary results through a number of necessary robustness checks.

[6]One example of criminal penalties would be the potential of four years imprisonment for defendants in a case regarding a gasoline-contaminated remediation project on Staten Island. `https://www.dec.ny.gov/press/109012.html`.

[7]Similar to Shimshack and Ward (2005), I also run a test for whether the self-reported stack test failure rate for plants changes in quarters where the regulator conducts their own tests at the plant and find no statistically significant difference.

[8]According to EPA documentation (Page 14, EPA Business Rule Compendium) this occurs when a plant hasn't been evaluated within the mandated timeframe, e.g. after two years for Title V plants.

3.2 Criteria for Priority Classification and Stylized Facts

This subsection discusses the criteria for classifying plants as priority violators, and which criteria are explicitly dynamic. Subsequently, two stylized facts are presented regarding the consistency of priority violation enforcement with dynamic enforcement. The first stylized fact is that, compared with compliant plants, and noncompliant plants that are non-priority, priority violators are subject to higher penalties and levels of monitoring. This is consistent with the idea that regulatory actions or penalties increase with a plant's history of noncompliance or violations. The second stylized fact is that, among noncompliant plants likely to be classified as priority violators, those eventually classified as such are more likely to be compliant one year later. This pattern is evidence of what the literature calls "enforcement leverage" (Shimshack, 2014), or the dynamic incentive for plants to undertake costly compliance investments, in order to lower the levels of monitoring and enforcement anticipated in the future.

Criteria for the Classification of Priority Violators

During the sample time period, plants could be classified as having "Priority Violations" based on 10 general criteria or 18 matrix criteria. These criteria can be separated into three broad types[9]: (i) explicitly dynamic criteria, such as "violation of direct surrogate of >25% for two reporting periods"; (ii) event criteria that refer to one major pollution event, such as "violation of allowed emissions limit"; (iii) paperwork criteria that do not necessarily indicate a major environmental event, such as "failure to obtain a PSD or NSR permit."[10]

As this study does not observe plants falling short of the thresholds for being classified with a priority violation, estimating the significance of dynamic incentives based on the prevalence of the dynamic criteria in the classification of priority violations is impractical. For

[9] I use the criteria descriptions on pages 56 - 57 of the EPA's AFS Business Compendium (2011) to guide this breakdown.

[10] A likely reason a plant would fail to obtain a permit and commit a paperwork violation is because they do not have the required abatement technology.

example, we do not observe how many plants were in "violation of direct surrogate of >25%" for one period, but then complied before the second period due to the dynamic incentive to avoid triggering that priority violation criteria. In addition, many of the criteria in the paperwork and event categories could also create dynamic incentives for plant compliance. For instance, small equipment failures that trigger a status of regular noncompliance, could be resolved due to the increased risk of incurring a larger violation that meets the criteria of an environmental event. Paperwork criteria could provide dynamic incentives for plants, if the plants "NSR or PSD permit" renewals are related to the installation of up-to-date pollution abatement technology, for example.

One additional aspect of the EPA's priority violator program that creates dynamic incentives is that plants can only exit their priority status once all outstanding violations at the facility have been resolved, not just the initial violation that triggered one of the criterion. Therefore, if priority violators are subject to higher levels of enforcement than other types of noncompliant plants, then any of the three criteria types has an underlying dynamic incentive by making subsequent violations more costly. In the following subsection I provide empirical evidence that priority violators are subject to dramatically higher levels of enforcement.

Fact #1: Priority plants are subject to the highest levels of enforcement.

Consistent with the theory of dynamic enforcement, priority violators are subject to increased monitoring and enforcement. The left panel of Figure 1 shows that plants classified as priority violators average 1.12 inspections per quarter. Noncompliant but non-priority plants receive 0.27, and compliant plants receive an average of 0.2 inspections per quarter. The center panel of Figure 1 shows a similar pattern for the quarterly averages of fines per each inspection. Priority plants averaged around $5,941 each quarter per inspection, which is significantly more than the average of \$302 for plants in regular noncompliance.[11] It is important to note that the quarterly penalties per inspection for compliant plants are generally small (an average of \$45.19 per quarter), which indicates that the penalized offense did not merit classification into the federal standard of noncompliance. The right panel of Figure 1 shows the same increasing pattern of the average number of quarterly stack tests across compliance classifications. Overall, these patterns of increasing enforcement with noncompliance statuses are consistent with the use of dynamic enforcement, but they do not rule out other possible theories. For instance, regulators may increase monitoring and enforcement for plants with higher compliance costs under a static enforcement scheme (no consideration of previous noncompliance history) if those plants' infractions are more severe. In this second case, the classification of priority violations would not depend on previous violations or a plant's history of noncompliance. However, the following stylized fact provides additional evidence in support of the presence of dynamic incentives with CAA enforcement.

Fact #2: Priority plants are more likely to be compliant within a year as compared to similar plants.

The second stylized fact provides evidence use of the priority violator status creates "enforcement leverage" (Shimshack, 2014). Specifically, since priority violator status is accompanied by additional scrutiny and the stipulation that, for the status to be lifted, plants must be "in compliance with all aspects of CAA requirements," priority violators are incentivized to comply more than other noncompliant plants. To demonstrate evidence of "enforcement leverage" empirically, I start with the transition function from Blundell, Gowrisankaran, and Langer (2018), to predict the transition of noncompliant but non-priority plants to priority violator status. This transition function is based on each plant's discounted violations, regulatory warnings, fines, and received inspections, as well as its specific industry and EPA

[11]The penalties in Figure 1 are also conditional on the plant receiving at least one inspection.

region Next, I group these noncompliant plants into five bins according to their predicted probability of being classified as a priority violator (0 – 20%, 20 – 40%, 40 – 60%, 60 – 80%, 80 – 100%). Finally, I regress the indicator for whether a plant is classified with a priority violation on the plant's compliance status one year later. Figure 2 plots the correlation coefficients of priority violator status for each of these five separate regressions. In all instances, the correlation is positive, and in certain bins the estimate is statistically significant. Although there may be differences in the unobservable characteristics between priority violators and non-priority plants, the bias would be expected to work against these results. Plants classified as priority violators should have worse unobservable characteristics, such as higher compliance costs, and therefore be less likely to be in compliance one year later.

Overall, these results are consistent with the idea that once plants are classified as priority violators they have a dynamic incentive to comply and therefore avoid a perpetually higher level of monitoring and enforcement. This contrasts with the null result one may expect under a static framework where otherwise similar plants would have similar levels of expected future enforcement and therefore similar rates of future compliance.

4 Theory

This section presents a model of how dynamic enforcement affects plant compliance. I alter the Harrington (1988) model to allow for stochasticity in the classification of a plant into the priority classification group. For simplicity, this model has only two groups rather than the three used by the EPA.[12] The goal is to provide a simple framework to think about how increased penalties for a subset of violators affects compliance behavior across different groups of plants. The strengths and limitations of this model are discussed with respect to the complex reality of CAA enforcement. Finally, the model clarifies predictions to take to the data. More details and proof are in the appendix.

[12] In compliance; noncompliant but not a priority-violator; and priority violator).

4.1 Two Group Model

The regulator classifies plants into two groups, G_1 and G_2, with plants in the second group facing higher levels of regulatory scrutiny. The inspection probability in G_i is P_i and the penalty for being found in violation is F_i, with $P_2 > P_1$ and $F_2 > F_1$. A violation in G_1 leads to classification in group G_2 with probability u. This means that being found in noncompliance does not guarantee being placed in G_2, which is consistent with the empirical setting of this paper. Finally, if no violation is discovered during an inspection in G_2 the plant is placed into G_1.

From the plant's perspective, this framework poses a Markov decision process. The plant will move from each classification group according to transition probabilities that depend on both the plant's current group classification and compliance decision. In choosing to comply, the plant will pay a compliance cost c. Otherwise, the plant risks being inspected and fined according to its classification. The plant's payoffs are constant over time and the plant discounts future payoffs by a factor $\beta \in (0, 1)$.

Under these conditions, Harrington (1988) shows that the plant's optimal policy, the policy that minimizes the expected discounted sum of plant costs, has a number of characteristics. First, the optimal policy is stationary. Second, for a given fixed compliance cost c, a plant's optimal policy is a constant mapping between the group, G_i, and compliance decision as follows: plants with low compliance costs, $c \leq P_1 \cdot F_1$, will always comply; medium cost plants, plants with costs $P_1 \cdot F_1 < c \leq L_1$, will alternate between groups by choosing $\{Don't\,Comply, Comply\}$; high cost plants, those with compliance costs $c > L_1$, will never comply by choosing $\{Don't\,Comply, Don't\,Comply\}$.

One key takeaway from Harrington (1988) is that the threshold between medium and high cost plants, L_1, exceeds the maximum expected single period penalty for priority group plants. Here L_1 is defined as:

$$L_1 = P_2 \cdot F_2 + \frac{P_2 \cdot \beta(P_2 \cdot F_2 - P_1 \cdot F_1)}{[1 - (1 - P_1 \cdot u) \cdot \beta]} > P_2 \cdot F_2 \tag{1}$$

Although the result that a plant with a compliance cost greater than the maximum expected penalty will sometimes comply seems counterintuitive, the reasoning is straightforward. By choosing to pay the compliance cost in the priority group G_2, the plant will benefit by facing lower future penalties while in G_1. This "dynamic enforcement leverage" allows regulators to achieve higher rates of compliance when maximum penalties are restricted, $P_2 \cdot F_2$, or enforcement budgets are fixed. Using this framework, I consider the impact of an increases in the level of penalties and monitoring for G_2 plants from F_2 to $F_2^{'}$ and from P_2 to $P_2^{'}$. In the appendix, I derive four propositions regarding this enforcement change which form the basis of the four predictions in the empirical analysis.

4.2 Relation to the Empirical Setting and Limitations

Before discussing the four empirical predictions from the two group model, I discuss a number of its limitations and explain how these limitations motivate the construction of the sample used in my primary empirical analysis. The first limitation stems from the fact that the EPA has three enforcement groups, plants in compliance, noncompliant but not-priority plants, and plants classified as priority violators. Framing this three group system within the Harrington (1988) correspondence between compliance history and unobserved compliance cost would mean that we should observe four types of plant histories corresponding to four cost types. That is, we should observe plants that are always in compliance, plants alternating between in compliance and non-priority noncompliance status, plants that alternate between all three classifications, and plants that are always priority violators. However, this fourth history of always a priority violator is not observed amongst Florida manufacturing plants during my sample period.[13]

Since I only observe a policy change for three compliance history types in Florida, applying the insights from the two-group model (with three observed compliance histories) is

[13]It is also extraordinarily rare among other U.S. states in my sample, so these observations are ultimately discarded.

relatively straightforward. For the first part of my empirical analysis, I use the mapping between unobserved plant compliance costs and observed compliance history in the theory model to characterize estimation samples. I define sample one plants as those never observed out of compliance in the pre-treatment period, sample two plants as plants observed as non-compliant but non-priority in the pre-treatment period, and sample three plants as plants with priority violations during the pre-treatment period. This stratification of samples allows for a direct comparison of the empirical results to predictions from the theoretical model. The second part of the analysis recombines the samples and uses a proxy for unobserved compliance costs, to further test the theory.

The second limitation of the two group model is that it abstracts from the dynamic nature of entry and exit decisions which impact the overall compliance rate. Plants with high unobserved compliance costs could exit to avoid additional penalties, raising the overall compliance rate over time. Additionally, potential entrants may base their decision to enter a market on its level of enforcement, resulting in new compositions of entrants when policies are amended. For these reasons, I limit my primary sample to plants that meet the following criteria: (i) plants indicated as "Operating" in the "AFS Facilities" snapshot; (ii) plants with at least one enforcement measure (inspection, stack test, penalty, warning) prior to the EPA review and one full compliance observation after. This adjustment should mitigate the potential endogeneity bias from entry and exit in my main results.

Third, the two-groups model does not consider that plant compliance costs may change over time. Such a model would not change the overall predictions to the extent that control plants in the empirical analysis would be representative of treated plants over time. To see this, consider the case of sample one plants (those observed as always compliant) prior to treatment. If treatment and control plants have similar underlying cost transition functions, then after the policy change a similar portion of treatment and control sample one plants would remain in compliance, a similar portion would transition to noncompliance, and a similar portion would become priority violators. Given that the cost threshold between choosing to never comply and sometimes comply only increases for treated plants, the overall post-treatment compliance rate for treated plants will be higher than that of the control plants. This last point would slightly alter the proof for Proposition IV where rather than no change among low cost plants there may in fact be some positive change in the compliance rate of low cost plants over time. However, this does not negate the overall insights derived from the two-group model.

Fourth, the compliance responses related to increased enforcement depend on the unobserved distribution of plant compliance costs, and the change in the inspection rate. In the empirical analysis, the actual distribution of plant compliance costs will dictate the size of the plant's treatment response. However, an uneven distribution of plant compliance costs would most likely bias the study from finding results in line with the predictions in the following subsection.

4.3 Predictions

My empirical application examines an increase in penalties and monitoring for plants classified as priority violators, which is the most severe noncompliance group. The two-groups model predicts that such a policy change may have several observable effects. Using the mapping between observed pre-treatment compliance history and compliance costs, I formulate four predictions based on the four propositions in the appendix.

Prediction I: The overall compliance rate will increase.

Prediction II: Compliance will increase among plants with a history of being classified with priority violations.

Prediction III: Compliance will increase among plants with a history of noncompliance, but no priority violations.

Prediction IV: The compliance response by plants with higher compliance costs will exceed the compliance response by plants with low compliance costs.

The first three predictions test whether the increased use of traditional enforcement increases compliance in a dynamic enforcement setting. However, these predictions would be expected under a static enforcement setting, as a static increase in penalties should lead to an increase in compliance. In contrast, Prediction IV tests whether a plant's compliance behavior is consistent with its adherence to dynamic incentives and reflects the opposite of what one would expect under a static enforcement setting. Under a static enforcement setting, where the regulator is unable to observe plant compliance costs, one would expect plants with higher costs of compliance to respond less to an increase in penalties than plants with lower costs of compliance. All four of the predictions outlined in this subsection are subject to the limitations discussed in the previous subsection. However, these and other concerns are mitigated to the extent that control plants represent the counterfactual compliance path of treated plants in the empirical investigation.

5 Empirical Design

The central empirical challenge of this article is to test whether plants respond to the use of traditional enforcement within a dynamic setting. Credibly measuring the impact of changes in the application of enforcement requires variation in plant exposure that is uncorrelated with unobserved determinants of compliance. This is empirically difficult to justify. For example, a simple comparison between states with varying use of enforcement may be biased if plants' locational decisions are determined by their own compliance costs and beliefs about a state's regulatory capability.

This paper solves the identification problem by exploiting an unexpected increase in enforcement for priority plants in Florida. Theory predicts that after the increase, all Florida plants should have an increased incentive to comply, regardless of their current regulatory status. However, the magnitude of the response depends on a plant's cost of compliance. Using two control groups, plants from other southern states and plants selected via a matching algorithm, I utilize the change in Florida within a difference-in-differences setting to separate out the effect of increased enforcement for priority violations from other time-varying determinants of compliance and regulator activity. I then use a Difference-in-Difference-in-Differences (DDD) framework to test if there is a differential change in the behavior of Florida's higher-compliance-cost plants that is consistent with the theory of dynamic enforcement.

5.1 The Natural Experiment

In June 2012, the EPA conducted a review of the Florida Department of Environmental Protection (DEP). The review was the DEP's second under the State Review Framework (SRF), and it found that several previous recommendations remained unaddressed. Details in the final report issued in May 2013 indicate that "Priority Violations were not addressed in a timely and appropriate manner."[14] Additionally, penalties were not "consistent with national policy and guidance."[15] Under the EPA's SRF tier response system, potential consequences for the DEP could include the EPA conducting enforcement activity within Florida, or a reduction in federal funding. Given the EPA's monthly communication with the DEP, and with the stated deadline of June 2013, the DEP would have been notified, and had every incentive to begin altering the treatment of priority violations, shortly after the 2012 review.[16] Subsequent EPA reviews indicate that the DEP did alter its treatment of priority facilities during this period, and that by 2015, the DEP had addressed priority violations in an "appropriate manner."

There is anecdotal evidence of firms' perceptions regarding the change in regulatory practices during this time. One Florida plant with a priority violation describes, in its 2013

[14]"State Review Framework and Integrated Clean Water Act Permit Quality Review Florida" (EPA, 2013).

[15]More specifically they did not follow the EPA's designated "penalty-and-financial models."

[16]"By June 30, 2013, DEP should submit and implement revised procedures to improve the timeliness of HPV addressing actions." Page 79 (EPA, 2013).

annual report, that CAA regulations were becoming "increasingly stringent."[17] This report coincided with the plant receiving a penalty four times greater than a previous priority violation penalty. A subsequent merger case for the firm that owns this plant also describes air pollution violations as an area of concern.[18] This perceived change in the practices of the DEP is also corroborated by descriptive statistics.

Focusing on the penalties paid by priority violator plants, Figure 3 shows that, from 2014 to 2017, the average quarterly penalties paid by these plants increased in comparison to the penalties from 2010 to 2013 (the one-year delay in the increase from the 2012 review is most likely explained by the length of the penalty process). Figure 4 shows a small but immediate increase in the monitoring of priority plants starting from 2013 that continues until the end of the data in 2017. Column 1 of Table I provides a comparison of the pre- and post-review means of these enforcement variables. The difference in the average number of quarterly stack tests and penalties for priority plants is positive in both cases. The increases in monitoring and enforcement in Florida represent a stark contrast with enforcement trends outside of Florida.

Table I provides evidence on the uniqueness of Florida's regulatory change by comparing enforcement measures, per violation, for three groups of priority facilities, pre- and post- the 2012 review[19]: (i) plants in Florida; (ii) plants in South Carolina, Alabama, and Georgia; and (iii) plants in the rest of the United States. Row 1 of Table I highlights that, although the difference-in-differences for the mean between groups pre- and post- penalties are not statistically significant (p-values of 0.174 and 0.289), there is a stark contrast between the increase in penalties for priority Florida plants, and the large downward trend in penalties for priority plants in nearby southern states, as well as the rest of the United States. Row

[17]"Annual Report December 2013," Page 50 (Ardagh Group, 2013).

[18]Alex Lawson, "FTC Approves Ardagh Antitrust Fix in 1.7B Saint Gobain Deal." Law360. June 18, 2014.

[19]Since the number of violations is not directly observable in the data, I follow the literature (Blundell, Gowrisankaran, and Langer, 2018) and use the 10% depreciated accumulated notices of violation as a proxy.

2 of Table I highlights that Florida plants experienced a substantial increase in stack tests compared to the other groups, with the difference between the change with other southern states being almost statistically significant at conventional levels (p-value = 0.121). Row 3 of Table I indicates that the differential change in inspections for Florida plants was positive and statistically significant at the 1% level compared with the rest of the United States. Therefore, it is reasonable to conclude that the changes in monitoring and enforcement for priority plants in Florida, contrasts with general enforcement trends.

I also examine enforcement trends for noncompliant but non-priority plants in Florida, to determine whether the changes in enforcement for priority plants coincided with a general increase in regulatory enforcement. Figures 5 and 6 demonstrate that this is not the case, as both monitoring and the average penalties per violation, paid by noncompliant but non-priority plants in Florida, do not demonstrate any discernible change in the post-review period. This helps to attribute any measured changes in the compliance behavior of Florida plants to the Florida DEP's application of dynamic enforcement in the primary empirical analysis.

5.2 Difference-in-Differences Estimation Framework

All plants in Florida are considered as treated after the second quarter of 2012. According to predictions made by the theoretical model, the size of the treatment effect will depend on the plant's pre-review history of noncompliance. Plants with no history of noncompliance are sample one. Plants with a history of noncompliance but no priority violations are sample two. Plants with a history of priority violations are sample three. The first step of the empirical analysis uses difference-in-differences in a linear regression framework to focus on the three samples of plants.

In all primary specifications, the outcome of interest is next quarter compliance ($Comp$). This outcome is binary, taking a value of zero if the plant is classified as noncompliant in the following quarter. Because compliance can take the value of zero for noncompliance at any point during the quarter, the full treatment effect associated with moving from noncompliant to compliant cannot be estimated without a review of the next quarter's compliance. The choice of next-quarter compliance is also consistent with other papers in the literature (Evans, 2016). Thus, each difference-in-differences specification is a variant of the following:

$$Comp_{i,t+1} = \alpha_0 + \alpha_1 \cdot FL \cdot Post + W'_{i,t} \cdot \beta + \gamma_i + \delta_t + \varepsilon_{i,t} \tag{2}$$

Post is an indicator taking a value of one for any quarter after the second quarter 2012 Florida review. *FL* is an indicator function that takes a value of one if the plant is located in the state of Florida. γ_i represents plant i's fixed effect, which controls for any unobserved plant specific time invariant characteristics that effect compliance, such as management practices. δ_t is the quarter dummy and $W'_{i,t}$ is a vector of controls to account for the heterogeneity between plants over time.

The controls for equation (2) include: (i) the annual average emissions for that plant's industry, as plants that are more pollutant have higher compliance costs; (ii) the nonattainment status of the county where the plant is located, because regulatory standards vary with a county's nonattainment status; and (iii) the number of manufacturing establishments within the county, which accounts for the finding in the literature that enforcement is less stringent when manufacturing is a major local employer (Deily and Gray, 1991).[20] The error term, $\varepsilon_{i,t}$, includes the unobserved determinants of plant compliance. I cluster at the plant level to account for plant-specific shocks to compliance, but provide permutation tests using both bootstrapped and state-level clustered standard errors as a robustness check in the appendix. Finally, the plant fixed effect will absorb any state specific (FL) effect on compliance and the quarter dummies will absorb any individual *Post* effect, hence their exclusion from the equation.

The DID estimate, therefore, is given by $\hat{\alpha}_1$, the coefficient on the interaction between

[20]Industry emissions is averaged at the six digit NAICS code level, the finest level of industry classification available in the data.

FL and *Post*. This approach teases out the Florida Plants' responses to the post-review increased enforcement, from other factors that changed during the period. Identification of α_1 requires two assumptions. First, control plants must satisfy the parallel trends assumption or represent the counterfactual change in compliance if no regulatory intervention had occurred. Second, no other changes to Florida plant compliance incentives coincided with the EPA's review. One example of how this second assumption could be violated is if the period overlapped with the term of a lame duck governor, as in Doyle et al. (2016), which is not the case with Florida. Another example of a possible violation of the second assumption would be if there were macroeconomic shocks unique to Florida that coincided with the review. Section six discusses robustness checks and other evidence for the validity of these two assumptions in my setting.

5.3 Difference-in-Difference-in-Difference

Ultimately, a simple DID specification will only test whether plants responded to an increase in the use of traditional enforcement in a dynamic setting. To test whether these responses are consistent with the underlying theory that higher-compliance-cost plants exhibit the largest response, the analysis is expanded with a difference-in-difference-in-differences estimation. Equation (2) is modified as follows:

$$Comp_{i,t+1} = \alpha_0 + \alpha_1 \cdot FL \cdot Post + \alpha_2 \cdot Treat \cdot Post + \alpha_3 \cdot FL \cdot Treat \cdot Post + W'_{i,t} \cdot \beta + \gamma_i + \delta_t + \varepsilon_{i,t} \quad (3)$$

Here, *Treat* is an indicator for whether a plant has a differing compliance cost. The key coefficient of interest is α_3, which is the interaction between *Treat* and whether a plant is in Florida during the post review period. This separates any differential response by higher compliance cost Florida plants to the increase in penalties for priority violations.

Additionally, assuming that no other treatments coincide, the identification of α_3 requires a stronger set of assumptions over the previous DID approach. In particular, two different trend assumptions must be met: (i) higher cost control plants represent the counterfactual compliance response for higher cost Florida plants; (ii) lower cost control plants represent the counterfactual response for lower cost Florida plants. Therefore, in order to identify this model, a measure of compliance cost is needed that will satisfy the two assumptions.

Since I do not observe plant compliance costs directly, I define *Treat* using three variables that are related to compliance costs: (i) whether the plant's parent company has another plant outside Florida; (ii) the number of manufacturing establishments in the county; or, (iii) the number of depreciated violations. The advantage of conducting three separate DDD analyses, using different definitions of $Treat$, is that it alleviates concerns about endogeneity in the timing of the event in Florida, and it defines higher-compliance-cost plants.

The first proxy for unobserved compliance costs comes from Rijal and Khanna (2017), who find that when a plant is subject to increased regulatory scrutiny due to priority violator classification, the plant's pollution is transferred to other facilities within the parent company. In this case, *Treat* takes a value of one if the plant's parent company has one or more manufacturing facilities located outside the state of Florida, since these plants have a lower cost of compliance with the 2012 policy change through their ability to move their pollution out of state. Using this first proxy variable, a finding that α_3 is negative and statistically significant would be consistent with Prediction IV of the theoretical model, plants with lower costs of compliance had a smaller compliance response to the EPA review. The parallel trends assumption for the identification of α_3 is then satisfied if Florida plants with parent company locations outside of Florida have a parallel compliance trend with control plants outside of Florida that have other non-Florida locations. In addition, higher compliance cost Florida and control plants, those with no additional locations outside Florida, must have a parallel compliance trend. These two assumptions are violated if the 2012 policy change coincided with other changes in the compliancy of multi-plant manufacturing firms, or if "Priority Violator" policy, applied differentially to plants with large parent companies, separate the differences in the underlying compliance costs over time. This latter concern is mitigated to the extent that there are no provisions regarding multi-plant firm violations within the "Priority Violator" criteria. This is consistent with conversations with EPA personnel that indicate initial investigations are not based on the compliance of multiple plants within the same parent company. The former concern is mitigated to the extent that control plants with a particular $Treat$ value are comparable to Florida plants with the same $Treat$ value.

The second proxy for unobserved plant compliance costs is the number of manufacturing establishments in a county, which is also an additional control in the DID discussion. Specifically, *Treat* is the total number of county manufacturing establishments, with larger values indicating that manufacturing is a more significant contributor to local employment and that local manufacturing plants should have a lower cost of compliance. The advantage of this second proxy is that the compliance decision of an individual plant has no impact on this measure of unobserved compliance costs. Similar to the first proxy, a finding that α_3 is negative and statistically significant when using this second proxy variable would be consistent with Prediction IV. Also similar to the first proxy variable, both parallel trends assumptions must be satisfied for the DDD specification to identify the heterogeneous effects of the 2012 change in Florida's CAA enforcement policy. Here, the parallel trends assumption would be violated if the 2012 policy change differentially impacted plants in manufacturing intensive counties in other ways, separate from plant compliance costs.

The third proxy is the total number of notices of violation a plant has received since they were last in compliance, depreciated at a rate of 10% per quarter. Here, *Treat* is a count where higher values indicate the plant is likely to have a higher unobserved cost of compliance. Unlike the first two proxy variables, a finding that α_3 is positive and statistically significant when using this third proxy variable would be consistent with Prediction IV of the theoretical model. In order for the third proxy variable to satisfy both parallel trends assumptions, the conditions for issuing a notice of violation must not have varied with the 2012 review or differentially varied between treatment and control plants. This assumption is supported by the fact that the 2012 review of the Florida Department of Environmental Protection does not indicate a need to change the notice of violation policy.

By testing for whether α_3 is statistically different from zero in the DDD model, I can determine whether there is significant heterogeneity in the response to enforcement of priority violations across plants with varying costs of compliance. This helps to understand whether plant compliance behavior is consistent with the use of dynamic enforcement. Results that are consistent with the presence of compliance response heterogeneity and with Prediction IV of the theoretical model would indicate that dynamic incentives may matter for the determination of plant compliance. A finding that α_3 is statistically indistinguishable from zero and therefore inconsistent with Prediction IV, would indicate that the increase in compliance following the EPA's review is best explained by a change in the static incentives for compliance. This result could be from an increase in the static expected total penalty for noncompliance, as well as a change in the marginal deterrence (Mookherjee and Png, 1994) between committing low-level (non-priority) violations and high-level (priority) violations.

6 Control Groups and Results

This section describes the empirical results. I begin in Section 6.1 with a descriptive summary of plant compliance rates, showing how Florida plants trended in a similar direction to southern control plants prior to the 2012 review and experienced a dramatic increase in their rate of compliance afterward. I then estimate the DID model which measures the treatment effect of the increased application of traditional enforcement on Florida plant compliance. Section 6.2 estimates the DID model using a set of control plants, selected via a matching algorithm. Section 6.3 provides a descriptive comparison of the heterogeneity across plant compliancy responses. The DDD model is then estimated, guided by the predictions from the theoretical model. Section 6.4 discusses the mechanisms by which plants achieve compliance, and the environmental results. Finally, Section 6.5 discusses robustness checks and the extension of the estimation sample using the ICIS data.

6.1 Control Plants from Other Southern States

First, the effect of the Florida DEP's reform is measured by comparing the outcomes of Florida plants to plants in the nearby southern states of South Carolina, Georgia, and Alabama. There are three appealing aspects for using these southern plants to form the control groups. First, plants in these states should face similar economic conditions to plants in Florida. Second, all four states had their EPA watchlist facilities revealed during the same period in 2011.[21] Third, these states' individual EPA reviews did not coincide with the timing of Florida's review. [22] These elements should make plants from these states comparable to Florida plants in other unobservable compliance determination factors.

Figure 7 plots the average compliance rate among all treatment and control plants in the eight quarters pre and post the EPA's 2012 review. Prior to the review, Florida plants and southern control plants' compliance followed a similar trend, with Florida plants exhibiting a slightly lower average rate of compliance. Following the review, Florida plant compliance increases to approximately 97% within four quarters, indicating that time was required for plants to adjust to the increased scrutiny. Overall, the post review increase in Florida plant compliance is roughly one percentage point larger than the increase in control group compliance. The appendix includes a comparison of the summary statistics for the pre-treatment period, by treatment status and sample. The largest statistically significant differences are those related to mean compliance, or the enforcement actions taken against sample three plants, which form the motivation behind the EPA's review of the Florida DEP.

Panel A of Table II reports the primary DID results with all samples (Column 1), sample one plants (Column 2), sample two plants (Column 3), and sample three plants (Column 4). All specifications include plant-level fixed effects and quarter fixed effects. Also included, as

[21]"Clean Air Act Facilities on the Active 2011 Watch List" (EPA, 2011).

[22]March 2010 for SC, August 2011 for GA, and September 2010 for AL.

additional controls, are average industry emissions, county nonattainment status, and county manufacturing activity. Standard errors are clustered at the plant-level. The estimated effect of the EPA review is an increase of 0.4% for all Florida plants, similar to the differential increase observed in Figure 7. The treatment effect is positive and statistically significant for sample two and three plants, with the estimated average increase ranging from 0.15% for sample one plants, to 9.4% for sample two plants. The small point estimate for sample one plants is consistent with the fact that all were compliant in the pre-treatment period, and only a small percentage are ever in noncompliance post-review. Across specifications, both county nonattainment status, and industry emissions level, negatively correlate with compliance, which proves consistent with these factors indicating higher underlying compliance costs for the plant. For county level manufacturing activity, the correlation with compliance has a mixed sign and an occasionally statistically insignificant coefficient.

Overall, the increase in plant compliance across groups is consistent with Predictions II and III. Plants with a history of noncompliance or priority violator status, improved their compliance following an increase in enforcement for priority violators. Given the pre-treatment compliance rates of 21.8% and 23.5% for sample two and three plants, these effects translate to a 43.1% and 30.6% decrease in noncompliance, respectively. These changes correspond with the compliance effects found in the literature. Evans (2016) estimates that the CAA watch list decreased noncompliance by 23%, and Gray and Shadbegian (2005) document deterrence effects that led to decreases in noncompliance for pulp and paper mills by 10%. Therefore, these findings are significant as they provide quasi-experimental evidence of a compliance response by Florida plants to increased levels of enforcement, in line with other effects in the literature.

6.2 DID With Matched Control Plants

A potential weakness of the primary DID estimation is the presence of unobserved or endogenous time-varying determinants of compliance that differ between treatment (Florida) and control plants. Specifically, the southern control plants may have a different distribution of unobserved compliance costs across estimation samples than Florida plants. This is a valid concern as the enforcement levels varied between Florida and the control states prior to treatment, and, under the theoretical framework, plants would sort themselves into different estimation samples according to the expected penalties and monitoring for priority violators in their state.

To address this concern, I implement a conditional DID estimator similar to Heckman et al (1998) and Cicala (2014). In my estimation, I match on a number of plant characteristics rather than a propensity score.[23] The intuition is that if these plant characteristics are strongly correlated with unobserved plant compliance costs, then the matched control groups will be more comparable to Florida plants than the southern control plants.

The following estimation adopts notation from the matching literature (Dehejia and Wahba 1999; Heckman, Ichimura, and Todd 1998). Consider the set of covariates for an individual Florida plant i, X_i. Let $1\{\}$ be an indicator function that takes a value of one if the bracketed statement is true. Let l be the index of facilities outside of Florida that are in the same subindustry as plant i.[24] Then,

$$\sum_l 1\{\|X_j - X_i\| \leq \|X_l - X_i\|\} = m \tag{4}$$

Equation (4) identifies the m closest facilities to plant i, located outside the state of Florida, according to the Mahalanobis distance metric. First, in the matching algorithm which selects control plants, three characteristics are matched; the number of manufacturing plants, the nonattainment status of the county where the plant is located, and the year of first observation in the ICIS (dating back to the 1970s, as older plants more likely have higher compliance costs).[25] The nearest match per Florida plant is collected, including plants that

[23]Which would be equivalent to predicting whether a plant is located in Florida or not.

[24]Four digit NAICS code.

[25]I do not include this as an additional control in the regression analysis since it would be

tie with the same matching value. These matching criteria result in 1,321 control plants in sample one, 243 control plants in sample two, and 51 control plants in sample three.[26] The appendix includes a comparison of the summary statistics between Florida plants and their matches in the pre-treatment period, revealing a more balanced set of plant characteristics, in comparison with the southern control plants.

Panel B of Table II presents the DID estimates with the matched control plants, using the same specifications as in Panel A. The presentation of the results is also analogous to Panel A, with the results for the whole sample (Column 1), sample one plants (Column 2), sample two plants (Column 3), and sample three plants (Column 4). In all specifications, the treatment effects are about 50% greater than the results with the southern state control plants; however, the coefficients are alike on the additional controls. These larger treatment effects are comparable with the other results from the literature. In particular, Deily and Gray (2007) document that enforcement pressure has a 33% effect on CAA compliance in the steel industry, and Telle (2013), using a field experiment, estimates a 37% effect from environmental enforcement in Norway. The positive estimated compliance increase of 2.2% for all Florida plants is statistically significant, indicating that that review led to an increase in the overall compliance rate, consistent with Prediction I. Overall, the results of this alternative estimation also provide evidence in support of both Predictions II and III.

6.3 DDD Results

This subsection considers heterogeneity in Florida plants' compliance responses to the increased enforcement of priority violations. Both the previous DID results, as well as predictions from the theoretical model, indicate that plants with a previous history of noncompliance or priority violator status will exhibit a large compliance response. Figure 8 shows the quarterly compliance trends of previously noncompliant or priority violator plants from

absorbed by the fixed effect.

[26]Some control plants matched with multiple Florida plants.

Florida and the southern control states, around the time of the EPA's 2012 review. Prior to the review, Florida plants and the southern control plants followed a comparable trend in compliance, with Florida plants exhibiting a lower average rate. Following the review, Florida plants with a history of noncompliance, or priority violator status, exhibited an approximately 10% larger increase in compliance, when compared with the southern control plants. As with the aggregate, Figure 7, for all plant samples, Florida plant improvements occur within one year of the review.

The DDD estimation outlined in section 5.3 is used to highlight the heterogeneity in compliance responses among sample two and sample three plants. My preferred specifications use three definitions for the third interaction $Treat$: (i) whether the plant's parent company has another plant outside of Florida; (ii) the number of county manufacturing establishments; or, (iii) the number of depreciated accumulated violations. If the compliance response heterogeneity is consistent with dynamic enforcement, then plants with higher/lower underlying compliance costs, should have larger/smaller responses to the 2012 policy change, under the conditions of no large right tail in plant compliance costs, or a large proportional change in penalties.[27]

Column 1 of Table III reports the primary DDD results, with $Treat$ being an indicator that takes a value of one if the parent company for the plant has at least one other manufacturing facility outside of Florida. The differential responses by these lower-compliance-cost plants in Florida is estimated to be -6.8% using the southern control plants (panel A) and -5.9% using the matched control plants (panel B). Both differential responses are statistically significant at the 10% level. The negative differential effects are smaller than the estimated baseline compliance increases for all Florida plants, 12.3% with southern control plants, and 21.1% with matched control plants. This indicates that lower-compliance-cost plants improved their overall rate of compliance after the 2012 review, but to a lesser extent than

[27]The first is supported by the fact that no Florida plant is observed to always be a Priority Violator, the latter by the large increase in penalties.

Florida plants with higher compliance costs.

Column 2 of Table III reports the main DDD results with $Treat$ representing the count of the number of manufacturing facilities in a plant's county. Similar to the previous specification, a higher value of $Treat$ indicates a lower underlying cost of compliance. Results using the matched control plants (Panel B) indicate that lower compliance cost Florida plants had a differential response of -1.1% (per 100 county establishments) as compared to higher compliance cost Florida plants. Furthermore, this estimated difference is statistically significant at the 10% level. Given the estimated 22.3% baseline compliance improvement for these Florida plants, these results indicate that when considering the matched control group, lower compliance cost Florida plants had a smaller increase in compliance than other Florida plants. The estimated differential compliance response using southern control plants is statistically indistinguishable from zero.

Column 3 of Table III reports the primary DDD results, with $Treat$ representing the count of depreciated violation notices. Unlike the previous two specifications, a higher $Treat$ value indicates a higher underlying compliance cost. Results using the southern control plants (Panel A) indicate that, in Florida, higher-compliance-cost plants had a large and statistically significant differential compliance increase of 14.1% (per each depreciated accumulated violation), compared with other plants. These results are not sensitive to the inclusion of the total number of depreciated violations as an additional control.[28]

One additional advantage to using the count of depreciated violation notices in the DDD, is that this count positively correlates with the classification of a priority violation plant, according to dynamic criteria.[29] In addition to having higher compliance costs, plants with a higher count of depreciated violations would have a higher likelihood of being classified according to dynamic criteria if they were classified as a priority violator. Therefore, the

[28]In the previous two specifications in Table III, $Treat$ would be collinear with the plant fixed effect or additional controls.

[29]I use the sample of 3,700 cases where priority violation classification was reported and control for industry, state, and quarter fixed effects.

southern control plant results in Column 3 of Table 3 (Panel A), suggest that dynamic incentives may matter in a way that extends beyond just being consistent with the predictions of the theoretical model. This is because plants with a higher value of $Treat$ in this specification have a higher likelihood of triggering the dynamic criteria for priority violator status and should be most responsive to a change in the underlying dynamic incentives. In the appendix, Table III specifications are reported without additional controls, with qualitatively similar results.

Overall, Florida plants' compliance decisions are consistent with Prediction IV from the theoretical model. Prediction IV states that plants with higher compliance costs should exhibit the largest compliance increases. The negative triple difference estimates in Columns 1 and 2, and the positive estimate in Column 3 of Table III, provide evidence that plants with higher compliance cost did have the largest compliance response to the policy change.

It should be noted that the results in Table III do not let us conclude that the increase in compliance following the 2012 review can be exclusively attributed to a change in the underlying dynamic incentives for plant compliance. This is because a plant can be classified as a priority violator due to meeting dynamic criteria, such as "violation of direct surrogate of >25% for two reporting periods," or from meeting criteria that presents no clear dynamic compliance incentive, such as "violation of allowed emissions limit." The EPA's review incentivized an increase in penalties for all priority violations, regardless of classification criteria, and plants with higher costs of compliance could be more likely to violate all types of criteria. This makes it reasonable to conclude that plants with higher underlying costs of compliance experienced a larger increase in both their dynamic and static compliance incentives as compared to plants with lower costs of compliance. Therefore, a more effective test of the importance of dynamic incentives would be one where plants are randomized between two enforcement schemes that differ according to how heavily they penalize priority violations classified via dynamic criteria, while leaving the overall expected penalties for priority violations constant. The necessity of keeping expected penalties for priority violations constant in this scenario would allow for a full decomposition of the effect of a change in the dynamic incentives separate from a change in the overall level of penalties. Unfortunately, this type of ideal experiment is not available in the data, which indicates that a more structural approach such as the one used in Blundell, Gowrisankaran, and Langer (2018) is necessary to evaluate the effectiveness of dynamic enforcement in this setting.

6.4 Mechanisms for Compliance

This section examines the different mechanisms behind plant compliance, specifically for two outcomes: (i) the plant-level emission of harmful pollutants, and (ii) the evaluation failure rate of a plant. These outcomes are associated with both production activity and abatement technology. The mechanisms through which plants achieve compliance has implications for understanding the benefits of enforcement. For example, if plants achieve compliance by reducing emissions, then there are potential health benefits from increased levels of compliance.

To assess whether Florida plants reduced their emissions following the 2012 review, the EPA's AFS Air Emissions dataset is used. The dataset is a compilation of facility-level emissions data from multiple databases, including the Greenhouse Gas Reporting Program (GHGRP), Clean Air Markets, National Emissions Inventory (NEI), and Toxic Release Inventory (TRI). The pollutants used for this analysis come from the latter two databases. In general, the information is a combination of plant self-reports, engineering estimates based on the facilities' utilized materials or processes, and actual measured emissions. Since the criteria for plants to report this information is more stringent than the criteria for the AFS, plants that are matched to the AFS are subject to a selection bias by being larger and subject to more enforcement. In particular, the fact that only larger facilities report this information renders the samples one and two plants with emissions information almost identical in terms of observables. Overall, the data are considered accurate and representative of plants' emissions. More detail regarding the data's strengths, weaknesses, the merging process, and the differing characteristics of matched plants, is discussed in the appendix. In particular, 49% of the manufacturing facilities reported in the emissions data, match to the AFS, which is comparable to the 77.4% match rate between the smaller NEI and Census of Manufacturers database, reported by Shapiro and Walker (2018).

This empirical analysis uses the log of a plant's emission of total annual criteria pollutants. From an empirical perspective, the focus on criteria pollutants is a useful outcome measure because these pollutants are detrimental to human health. The relationship between CAA enforcement and these harmful pollutants provides information on the value of enforcement policy. The annual nature of the emissions data is accounted for by modifying equations (2) and (3) to denote the time subscript as year, as well as by dropping all 2012 observations, due to the mid-year EPA review. Panel A of Table IV reports the results of the change in the log of plant emissions, using the southern control group. There is a clear pattern of Florida plants decreasing their emissions following the intervention with the reductions ranging from 11.5% to 31.8%. The 11.5% estimated reduction for sample three plants is insignificant at conventional levels (p-value of 0.23), which should be attributed to the far right tail in emissions for these previously priority plants, making a large percentage reduction difficult. Results using the DDD specification, and $Treat$ being defined on whether a plant's parent company has another non-Florida location, show that plants with lower compliance costs had smaller emissions reductions, which is in line with the results in Table III. These emissions results indicate that Florida plants reduced their emissions following the mandate, with effects similar to the 11% to 16% decrease in emissions through nonattainment classification noted by Chay and Greenstone (2003).

Finally, I focus on whether a plant failed a quarterly evaluation. Plant evaluation failures (inspections or stack tests) are a useful outcome because they indicate whether the plant failed to use the required abatement technology, or if it emitted over the permitted levels. Plants can be evaluated for numerous reasons, although they are primarily due to the EPA's requirements for regular inspections, increased scrutiny of noncompliant plants, or for case building. These requirements could indicate a selection issue in the subsample of evaluated plants, as larger or more noncompliant plants have a higher likelihood of being evaluated. Although a sample selection issue seems unlikely to lead toward a negative bias, due to Florida's previous priority violators, plants with the worst performance being more likely to be evaluated after the review. Instead, the bias should be from finding a negative effect of the review on evaluation failures. Furthermore, there is no reason to believe that the selection of evaluated plants changed for the control groups, before or after the review.

Panel B of Table IV shows the regression estimates of the change in evaluation failures using the southern control group. The estimated effect of the review is negative and statistically significant for plants with a previous history of noncompliance, but no priority violations, with a mixed but statistically insignificant sign across the other specifications. The appendix reports results for both outcomes with the matched control group; the results are qualitatively similar but noisy due to the low number of observations. Overall, these results assist in explaining the underlying mechanisms for the Florida plants' improved compliance. Florida plants improved their compliance following the change in enforcement policy, by reducing their emissions of harmful pollutants. There is also some evidence that the review led to a decrease in the rate of Florida plants' evaluation failures.

6.5 Robustness

The robustness of the results are explored through several additional tests, reported in the appendix, that focus on the estimates obtained using the southern state control plants. First, the statistical significance of the DID results is examined with southern control plants, using both bootstrapped standard errors and standard errors clustered at the state level. The permutation tests show that the use of standard errors, which account for facility level shocks consistent with other papers in this literature, are not qualitatively different from other reasonable approaches.

Second, the validity of the estimation procedure is tested by estimating a series of falsification tests. For each state, I redefine the $FL \cdot Post$ term from equation (2), to take the value of one if a plant is located in a certain state after the second quarter of 2012, and zero otherwise. For the pooled sample of all plants, only three of the 49 other states had positive and statistically significant treatment results at the 10% level, which is in line with the number of false positives expected under a valid test, given the singularity of the Florida review.

Third, a placebo test is estimated of the DID and DDD models. The $Post$ term is redefined in equations (2) and (3), to take the value of one for starting in the second quarter of 2010. Because the quarters after the EPA review are not included, this provides an identical number of post treatment quarters to the original analysis. As anticipated, the estimated treatment effects are statistically indistinguishable from zero.

Fourth, I explore the sensitivity of the DID and DDD results to recovery from the Great Recession. To test the robustness of the DID results, equation (3) is modified to take the value of one if the county in which the plant is located had an unemployment rate exceeding 10% during the Great Recession. The robustness of the DDD results is tested by estimating a quadruple difference of the interaction of the high unemployment indicator with the indicator for whether a plant's parent company has an additional facility outside of Florida. These specifications should distinguish whether the compliance improvement of Florida plants is explained by a difference in the economic impacts of the Great Recession across states, rather than the EPA's review. Results indicate no differential compliance trends for Florida plants in counties with high unemployment during the Great Recession.

Fifth, I test the sensitivity of the DID results to the coding of missing compliance values as well as the dropping of plants due to entry or exit concerns. Results are qualitatively similar across samples and the exclusion of missing compliance values.

Finally, I extend the sample period from the end of Q3 2014 to Q4 2017 by using the ICIS data. The ICIS data do not contain a comparable measure of quarterly compliance. Therefore, a plant is defined as being in noncompliance if they have a priority violation, or Federally Reportable Violation (FRV). This is a restrictive definition of noncompliance because plants could be classified in the AFS as being noncompliant for the quarter, without committing a violation classified as an FRV. The results from this extended sample are qualitatively similar to the preferred DID and DDD specifications, using nearby southern state plants as a control group. The robustness of the primary results to the inclusion of additional years, indicates that this study's findings are unlikely to be driven by unobserved shocks to CAA enforcement coinciding with the Florida review, because a contemporaneous shock would be attenuated by an extended sample period. Therefore, these results rule out contemporaneous regulatory or state-specific shocks that could have driven the Florida compliance effects observed in the post-treatment period.

7 Conclusion

Recently, there has been a downward trend in regulators' enforcement resources. This has led to an expansion in the use of non-traditional enforcement measures (e.g., citizen monitoring), and a targeted use of traditional enforcement measures (e.g., penalties). Since October 2014, the EPA has made several changes to its enforcement policies, including revising the priority violation criteria, discontinuing the watch list for a subset of priority violations, and clarifying the use of compliance tools in enforcement settlements.[30] The stated purpose of these modifications was to emphasize the importance of targeting severe violators, and thereby deter noncompliance. Theoretical work has shown that this type of dynamic enforcement will encourage compliance among both high- and low-compliance-cost plants. However, there is little empirical evidence to support this theory.

Overall, this article provides some of the first quasi-experimental evidence on the role of traditional enforcement in determining plant compliance within a dynamic setting. I find that after an increase in penalties for plants with priority violations, the rate of compliance

[30]"Revision on Timely and Appropriate Enforcement Response to High Priority Violations" (EPA, 2014).

increased amongst all plants. As a consequence, there was a decline in the emission of pollutants harmful to human health. The change in enforcement policy also had the largest impacts on plants with higher expected compliance costs, which suggests that dynamic incentives may matter. This last result is particularly significant for policymaking, as the benefits of dynamic enforcement depend on the extent to which plants that regularly commit violations improve their performance.

The results indicate the need for further research on the effectiveness of enforcement in settings where dynamic enforcement is used. Some examples where this could be applied include the Clean Water Act, the Resource Conservation and Recovery Act, the enforcement of oil spills, and tax compliance. The importance of dynamic enforcement presumably varies with industrial and regulatory characteristics that determine the extent firms can respond to the use of enforcement. Recent trends in federal hiring and budgets, with a 15.7% decline in the number of federal EPA enforcement and compliance personnel since January 2017, indicate a possible reduction in enforcement capacity.[31] To the extent that it maximizes the value of enforcement resources and complements non-traditional enforcement activity, it may be in regulators' best interest to respond by increasing the use of dynamic enforcement. Further work in other settings will help to gauge the extent of implementation, but this study supports the effects of traditional enforcement in an important sector where dynamic enforcement is used.

References

[1] Gary S Becker. Crime and punishment: An economic approach. *Journal of Political Economy*, 76(2):169–217, 1968.

[2] Randy A Becker. Air pollution abatement costs under the clean air act: evidence from

[31] Juliet Eilperin "With a shrinking EPA, Trump delivers on his promise to cut government." Washington Post. September 8, 2018.

the pace survey. *Journal of environmental economics and management*, 50(1):144–169, 2005.

[3] Wesley Blundell, Gautam Gowrisankaran, and Ashley Langer. Escalation of scrutiny: The gains from dynamic enforcement of environmental regulations. Technical report, National Bureau of Economic Research, 2018.

[4] Kenneth Y Chay and Michael Greenstone. Air quality, infant mortality, and the clean air act of 1970. Technical report, National Bureau of Economic Research, 2003.

[5] Steve Cicala. When does regulation distort costs? lessons from fuel procurement in us electricity generation. *The American Economic Review*, 105(1):411–444, 2014.

[6] Rajeev H Dehejia and Sadek Wahba. Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, 94(448):1053–1062, 1999.

[7] Mary E Deily and Wayne B Gray. Enforcement of pollution regulations in a declining industry. *Journal of Environmental Economics and Management*, 21(3):260–274, 1991.

[8] Mary E Deily and Wayne B Gray. Agency structure and firm culture: Osha, epa, and the steel industry. *The Journal of Law, Economics, & Organization*, 23(3):685–709, 2006.

[9] Avinash Dixit. Governance institutions and economic activity. *The American Economic Review*, 99(1):3–24, 2009.

[10] Matthew Doyle, Corrado Di Maria, Ian Lange, and Emiliya A Lazarova. Electoral incentives and firm behavior: Evidence from us power plant pollution abatement. 2016.

[11] Dietrich Earnhart. Panel data analysis of regulatory factors shaping environmental performance. *The Review of Economics and Statistics*, 86(1):391–401, 2004.

[12] Heather Eckert. Inspections, warnings, and compliance: the case of petroleum storage regulation. *Journal of Environmental Economics and Management*, 47(2):232–259, 2004.

[13] Mary F Evans. The clean air act watch list: An enforcement and compliance natural experiment. *Journal of the Association of Environmental and Resource Economists*, 3 (3):627–665, 2016.

[14] Matthew Gibson. Regulation-induced pollution substitution. *Review of Economics and Statistics*, (0), 2015.

[15] Wayne B Gray and Ronald J Shadbegian. When and why do plants comply? paper mills in the 1980s. *Law & Policy*, 27(2):238–261, 2005.

[16] Wayne B Gray and Jay P Shimshack. The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence. *Review of Environmental Economics and Policy*, 5(1):3–24, 2011.

[17] James T Hamilton. Pollution as news: Media and stock market reactions to the toxics release inventory data. *Journal of Environmental Economics and Management*, 28(1): 98–113, 1995.

[18] Winston Harrington. Enforcement leverage when penalties are restricted. *Journal of Public Economics*, 37(1):29–53, 1988.

[19] James J Heckman, Hidehiko Ichimura, and Petra Todd. Matching as an econometric evaluation estimator. *The Review of Economic Studies*, 65(2):261–294, 1998.

[20] Eric Helland. The enforcement of pollution control laws: Inspections, violations, and self-reporting. *The Review of Economics and Statistics*, 80(1):141–153, 1998.

[21] Kilkon Ko, John Mendeloff, and Wayne Gray. The role of inspection sequence in compliance with the us occupational safety and health administration's (osha) standards: Interpretations and implications. *Regulation & Governance*, 4(1):48–70, 2010.

[22] Shameek Konar and Mark A Cohen. Does the market value environmental performance? *The Review of Economics and Statistics*, 83(2):281–289, 2001.

[23] Michael Landsberger and Isaac Meilijson. Incentive generating state dependent penalty system: The case of income tax evasion. *Journal of Public Economics*, 19(3):333–352, 1982.

[24] Benoıt Laplante and Paul Rilstone. Environmental inspections and emissions of the pulp and paper industry in quebec. *Journal of Environmental Economics and Management*, 31(1):19–36, 1996.

[25] Ardagh Packaging Holdings Limited. Annual report for the year ended 31 december 2013. 2013. URL `https://www.ardaghgroup.com/userfiles/files/aphl_fy_2013.pdf`.

[26] John A List. Field experiments: a bridge between lab and naturally occurring data. *The BE Journal of Economic Analysis & Policy*, 5(2), 2007.

[27] Dilip Mookherjee and Ivan PL Png. Marginal deterrence in enforcement of law. *Journal of Political Economy*, 102(5):1039–1066, 1994.

[28] Wallace E Oates. An essay on fiscal federalism. *Journal of Economic Literature*, 37(3): 1120–1149, 1999.

[29] State of North Dakota. 2016 rule changes. 2016. URL `https://www.dmr.nd.gov/oilgas/or27349.pdf`.

[30] Elinor Ostrom. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge University Press, 1990.

[31] A Mitchell Polinsky and Daniel L Rubinfeld. A model of optimal fines for repeat offenders. *Journal of Public Economics*, 46(3):291–306, 1991.

[32] Binish Rijal and Neha Khanna. High priority violations and intra-firm pollution substitution. 2017.

[33] Joseph S Shapiro and Reed Walker. Why is pollution from us manufacturing declining? the roles of environmental regulation, productivity, and trade. *American Economic Review*, 108(12):3814–54, 2018.

[34] Jay P Shimshack. The economics of environmental monitoring and enforcement. *Annu. Rev. Resour. Econ.*, 6(1):339–360, 2014.

[35] Jay P Shimshack and Michael B Ward. Regulator reputation, enforcement, and environmental compliance. *Journal of Environmental Economics and Management*, 50(3): 519–540, 2005.

[36] Jay P Shimshack and Michael B Ward. Enforcement and over-compliance. *Journal of Environmental Economics and Management*, 55(1):90–105, 2008.

[37] Takayoshi Shinkuma and Shunsuke Managi. Effectiveness of policy against illegal disposal of waste. *Environmental Economics and Policy Studies*, 14(2):123–145, 2012.

[38] Sarah L Stafford. The effect of punishment on firm compliance with hazardous waste regulations. *Journal of Environmental Economics and Management*, 44(2):290–308, 2002.

[39] Kjetil Telle. Monitoring and enforcement of environmental regulations: Lessons from a natural field experiment in norway. *Journal of Public Economics*, 99:24–34, 2013.

[40] Environmental Protection Agency USA. Issuance of policy on timely and appropriate enforcement response to high priority violations. 1998. URL `https://www.epa.gov/sites/production/files/documents/issue-ta-rpt.pdf`.

[41] Environmental Protection Agency USA. The air facility system (afs) business rules

compendium. 2011. URL https://www.epa.gov/sites/production/files/2013-10/documents/afsbusinessrulescompendium.pdf.

[42] Environmental Protection Agency USA. Clean air act facilities on the active 2011 watch list. 2011. URL https://echo.epa.gov/system/files/CAASept2011WatchListPost.pdf.

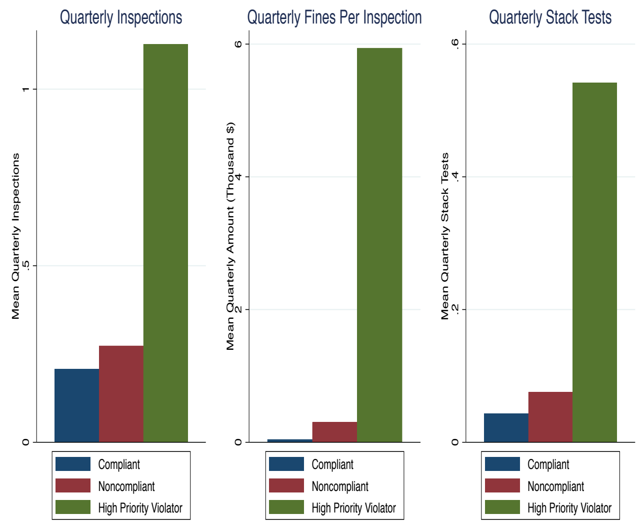
[43] Environmental Protection Agency USA. State review framework and integrated clean water act permit quality review florida. 2012. URL https://www.epa.gov/sites/production/files/2013-08/documents/srf-rd2-rev-fl.pdf.

[44] Environmental Protection Agency USA. National strategy for improving oversight of state enforcement performance. 2013. URL https://www.epa.gov/sites/production/files/2014-06/documents/state-oversight-strategy.pdf.

[45] Environmental Protection Agency USA. Revision on timely and appropriate enforcement response to high priority violations. 2014. URL https://www.epa.gov/sites/production/files/2015-01/documents/hpvpolicy2014.pdf.

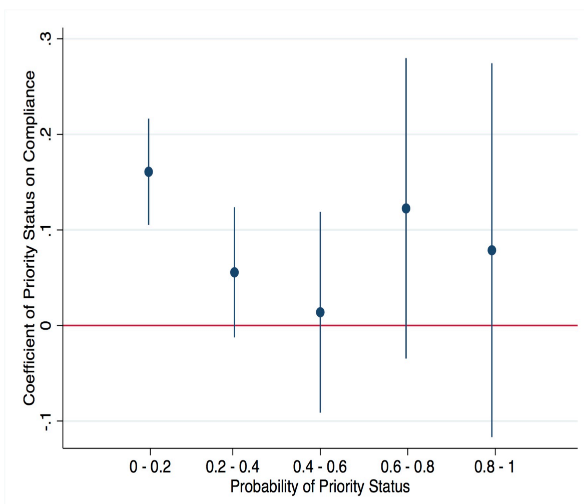
Figures and Tables

Figure 1: Differences in Enforcement by Regulatory Status



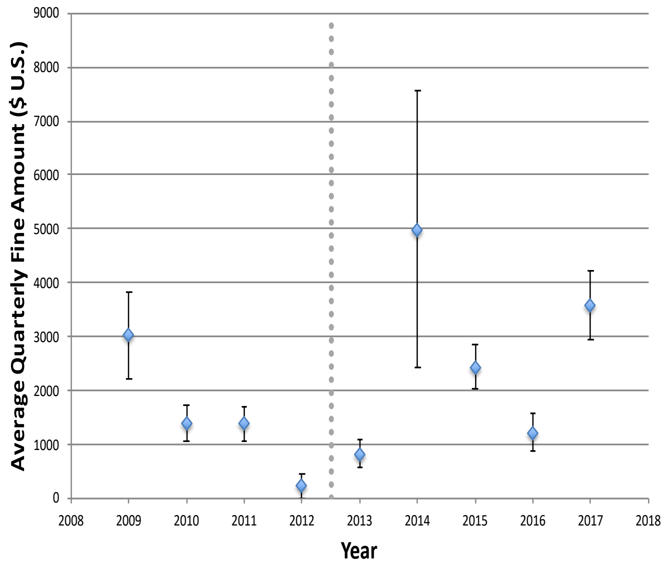
Notes: This figure shows the mean levels of enforcement activity by compliance status. Compliant or noncompliant but non-priority plants experience relatively low levels of monitoring (stack tests or inspections) and penalties as compared to plants with priority violations. The data comes from the AFS.

Figure 2: Priority Violator Compliance Effect by Predicted Probability of Classification



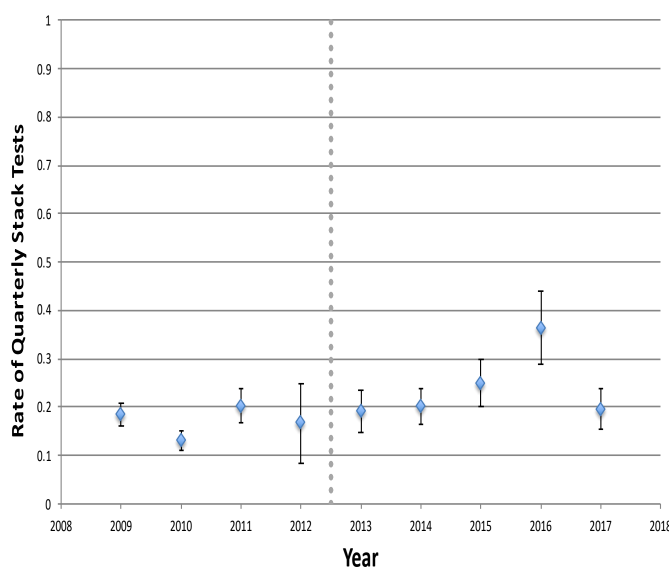
Notes: This figure plots correlation coefficients of the relationship between a noncompliant plant being classified as having a priority violation and whether they are compliant one year later. The reported coefficients are from five separate regression samples, stratified according to the predicted probability the noncompliant plant would be classified as having a priority violation. The predicted probability is based on the "status transition" regressions in Blundell, Gowrisankaran, and Langer (2018). There is a clear pattern of plants classified as having a priority violation being more likely to be in compliance one year later compared to the otherwise similar plants not classified as such. The regressions correspond to 2,524, 768, 257, 179, and 256 observations from the AFS data.

Figure 3: Fines for Priority Violation Plants in Florida



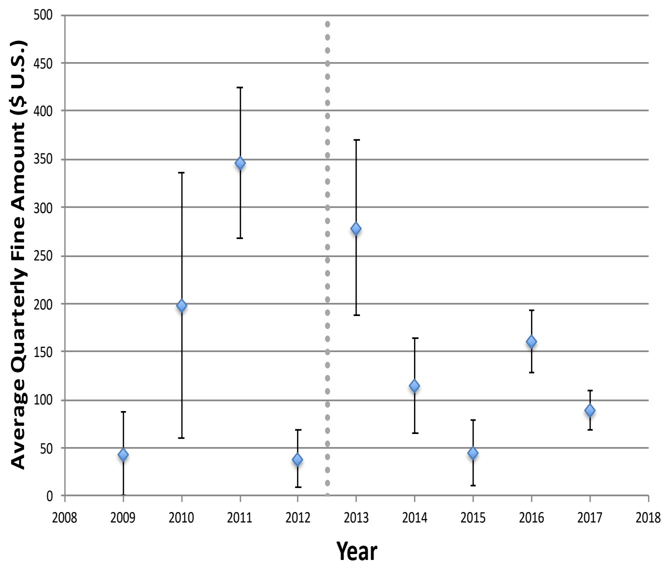
Notes: This figure plots the average amount of fines paid per quarter for Florida plants with priority violations over the 2009 - 2017 fiscal year time period, with the omission of the post review 2012 quarters. The data comes from the AFS and ICIS. The black bars represent the standard errors of the mean observations within the year.

Figure 4: Quarterly Stack Test Rate for Priority Plants in Florida



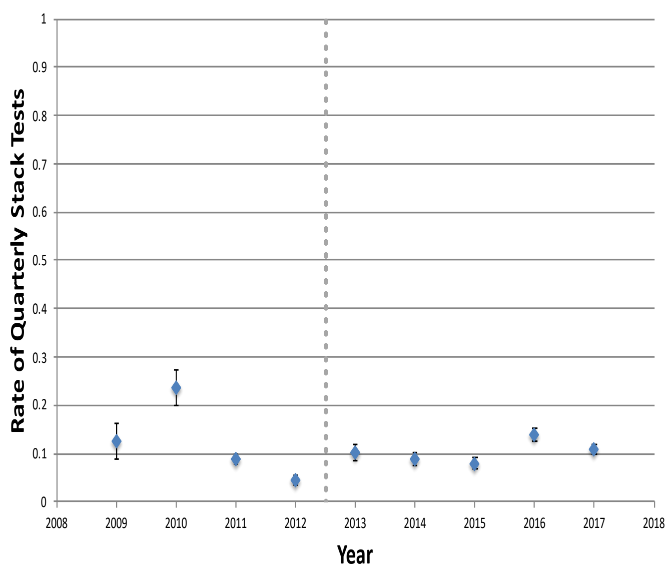
Notes: This figure plots the rate of plants having at least one-stack test per quarter for priority violation Florida plants over the 2009 - 2017 fiscal year time period, with the omission of the post review 2012 quarters. The black bars represent the standard errors of the mean observations within the year. The data comes from the AFS and ICIS.

Figure 5: Fines for Non-Priority Violation Plants in Florida



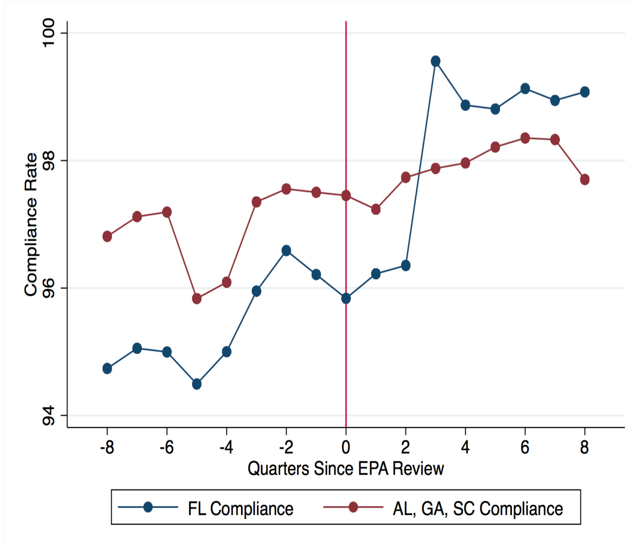
Notes: This figure plots the average amount of fines paid per quarter for Florida plants with Federally Reportable Violations but not priority violations over the 2009 - 2017 fiscal year time period, with the omission of the post review 2012 quarters. The black bars represent the standard errors of the mean observations within the year. The data comes from the ICIS.

Figure 6: Quarterly Stack Test Rate for Non-Priority Plants in Florida



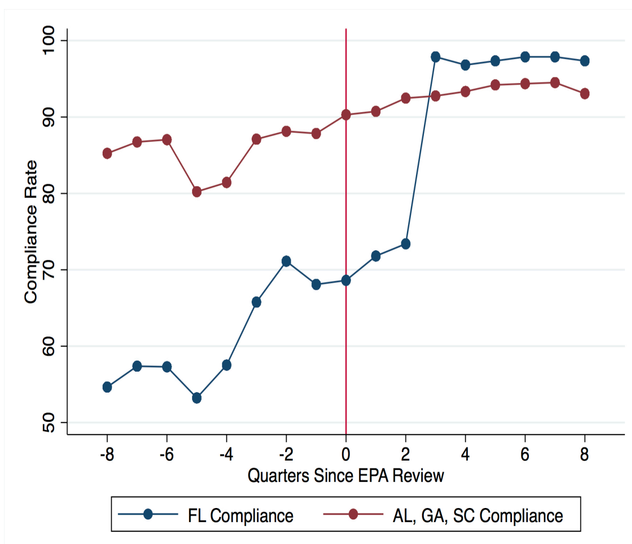
Notes: This figure plots the rate of plants having at least one-stack test per quarter for Florida plants with Federally Reportable Violations but not priority violations over the 2009 - 2017 fiscal year time period, with the omission of the post review 2012 quarters. The black bars represent the standard errors of the mean observations within the year. The data comes from the ICIS.

Figure 7: Compliance for all Florida Plants



Notes: This figure plots the average quarterly rate of observed compliance for all plants in Florida, Georgia, Alabama, and South Carolina over time. The red bar marks the quarter of the EPA's review of the Florida Department of Environmental Protection. Prior to the review, the Florida plant compliance rate had a comparable trend with other southern plants. Three quarters after the review, there is a large increase in the compliance rate for Florida plants. The data comes from the AFS.

Figure 8: Compliance for all Pre-Review Noncompliant Florida Plants



Notes: This figure plots the average quarterly rate of observed compliance for the previously noncompliant or priority violator subsample of plants in Florida, Georgia, South Carolina, and Alabama over time. Plants had to be in noncompliance or be a priority violator at least once prior to the 2012 review. The red bar marks the quarter of the EPA's review of the Florida Department of Environmental Protection. Prior to the review, Florida plant compliance had a comparable trend with other southern plants. Three quarters after the review, there is a large increase in Florida plant compliance. The data comes from the AFS.

TABLE I. Comparison of Changes in Priority Violator Enforcement Across States

	Mean Diff in Florida (1)	Mean Diff in GA,SC,AL (2)	Mean Diff in Rest of U.S. (3)	Florida vs. GA,AL,SC (4)	Florida vs. Rest of U.S. (5)
Penalties (U.S. $)	445.44	-925.71	-494.79	1,371.16	940.23
	(606.64)	(803.02)	(651.08)	(1,006.41)	(889.90)
Stack Tests	0.15	-0.15	-0.02	0.31	0.17
	(0.19)	(0.07)	(0.09)	(0.20)	(0.21)
Inspections	-0.28	-0.13	-0.47	-0.15	0.18
	(0.06)	(0.03)	(0.01)	(0.06)	(0.06)

Notes: This tables shows the differences in mean enforcement characteristics for priority violator plants pre and post the 2012 EPA review of the Florida Department of Environmental Protection. The timeframe is 2007 - 2017 using data from both the AFS and ICIS databases. Column (1) corresponds to 417 plant-quarter priority violator observations from Florida. Column (2) corresponds to 1,429 plant-quarter priority violator observations from Alabama, Georgia, and South Carolina. Column (3) corresponds to 15,033 plant-quarter priority violator observations from the rest of the U.S. Column (4) corresponds to the difference in differences of the changes in enforcement between Florida and the southern states. Column (5) corresponds to the difference in differences of the changes in enforcement between Florida and the rest of U.S.. All of these enforcement variables are quarter averages and scaled by the number of depreciated violations for the plant. Standard errors of these differences are in parentheses.

TABLE II. D.I.D. Regression Results: Impact of the EPA Review on Plant Compliance

	(1)	(2)	(3)	(4)
Panel A: Southern Plant Control Group				
FL·Post	0.004	0.001	0.094***	0.072*
	(0.003)	(0.002)	(0.019)	(0.038)
Industry Emissions	-0.014***	-0.005	-0.027***	-0.026***
	(0.003)	(0.003)	(0.008)	(0.009)
Nonattainment County	-0.011	-0.001	-0.081	-0.025
	(0.008)	(0.001)	(0.067)	(0.056)
County Establishments	0.003	-0.002**	0.028	0.018
	(0.000)	(0.000)	(0.000)	(0.001)
Panel B: Matched Plant Control Group				
FL·Post	0.022***	-0.002	0.188***	0.216***
	(0.007)	(0.007)	(0.022)	(0.043)
Industry Emissions	-0.009***	-0.004	-0.014***	-0.010***
	(0.003)	(0.006)	(0.005)	(0.004)
Nonattainment County	-0.009	-0.001	-0.142	-0.060
	(0.012)	(0.007)	(0.091)	(0.093)
County Establishments	-0.018**	-0.020***	-0.002	-0.075
	(0.007)	(0.007)	(0.038)	(0.050)
Plant Fixed Effects	Y	Y	Y	Y
Quarter Dummies	Y	Y	Y	Y

Notes: This table reports D.I.D regressions of next quarter plant-level compliance on whether the plant was in Florida in the post review period during the sample time frame of 2007 - 2014. The indicator FL·Post takes a value of one if the time period is after the second quarter of 2012 and the plant resides in Florida. "County Establishments" is scaled to be per 100 sites. "Industry Emissions" is scaled to be per 1,000 tons. Standard errors are clustered at the plant level. $***$ indicates significance at the 1% level, $**$ at the 5% level, and $*$ at the 10% level. **Sample:** Columns (1) to (4) only include plants that were in operation both prior to and after the review and were indicated as "Operating" in the 2014 AFS facilities file. Column (2) includes the subsample of 1,140 Florida plants (30,837 observations) with no history of noncompliance prior to the review. There are 2,200 control plants (57,324 observations) in Panel A and 1,321 control plants (40,951 observations) in Panel B. Column (3) corresponds to the subsample of 98 Florida plants (2,657 observations) with a history of noncompliance, but not priority violator status, prior to the review. There are 386 control plants (9,589 observations) in Panel A and 243 control plants (7,533 observations) in Panel B. Column (4) corresponds to the subsample of 48 Florida plants (1,422 observations) that were priority violators prior to the review. There are 182 control plants (4,533 observations) in Panel A and 51 control plants (1,581 observations) in Panel B. Column (1) includes all plants from the corresponding panel in Columns (2) to (4).

TABLE III. D.D.D. Results: Heterogeneous Impact of the EPA Review on Plant Compliance

	(1)	(2)	(3)
Panel A: Southern Plant Control Group			
FL·Post	0.123***	0.083***	0.076***
	(0.024)	(0.025)	(0.016)
FL·Post·Treat	-0.068*	0.007	0.141***
	(0.035)	(0.007)	(0.051)
Post·Treat	0.034*	-0.012**	0.002
	(0.021)	(0.005)	(0.057)
Industry Emissions	-0.031***	-0.031***	-0.027***
	(0.006)	(0.006)	(0.004)
Nonattainment County	-0.050	-0.059	-0.069
	(0.045)	(0.046)	(0.047)
County Establishments	0.027	-0.012	0.033
	(0.026)	(0.033)	(0.025)
Panel B: Matched Plant Control Group			
FL·Post	0.211***	0.223***	0.172***
	(0.022)	(0.023)	(0.018)
FL·Post·Treat	-0.059*	-0.011*	-0.035
	(0.036)	(0.007)	(0.107)
Post·Treat	0.026	0.003	0.072
	(0.022)	(0.004)	(0.072)
Industry Emissions	-0.012***	-0.012***	-0.012***
	(0.004)	(0.004)	(0.003)
Nonattainment County	-0.082	-0.079	-0.136**
	(0.070)	(0.072)	(0.068)
County Establishments	-0.026	-0.040	-0.036
	(0.029)	(0.032)	(0.029)
Plant Fixed Effects	Y	Y	Y
Quarter Dummies	Y	Y	Y

Notes: This table reports D.D.D regressions of next quarter plant-level compliance on whether the plant was in Florida in the post review period and had a differential compliance cost (Treat). The indicator Post takes a value of one if the time period is after the second quarter of 2012. The indicator FL takes a value of one if the plant resides in Florida. "County Establishments" is scaled to be per 100 sites. "Industry Emissions" is scaled to be per 1,000 tons. Column (3) includes a plants depreciated accumulated violations as an additional control. In Column (1) the indicator Treat takes a value of one if the plant's parent company has another facility outside the state of Florida. In Column (2) the indicator Treat corresponds to the number of manufacturing establishments in the county. In Column (3) the indicator Treat corresponds to the depreciated number of accumulated violations since the plant was last in compliance. Standard errors are clustered at the plant level. ∗∗∗ indicates significance at the 1% level, ∗∗ at the 5% level, and ∗ at the 10% level. **Sample:** Columns (1) to (3) only includes plants that were observed both prior to and after the review and indicated as "Operating" in the 2014 AFS facilities file. All three columns correspond to the subsample of 146 Florida plants (4,079 observations) with a history of noncompliance or priority violator status prior to the review. There are 568 control plants (14,122 observations) in Panel A and 294 control plants (9,114 observations) in Panel B.

TABLE IV. Results for Compliance Mechanisms Southern Plant Control Group

	(1)	(2)	(3)	(4)
Panel A: Log of Annual Emissions				
FL·Post	-0.310***	-0.318***	-0.115	-0.476**
	(0.044)	(0.093)	(0.096)	(0.206)
FL·Post·Treat				0.501**
				(0.225)
Post·Treat				-0.119
				(0.205)
Panel B: Evaluation Failure Rate				
FL·Post	-0.000	-0.021*	0.028	-0.011
	(0.003)	(0.012)	(0.022)	(0.017)
FL·Post·Treat				0.029
				(0.032)
Post·Treat				-0.006
				(0.011)
Plant Fixed Effects	Y	Y	Y	Y
Time Dummies	Y	Y	Y	Y
Additional Controls	Y	Y	Y	Y

Notes: This table reports D.I.D and D.D.D regressions of two environmental outcomes (denoted by the panel title) on whether the plant was in Florida in the post review period and in the lower cost (Treat) group during the sample timeframe of 2007 - 2014 (Panel B omits the year 2012). The indicator Post takes a value of one if the time period is after the second quarter of 2012. The indicator FL takes a value of one if the plant resides in Florida. In Column (4) the indicator Treat takes a value of one if the plant's parent company has one additional facility outside the state of Florida. All columns use average annual plant emissions at the industry level, the nonattainment status of the county where the plant resides, and the number of manufacturing establishments in the county where the plant resides as additional controls. Standard errors are clustered at the plant level. $***$ indicates significance at the 1% level, $**$ at the 5% level, and $*$ at the 10% level. **Sample:** The estimation samples only includes plants that were observed both prior to and after the review and indicated as "Operating" in the 2014 AFS facilities file. Column (4) contains all plants in Columns (2) - (3). Column (3) examines plants that were priority violators prior to the review. Panel A includes 42 Florida plants (378 observations) along with 110 control plants (990 observations). Panel B includes 48 Florida plants (626 observations) along with 181 control plants (2,037 observations). Column (2) examines plants that were noncompliant but not priority violators prior to the review. Panel A includes 52 Florida plants (468 observations) along with 120 control plants (1,080 observations). Panel B includes 97 Florida plants (373 observations) along with 349 control plants (2,579 observations). Column (1) examines plants that were always compliant prior to the review. Panel A includes the 488 Florida plants (4,392 observations) along with 450 control plants (4,050 observations). Panel B includes 1,140 Florida plants (5,197 observations) along with 2,066 control plants (11,971 observations).

Appendix

A1. Theory

Here I start with the steady state compliance rate under the Harrington (1988) framework, but with a modified transition probability u, which is equal to:

$$\int_0^{\infty} g(c)f(c)dc = F(P_1 \cdot F_1) + \frac{P_1 \cdot u}{P_1 \cdot u + P_2} \cdot [F(P_2 \cdot F_2 + \frac{P_2 \cdot \beta(P_2 \cdot F_2 - P_1 \cdot F_1)}{[1-(1-P_1 \cdot u)\beta]}) - F(P_1 \cdot F_1)] \quad (1)$$

The first term, $F(P_1 \cdot F_1)$, represents plants with low compliance costs that always comply. The second term represents plants with medium compliance costs and the proportion of time they comply. Where, $1 - \frac{P_1 u}{P_1 u + P_2}$, is the steady state probability that a medium cost plant, having a compliance cost less than or equal to L_1 but greater than $P_1 \cdot F_1$, is observed in G_1. Using these equations, I consider the impact of an increase in group two penalties and monitoring in the following subsection.

Since it is never optimal for a plant to comply in G_1, and not comply while in G_2 (choose $\{Comply,\ Don't\,Comply\}$) because plants with lower costs will be better off choosing to always comply, while higher cost plants will be better off choosing to never comply. I solve for the cutoff (L_1) where a plant is indifferent between choosing the policy $\{Don't\,Comply,\ Don't\,Comply\}$ and $\{Don't\,Comply,\ Comply\}$ by setting the calculated G_2 present values for each of those policies equal and solving for the compliance cost c. As stated in the text this cutoff is:

$$L_1 = P_2 \cdot F_2 + \frac{P_2 \cdot \beta(P_2 \cdot F_2 - P_1 \cdot F_1)}{[1-(1-P_1 \cdot u) \cdot \beta]} \quad (2)$$

Thus, plants with compliance costs exceeding that threshold (high cost plants) will be better off never complying and will therefore always be observed in group G_2 within the steady state. Medium cost plants, plants with compliance costs less than or equal to L_1 but greater than $P_1 \cdot F_1$, will alternate between groups with the steady state probability of being observed in G_1 as follows:

$$Pr(G_1 | P_1 \cdot F_1 < c \leq L_1) = 1 - \frac{P_1 u}{P_1 u + P_2} \tag{3}$$

Finally, low cost plants will always be observed in group G_1 within the steady state. Given the solved probabilities, steady state compliance under dynamic enforcement is equal to:

$$\int_0^\infty g(c)f(c)dc = F(P_1 \cdot F_1) + \frac{P_1 \cdot u}{P_1 \cdot u + P_2} \cdot [F(P_2 \cdot F_2 + \frac{P_2 \cdot \beta(P_2 \cdot F_2 - P_1 \cdot F_1)}{[1 - (1 - P_1 \cdot u)\beta]}) - F(P_1 \cdot F_1)] \tag{4}$$

The first term, $F(P_1 \cdot F_1)$, represents plants with low compliance costs that always comply. The second term represents plants with medium compliance costs and the proportion of time they comply. Using these equations, I can make predictions regarding the effect of increased enforcement of priority group plants. My empirical application examines an increase in the penalties and monitoring of plants in the most severe priority noncompliance classification. This model predicts a number of potentially observable effects under this change in regulatory policy.

A.1.2 The Effects of Increases in Group Two Monitoring and Penalties

I motivate the four predictions in the main empirical text from the following assumptions. Now assume there are increases in the level of penalties and monitoring for group two plants from F_2 to $F_2^{'}$ and from P_2 to $P_2^{'}$. In this case, observed plant behavior will change on a number of margins.

Proposition 1. *The increase in penalties for group two plants will increase the overall compliance rate.*

The second term of equation (3) above shows how an increase in the priority group penalties (F_2) will increase the total number of plants that are better off choosing the policy $\{Don't\,Comply,\, Comply\}$ and sometimes be in compliance, over the policy of never being in compliance. These firms will choose to comply because the future value of moving to the lower penalty and monitoring regime in group one is relatively higher. The second proposition addresses the transition and group composition of these plants.

Proposition 2. *The increase in penalties and monitoring for group two plants leads to some high costs plants, those only observed in group two in the original steady state, to transition to group one.*

Under an increase in penalties from F_2 to F_2' and an increase in monitoring from P_2 to P_2' the cutoff between medium and high cost plants L_1 will increase to $L_1^{'}$. Once these formerly high cost plants comply they will transition to group one. However, formerly high costs plants are not the only plants whose compliance behavior will be affected, the observed behavior of formerly medium cost plants will change as well.

Proposition 3. *The increase in monitoring for group two plants will increase the rate that formerly medium cost plants are observed in group one.*

This stems from the fact that the steady state probability for a medium cost plant to be observed in group one, $1 - \frac{P_1 u}{P_1 u + P_2}$, is increasing priority group monitoring (P_2). Originally, these plants alternated their choice of compliance based on their group classification so an increase in their monitoring while in the priority group will speed up their transition back to the non-priority group. Additionally, the increased penalties for plants in the priority group will not alter this prediction since medium cost plants never pay the group two penalties in equilibrium. This last point about medium cost plants not paying the group two penalties in equilibrium also applies to low cost plants, leading to the fourth proposition.

Proposition 4. *The increase in penalties and monitoring for group two plants will have a larger impact on formerly medium and high cost plants as compared to low cost plants.*

The compliance cost cutoff for a low cost plant, which always chooses to comply, is a function of non-priority group penalties and monitoring $(P_1 \cdot F_1)$. This implies low cost plants should exhibit no change in their behavior following an increase in priority group penalties or monitoring. Since both Proposition 2 and Proposition 3 indicate a change in the compliance behavior of formerly medium and high cost plants, this will exceed the expected response by low cost plants.

A1.3 Conditions for Ambiguity Between Medium and High Cost Plants

To demonstrate how high cost plants should exhibit the largest compliance increases after the policy change, I derive conditions for when medium cost plants would experience a larger increase in the probability that they are in group one.

-Assume there are increases in the level of penalties and monitoring for group two plants from F_2 to $F_2^{'}$ and from P_2 to $P_2^{'}$.

-I use the following condition for the rate at which medium cost plants are observed in group 1 as follows:

$$Pr(G_1|P_1 \cdot F_1 < c \leq L_1) = 1 - \frac{P_1 \cdot u}{P_1 \cdot u + P_2} \tag{5}$$

For simplicity it is easier to rewrite this as:

$$Pr(G_1|P_1 \cdot F_1 < c \leq L_1) = \frac{P_2}{P_1 \cdot u + P_2} \tag{6}$$

This means that the change in the group one rate of these medium cost plants can be written as:

$$[\frac{P_2'}{P_1 \cdot u + P_2'}] - [\frac{P_2}{P_1 \cdot u + P_2}] \tag{7}$$

Which is equal to:

$$\frac{P_1 \cdot u \cdot (P_2' - P_2)}{(P_1 \cdot u + P_2) \cdot (P_1 \cdot u + P_2')} \tag{8}$$

For the high cost plants. The change in the group one rate is going to be the steady state probability of being a medium cost plant under the new policy multiplied by the probability a high cost plant transitioned to being a medium cost plant. In this case, the probability a high cost plant transitions to being medium cost is:

$$\frac{F(L_1') - F(L_1)}{1 - F(L_1)} \tag{9}$$

I then multiply this by the steady-state group one rate for medium cost plants (equation (6)) to get the change:

$$[\frac{P_2'}{P_1 \cdot u + P_2'}] \cdot \frac{F(L_1') - F(L_1)}{1 - F(L_1)} \tag{10}$$

Thus, the change for originally medium cost plants is higher than the increase for originally high cost plants if the difference between equation (8) and equation (10) is positive.

For simplicity, I rewrite this difference to be the following condition:

$$\frac{1 - \frac{P_2}{P_2'}}{1 + \frac{P_2}{P_1 \cdot u}} > \frac{F(L_1') - F(L_1)}{1 - F(L_1)} \tag{11}$$

Here, the right hand side or L_1 terms are a function of both penalties and monitoring. While the left hand side is dependent on just the changes in monitoring. This inequality indicates that high cost plants should have the largest compliance increase as long as the proportional increase in monitoring is small, $\frac{P_2}{P_2'}$, or the penalty increase is large with no significant right tail in plant compliance costs making the distributional increase $F(L_1') - F(L_1)$ large. Both of these conditions seem plausible given that monitoring for Florida priority violators was relatively high prior to the review and there was a large increase in penalties after the review. In addition, there was no Florida plant that was a priority violator during the entire sample period, indicating there is no significant right tail in Florida plant compliance costs that would make the right hand side of equation (11) small.

A2. Data

A2.1 AFS Data

For my empirical analysis, I primarily use the AFS data from the EPA's Environmental Compliance History Online (ECHO) database. The AFS data is comprised of five datasets: (1) the *AFS Facilities* dataset, (2) the *AFS Action*s dataset, (3) the *AFS Air Programs* dataset, (4) the *AFS HPV History* dataset, and (5) the *AFS Air Program Historical Compliance* dataset. The data contain information for 37,000 different manufacturing facilities from the fourth quarter of 2006 to the end of the third quarter of 2014, or fiscal years 2007 - 2014. I describe their use as follows.

First, the *AFS Facilities* dataset is a list of all plants. The dataset contains information on the facility name, address, industry, AFS id code (for linking with other datasets), and an operating status code.

Second, the *AFS Actions* dataset records the history of regulatory actions for each facility. The unit of observation in this dataset is by AFS id, action, and date. These actions are primarily characterized as either inspections, penalties, stack tests, and warnings.

Third, the *AFS Air Programs* dataset contains data on what air programs each plant falls under as well as which pollutants from the plant dictate the air program designation. Since this data contains both plant id and AFS id, it allows me to form the key linkage file to aggregate the actions data and other data, which are at the air program level, to the plant-quarter level for the final analysis.

Fourth, the *AFS Air Program Historical Compliance* dataset records quarterly compliance information by AFS id, air program, and quarter. The main text describes how I follow the literature and treat this self-reported data as accurate.

Finally, the *AFS HPV History* dataset records the date each plant enters and exits High Priority Violator status. This is a direct measure of whether a plant was classified as a Priority Violator (the most severe form of noncompliance) in a given quarter. Since the

AFS version of this dataset has missing 2014 data, due to some states failing to report this information prior to the switch to the ICIS in October 2014, I use the analogous ICIS HPV history for 2014 to complete the plant-quarter Priority Violation history used in the main analysis. I describe the merging between the AFS and ICIS data later in this section.

A2.2 AFS Air Emissions Data

To examine environmental outcomes related to enforcement activity, I obtained information on annual plant emissions from the AFS Air Emissions dataset. The dataset is a compilation of facility level emissions data from multiple databases, including the Greenhouse Gas Reporting Program (GHGRP), Clean Air Markets, National Emissions Inventory (NEI), and Toxic Release Inventory (TRI). The pollutants I use for this analysis come from the latter two databases.

I merge the emissions data to the AFS using a multi-step procedure. First, I merge the lists using a linkages file provided by the EPA.

Second, I use a multi-step matching procedure between the NEI and the AFS using the plant level characteristics in both.

1. First, I merge the lists using a linkages file provided by the EPA.

2. I merge plants based on whether they have an exact match on name and street address.

3. I merge plants based on whether they have an exact match on street address and zip code.

4. I look for cases where there is only one unique industry observation (six digit NAICS-code) within a five digit zip code found in both datasets and match the plants accordingly.

5. I conduct a fuzzy matching on plant-name, six digit NAICS-code, plant street address, and five digit zip code with the restriction that there must be an exact match on city.

For the fuzzy matching I use a biagram string comparator which assigns a probabilistic match score reflecting the degree of difference between two match strings. To minimize the number of false positives in the matching I only consider matches with a score above

.85 (with one being an exact match). To place a greater weight on plant name and street address, I perform the fuzzy matching on those variables first, then remove any new matches before including the five digit zip code and six digit NAICS-code in matching the residual observations.

Match failures can occur for three reasons: (1) identifying plant information is not harmonized across the two datasets; (2) the plant did not meet the Air Emissions Reporting rules, a different set of standards than the AFS; and (3) there are missing observations.

At the end of the matching process, I have matched 24,873 plants out of 50,485 emissions plants to the 37,000 AFS plants. Table AV provides a comparison of the regulatory characteristics for matched and unmatched Florida plants. The typical unmatched plant appears to be subject to less regulatory activity than matched plants. This indicates that the main results regarding pollution reductions are characteristic of plants more likely to be subjected to regulatory enforcement, which reinforces the findings that environmental outcomes improved more for plants with greater compliance costs.

A2.3 ICIS Data

In October 2014 the AFS was retired and the EPA began using the ICIS-AIR system to track and record CAA violations and enforcement information. The system contains plant-level data on inspections, stack tests, stack test results, inspection results, warnings, and penalties that is analogous to the AFS enforcement data. However, compliance information in the ICIS differs from the AFS by only recording instances of federally reportable violations, a more coarse and severe definition of noncompliance, in addition to High Priority Violations. Since this compliance data does not contain a comparable outcome variable to the AFS definition of compliance, I exclude it from the main empirical estimation but provide it as a robustness check in Table X of the appendix. By incorporating the ICIS data, I am able to document that the change in the Florida Department of Environmental Protection's enforcement practices discussed in the main text continued for a number of years after the retirement of the AFS data in 2014.

The EPA's Integrated Compliance Information System Air database is the continuation of the EPA's air enforcement tracking efforts after the AFS was retired in October of 2014. The ICIS is comprised of seven datasets: (1) the *ICIS Air Facilities* dataset, (2) the *ICIS Air Informal Actions* dataset (warnings), (3) the *ICIS Air Formal Action*s dataset (fines or penalties), (4) the *ICIS Air Stack Tests* dataset (stack tests), (5) the *ICIS AIR Full Compliance Evaluations and Partial Compliance Evaluations* dataset (inspections), (6) the *ICIS Air Programs* dataset, and (7) the *ICIS Air Violation History* dataset. The data contains plant violation and enforcement information from as early as 1975 to the fourth quarter of 2017, although I only use this data to augment my primary analysis to include fiscal years 2015, 2016, and 2017.

The unit of observation in the ICIS data is at the program system id and day level, where a program system id corresponds to a plant-air program combination. Therefore, I compile these data by constructing a linkage between the plant id (FRS number) and the program system id using the *Air Programs* dataset. I then use the linkage file to construct a dataset on plant compliance, inspections, warnings, penalties, and stack tests at the plant-quarter level that is similar to the primary AFS enforcement dataset. The key difference between the AFS and the ICIS is the definition of compliance. The ICIS only records High Priority Violations and Federally Reportable Violations which are more severe forms of noncompliance than some of the AFS compliance codes which would result in a plant being classified as noncompliant in my primary analysis. However, I still formulate a definition of quarterly noncompliance, based on whether a plant has one of these more severe violations at any point in the quarter, to check for robustness of the main results. I am able to merge 85.4% of the plants in the primary dataset with the ICIS data by using a linkage file between FRS number and AFS ids provided by the EPA.

A2.4 National Organization File

To obtain information on the location and number of facilities under a plant's parent company I use the Facility Registry System's "National Organization File." The dataset contains information on a facility's "Organization Name," the "Affiliation Type" (whether the organization is the parent company), as well as the Dun and Bradstreet identification number or "DUNS Number." I merge the organization information to the AFS using a multi-step procedure. First, I merge the plant's that matched to the "National Organization File" using their frsnumber and then used the "DUNS Number" to associate those plants to their respective parent company.

Second, I use a string matching between the plant names in the AFS and the organization names in the "National Organization File" with a high tolerance of 0.95 to classify a match. Then, I associate these string matched plants to their parent company using the corresponding "DUNS Number" in the "National Organization File." Finally, I record the number of non-Florida plant's for each parent company before merging this count to plants in the AFS data.

A2.5 Other Data

I supplement the analysis with data from the Census of Manufacturers, the EPA's Greenbook, and the Bureau of Labor Statistics. I formulate a control variable for county manufacturing activity by using the 2007 and 2012 publicly available versions[1] of the Census of Manufacturers to obtain the number of manufacturing facilities operating in each county during the sample timeframe. I use the EPA's Greenbook to create a panel of annual county nonattainment status. The county nonattainment status variable takes a value of one if the county is in nonattainment for any of the NAAQS emissions standards. Finally, I merge the Bureau of Labor Statistics county-year panel of each county's unemployment rate with the primary AFS data to examine the robustness of the main results as discussed in the following section.

[1] Versions EC0731A1 and EC1231A1.

A3. Robustness

A3.1 Permutation Tests

Table AVIII shows permutation tests of the primary DID results using southern control plants. I present the original standard errors clustered at the plant level, as well as bootstrapped standard errors and standard errors clustered at the state level. Bootstrapping was done by sampling plants and all corresponding plant-quarter observations with replacement. The statistical significance of the primary results remains consistent using these alternative standard errors.

A3.2 Falsification Tests

To explore the validity of the estimation procedure I estimate a series of 49 falsification tests. For each test, I redefine the $FL \cdot Post$ term from equation (2) in the main text to take a value of one if a plant resides in that particular state after the second quarter of 2012 and zero otherwise, cycling through all 49 states (excluding Florida). Each test uses the whole sample of U.S. manufacturing plants, making the plants in the non-treated states in each test the control group. As anticipated, only four of the test coefficients were positive and statistically significant at the 10% level, which is in line with the number of false positives one should expect under a valid test. This provides evidence that the primary results are not driven by the choice of estimation procedure.

A3.3 Impact of the Great Recession

Table AIX shows sensitivity results examining the impact of the Great Recession on plant compliance. The variable Treat takes a value of one if the plant resides in a county that had an unemployment rate higher than 10% during the great recession. In all specifications, none of the $FL{\cdot}Treat{\cdot}Post$ terms are statistically significant at conventional levels. This indicates that Florida plants residing in counties with Great Recession unemployment rates exceeding

10% did not have any differential trend in their post review compliance. In addition, the point estimates on the $FL \cdot Post$ are qualitatively similar to the Table II results in the main text. For the quadruple difference estimates the variable High takes a value of one if the plant has a parent company with one additional facility outside the state of Florida. Overall, differences in the recovery conditions of Florida manufacturing plants does not explain the vast increase in plant compliance following the EPA's 2012 review.

A3.4 Extended Sample with ICIS Data

Table AX shows the DID and DDD results using the ICIS data to extend the sample timeframe. I omit the inclusion of the additional controls (county manufacturing intensity, nonattainment, and county unemployment) used in the main analysis as this information isn't available for 2014 - 2017 time period. The estimated treatment effect remains positive and statistically significant with the inclusion of the additional observations. These results provide evidence that the findings in the main analysis are not driven by the limited number of post review quarters in the AFS data.

A3.5 Placebo Test

Table AIII shows the DID and DDD results from a placebo test. I redefine the *Post* term in equations (2) and (3) from the main text to take a value of 1 for starting in the second quarter of 2010 because this gives me an identical number of post treatment quarters as the original analysis since I don't include quarters after the EPA review. As anticipated, the estimated treatment effects are statistically indistinguishable from 0 which provides evidence of no difference in compliance pre trends between Florida plants and their southern control counterparts.

Appendix Tables

TABLE AI. Pre-Treatment Summary Statistics by Group

	Florida Mean	Florida SE	GA,AL,SC Mean	GA,AL,SC SE	Matched Mean	Matched SE
Panel A: Sample Three						
Compliance	0.76	0.013	0.79	0.0071	0.76	0.012
Warnings	0.057	0.0075	0.043	0.0036	0.018	0.0044
Penalty (U.S. $)	655.4	176.1	653.2	110.3	234.9	171.6
Inspections	0.80	0.038	0.54	0.016	0.35	0.031
Stack Test Rate	0.16	0.011	0.072	0.0045	0.035	0.0054
Industry Emissions	263.4	29.7	175.9	10.3	277.1	57.8
Nonattainment	0.067	0.0077	0.18	0.0067	0.22	0.012
Establishments	505.5	11.1	156.5	3.07	452.5	9.44
Initial Year	1997.4	0.17	1995.6	0.16	1998.7	0.18
Outside FL	0.40	0.015	0.78	0.0072	0.76	0.012
Observations	1,051		3,317		1,173	
Panel B: Sample Two						
Compliance	0.78	0.0093	0.84	0.0044	0.92	0.0036
Warnings	0.0041	0.0016	0.011	0.0012	0.0075	0.0013
Penalty (U.S. $)	0.99	0.78	37.6	10.7	11.1	5.47
Inspections	0.074	0.0093	0.21	0.0074	0.14	0.0067
Stack Test Rate	0.0056	0.0017	0.020	0.0017	0.0029	0.00071
Industry Emissions	106.1	17.1	78.5	4.46	62.5	1.45
Nonattainment	0.040	0.0044	0.25	0.0051	0.037	0.0025
Establishments	410.9	8.55	186.2	2.47	148.2	3.44
Initial Year	2007.1	0.12	1997.3	0.11	2006.1	0.049
Outside FL	0.24	0.0097	0.79	0.0049	0.51	0.0067
Observations	1,970		7,052		5,589	
Panel C: Sample One						
Compliance	1	0	1	0	0.89	0.0018
Warnings	0	0	0.00027	0.000083	0.012	0.00065
Penalty (U.S. $)	0	0	0.57	0.28	70.3	16.7
Inspections	0.067	0.0024	0.099	0.0021	0.16	0.0031
Stack Test Rate	0.0048	0.00046	0.0042	0.00032	0.011	0.00060
Industry Emissions	49.2	1.40	59.1	1.67	44.9	1.17
Nonattainment	0.035	0.0012	0.18	0.0019	0.046	0.0012
Establishments	454.4	3.14	187.7	0.95	205.1	2.05
Initial Year	2005.9	0.038	1999.1	0.044	2004.6	0.028
Outside FL	0.26	0.0029	0.69	0.0023	0.51	0.0029
Observations	22,915		40,052		30,383	

Notes: This tables shows the pre-treatment summary statistics for the estimation samples of Florida plants, southern control plants, and the matched control plants from the AFS data. The variable "Initial Year" is the first year the plant was observed in the ICIS data. The variable "Outside FL" is an indicator that takes a value of one if the plant's parent company has an additional facility that is located outside the state of Florida.

TABLE AII. D.I.D. Regression Results: Table II Without any Additional Controls

	(1)	(2)	(3)	(4)
Panel A: Southern Plant Control Group				
FL·Post	0.002	0.002	0.084***	0.065*
	(0.003)	(0.001)	(0.018)	(0.033)
Panel B: Matched Plant Control Group				
FL·Post	0.022***	-0.002	0.188***	0.216***
	(0.007)	(0.007)	(0.022)	(0.043)
Plant Fixed Effects	Y	Y	Y	Y
Quarter Dummies	Y	Y	Y	Y

Notes: This table reports D.I.D regressions of next quarter plant-level compliance on whether the plant was in Florida in the post review period over the time period of 2007 - 2014. This is the same as the specifications presented in Table II of the main text but without any additional controls. The indicator FL·Post takes a value of one if the time period is after the second quarter of 2012 and the plant resides in Florida. Standard errors are clustered at the plant level. $***$ indicates significance at the 1% level, $**$ at the 5% level, and $*$ at the 10% level. **Sample:** Columns (1) to (4) only includes plants that were in operation both prior to and after the review and were indicated as "Operating" in the 2014 AFS facilities file. Column (2) includes the subsample of 1,140 Florida plants (30,837 observations) with no history of noncompliance prior to the review. There are 2,200 control plants (57,324 observations) in Panel A and 1,321 control plants (40,951 observations) in Panel B. Column (3) corresponds to the subsample of 98 Florida plants (2,657 observations) with a history of noncompliance, but not priority violator status, prior to the review. There are 386 control plants (9,589 observations) in Panel A and 243 control plants (7,533 observations) in Panel B. Column (4) corresponds to the subsample of 48 Florida plants (1,422 observations) that were Priority Violators prior to the review. There are 182 control plants (4,533 observations) in Panel A and 51 control plants (1,581 observations) in Panel B. Column (1) includes all plants from the corresponding panel in Columns (2) to (4).

TABLE AIII. Placebo Tests: Treatment 8 Eight Quarters Prior to the 2012 Review

	(1)	(2)	(3)	(4)
Panel A: Southern Plant Control Group				
FL·Post	0.058	0.056	0.080	0.056
	(0.036)	(0.049)	(0.050)	(0.046)
FL·Post·Treat				0.008
				(0.073)
Post·Treat				-0.025
				(0.029)
Industry Emissions	-0.027***	-0.028***	-0.024**	-0.030***
	(0.007)	(0.007)	(0.010)	(0.006)
Nonattainment County	-0.088	0.069	-0.143	-0.091
	(0.101)	(0.190)	(0.131)	(0.105)
County Establishments	-0.012	0.010	-0.051	-0.007
	(0.039)	(0.050)	(0.084)	(0.042)
N	12,829	8,656	4,173	12,829
Plant Fixed Effects	Y	Y	Y	Y
Quarter Dummies	Y	Y	Y	Y

Notes: This table reports the results of a placebo test of the southern control plant DID and DDD specifications over the time period of 2007 - 2012 (just prior to the review). The indicator Post takes a value of one if the time period is after the first quarter of 2010. The indicator FL takes a value of one if the plant resides in Florida. "County Establishments" is scaled to be per 100 sites. "Industry Emissions" is scaled to be per 1,000 tons. In Column (4) the indicator Treat takes a value of one if the plant's parent company has at least one additional non-Florida facility. Standard errors are clustered at the plant level. $***$ indicates significance at the 1% level, $**$ at the 5% level, and $*$ at the 10% level. **Sample:** Columns (1) to (3) only includes plants that were observed both prior to and after the review and indicated as "Operating" in the 2014 AFS facilities file and had a history of noncompliance or priority violator status prior to the 2012 review. Columns (1) and (4) corresponds to all plants in Columns (2) and (3). Column (2) corresponds to the subsample of plants with a history of noncompliance but not priority violator status prior to the 2012 review. Column (3) corresponds to the subsample of plants with a history of priority violator status prior to the review.

TABLE AIV. D.I.D. Regression Results: Alternative Samples

	(1)	(2)	(3)	(4)
Panel A: All Plant Observations				
FL·Post	0.006*	0.000	0.113***	0.064*
	(0.003)	(0.001)	(0.016)	(0.036)
Industry Emissions	-0.012***	-0.003	-0.027***	-0.027***
	(0.002)	(0.003)	(0.007)	(0.008)
Nonattainment County	-0.012*	-0.004	-0.053	-0.042
	(0.007)	(0.003)	(0.058)	(0.048)
County Establishments	0.004*	-0.001	0.032	0.012
	(0.003)	(0.001)	(0.023)	(0.047)
N	141,444	118,647	15,687	7,110
Panel B: Known Compliance Observations				
FL·Post	0.020***	0.001	0.298***	0.065*
	(0.005)	(0.002)	(0.036)	(0.039)
Industry Emissions	-0.014***	-0.005	-0.025***	-0.026***
	(0.003)	(0.003)	(0.007)	(0.009)
Nonattainment County	0.003	-0.002	0.110*	-0.028
	(0.009)	(0.002)	(0.065)	(0.057)
County Establishments	0.004	-0.002**	0.010	0.015
	(0.005)	(0.001)	(0.039)	(0.067)
N	90,714	73,790	11,178	5,746
Plant Fixed Effects	Y	Y	Y	Y
Quarter Dummies	Y	Y	Y	Y

Notes: This table reports D.I.D regressions of next quarter plant-level compliance on whether the plant was in Florida in the post review period over the time period of 2007 - 2014. "County Establishments" is scaled to be per 100 sites. "Industry Emissions" is scaled to be per 1,000 tons. Specifications are identical to those in Panel A of Table II in the maint text. The indicator FL·Post takes a value of one if the time period is after the second quarter of 2012 and the plant resides in Florida. Standard errors are clustered at the plant level. $***$ indicates significance at the 1% level, $**$ at the 5% level, and $*$ at the 10% level. **Sample:** Panel A corresponds to all Florida and southern control plants observed in the AFS data over the 2007 - 2014 time period. Panel B restricts the southern control group samples in Panel A of Table II in the main text to only those plants whose compliance status was directly reported.

TABLE AV. Environmental Outcomes Sample Comparisons

	Outside Sample Mean	Outside Sample SE	Sample Mean	Sample SE
Panel A: Air Emissions Analysis				
Compliance	0.98	0.00055	0.96	0.0011
Warnings	0.0020	0.00017	0.0076	0.00052
Penalty (U.S. $)	15.9	3.01	85.9	14.5
Inspections	0.071	0.0013	0.26	0.0040
Stack Test Rate	0.0053	0.00026	0.026	0.00091
Industry Emissions	47.3	1.00	117.5	2.84
Nonattainment	0.12	0.0012	0.15	0.0021
Establishments	266.9	1.20	262.6	1.75
Initial Year	2002.7	0.029	1997.0	0.052
Observations	76,533		29,829	
Panel B: Evaluated Failure Analysis				
Compliance	0.98	0.00048	0.91	0.0031
Warnings	0.0014	0.00012	0.028	0.0019
Penalty (U.S. $)	20.5	3.79	207.0	37.5
Inspections	0	0	1.41	0.0095
Stack Test Rate	0	0	0.14	0.0037
Industry Emissions	55.7	0.80	195.4	9.57
Nonattainment	0.13	0.0011	0.15	0.0039
Establishments	270.7	1.05	208.7	2.94
Initial Year	2001.6	0.027	1994.7	0.087
Observations	97,780		8,582	

Notes: This table compares the summary statistics for plants used in the analysis for Table IV in the main text for plants outside those samples but among the plants used in the southern control group DID and DDD estimations. Panel A compares the plants that were matched with the "Air Emissions" data with plants that were not matched and therefore not used in the analysis of plant emissions in Table IV of the main text. Panel B compares evaluated plants to plants that are not evaluated and therefore not in the evaluation failure analysis in Table IV of the main text. The variable "Initial Year" is the first year the plant was observed in the ICIS data. The differences observed in Panel A indicate plants that are older, from more polluting industries, and with a greater history of noncompliance, are more likely to be in the emissions dataset. The differences observed in Panel B indicate that evaluated plants are more likely to be noncompliant, are older, and more likely to be from polluting industries.

TABLE AVI. D.D.D. Regression Results: Table III Without any Additional Controls

	(1)	(2)	(3)
Panel A: Southern Plant Control Group			
FL·Post	0.113***	0.082***	0.063***
	(0.023)	(0.025)	(0.015)
FL·Post·Treat	-0.067*	0.005	0.138***
	(0.035)	(0.006)	(0.052)
Post·Treat	0.034*	-0.011**	0.001
	(0.021)	(0.005)	(0.058)
Panel B: Matched Plant Control Group			
FL·Post	0.214***	0.223***	0.171***
	(0.020)	(0.024)	(0.017)
FL·Post·Treat	-0.065*	-0.012*	-0.045
	(0.036)	(0.006)	(0.109)
Post·Treat	0.031	0.004	0.078
	(0.022)	(0.004)	(0.072)
Plant Fixed Effects	Y	Y	Y
Quarter Dummies	Y	Y	Y

Notes: This table reports D.D.D regressions of next quarter plant-level compliance on whether the plant was in Florida in the post review period and in the differing cost (Treat) group over the time period of 2007 - 2014. The specifications and samples are identical to Table III in the main text but differ by not using any additional control variables. The indicator Post takes a value of one if the time period is after the second quarter of 2012. The indicator FL takes a value of one if the plant resides in Florida. In Column (1) the indicator Treat takes a value of one if the plant's parent company has another facility outside the state of Florida. In Column (2) the indicator Treat corresponds to the number of manufacturing establishments in the county. In Column (3) the indicator Treat corresponds to the depreciated number of accumulated violations since the plant was last in compliance. Standard errors are clustered at the plant level. $***$ indicates significance at the 1% level, $**$ at the 5% level, and $*$ at the 10% level. **Sample:** Columns (1) to (3) only includes plants that were observed both prior to and after the review and indicated as "Operating" in the 2014 AFS facilities file. All three columns correspond to the subsample of 146 Florida plants (4,079 observations) with a history of noncompliance or priority violator status prior to the review. There are 568 control plants (14,122 observations) in Panel A and 294 control plants (9,114 observations) in Panel B.

TABLE AVII. Results for Compliance Mechanisms with the Matched Plant Control Group

	(1)	(2)	(3)	(4)
Panel A: Log of Annual Emissions				
FL·Post	-0.001	0.228	-0.093	-0.099
	(0.045)	(0.197)	(0.167)	(0.544)
FL·Post·Treat				0.549
				(0.571)
Post·Treat				-0.371
				(0.467)
N	6,273	612	432	1,044
Panel B: Evaluation Failure Rate				
FL·Post	0.011	0.011	0.066	0.014
	(0.007)	(0.023)	(0.048)	(0.015)
FL·Post·Treat				-0.014
				(0.037)
Post·Treat				-0.167
				(0.563)
N	8,212	491	679	1,170
Plant Fixed Effects	Y	Y	Y	Y
Time Dummies	Y	Y	Y	Y
Additional Controls	Y	Y	Y	Y

Notes: This table reports D.I.D and D.D.D regressions of two environmental outcomes (denoted by the panel title) on whether the plant was in Florida in the post review period and in the higher cost (Treat) group over the time period of 2007 - 2014 (Panel B omits the year 2012). This table is analogous to Table IV in the main text but uses the matched plant control group. The indicator Post takes a value of one if the time period is after the second quarter of 2012. The indicator FL takes a value of one if the plant resides in Florida. In Column (4) the indicator Treat takes a value of one if the plant has a parent company with at least one additional location outside of Florida. All columns use average annual plant emissions at the industry level, the nonattainment status of the county where the plant resides, and the number of manufacturing establishments in the county where the plant resides as additional controls. Standard errors are clustered at the plant level. $***$ indicates significance at the 1% level, $**$ at the 5% level, and $*$ at the 10% level.
Sample: The estimation samples only includes plants that were observed both prior to and after the review and indicated as "Operating" in the 2014 AFS facilities file. Column (4) contains all plants in Columns (2) - (3). Column (3) examines plants that were priority violators prior to the review. Column (2) examines plants that were noncompliant but not priority violators prior to the review. Column (1) examines plants that were always compliant prior to the review.

TABLE AVIII. Permutation Tests: DID Analysis with the Southern Control Group

	(1)	(2)	(3)	(4)
Panel A: Southern Plant Control Group				
FL·Post	0.002	0.002	0.084	0.065
Plant Level SE	(0.003)	(0.001)	(0.018)	(0.033)
State Level SE	(0.012)	(0.001)	(0.040)	(0.027)
Bootstrapped SE	(0.004)	(0.001)	(0.018)	(0.037)
Plant Fixed Effects	Y	Y	Y	Y
Quarter Dummies	Y	Y	Y	Y

Notes: This table reports permutation tests for the D.I.D regressions of next quarter plant-level compliance. The estimated effects are the same as those from the specifications presented in Table AII Panel A of the appendix, which is the results of Table II in the main text but without additional controls. Different standard errors are provided for reference. The indicator FL·Post takes a value of one if the time period is after the second quarter of 2012 and the plant resides in Florida. The results here indicate that the statistical significance of the primary results hold across different levels of clustered standard error. **Sample:** Columns (1) to (4) only includes plants that were in operation both prior to and after the review and were indicated as "Operating" in the 2014 AFS facilities file. Column (2) includes the subsample of 1,140 Florida plants (30,837 observations) with no history of noncompliance prior to the review, there are 2,200 control plants (57,324 observations). Column (3) corresponds to the subsample of 98 Florida plants (2,657 observations) with a history of noncompliance, but not priority violator status, prior to the review, there are 386 control plants (9,589 observations). Column (4) corresponds to the subsample of 48 Florida plants (1,422 observations) that were Priority Violators prior to the review, there are 182 control plants (4,533 observations). Column (1) includes all plants from the corresponding panel in Columns (2) to (4).

TABLE AIX. Great Recession Shock: Impact of Great Recession on Compliance

	(1)	(2)	(3)	(4)
Panel A: Southern Plant Control Group				
FL·Post	0.001	0.077***	0.079	0.160***
	(0.002)	(0.031)	(0.060)	(0.038)
FL·Post·Treat	0.001	-0.059	-0.006	-0.052
	(0.003)	(0.037)	(0.079)	(0.048)
Post·Treat	-0.001	-0.001	-0.009	0.000
	(0.002)	(0.030)	(0.040)	(0.035)
FL·Post·Treat·High				0.033
				(0.072)
FL·Post·High				-0.092
				(0.057)
Post·High				0.035
				(0.031)
Post·Treat·High				-0.001
				(0.041)
N	88,161	12,246	5,955	18,201
Plant Fixed Effects	Y	Y	Y	Y
Quarter Dummies	Y	Y	Y	Y
Additional Controls	Y	Y	Y	Y

Notes: This table reports the DDD and quadruple difference tests on the relationship between a plant residing in a high unemployment county and their compliance trend. The indicator Treat takes a value of one if the plant resides in a county that had an unemployment rate greater than 10% during the great recession. The indicator High takes a value of one if the plant has a parent company with at least one additional plant outside Florida. The results indicate that differential economic conditions during the Great Recession does not explain the improvement in Florida plant compliance after 2012. All columns use average annual plant emissions at the industry level, the nonattainment status of the county where the plant resides, and the number of manufacturing establishments in the county where the plant resides as additional controls. Standard errors are clustered at the plant level. $***$ indicates significance at the 1% level, $**$ at the 5% level, and $*$ at the 10% level.
Sample: Columns (1) to (3) correspond identically to the southern plant control group specifications in Table II of the main text. Column (4) contains the sample of all plants in Columns (2) and (3).

TABLE AX. D.I.D. and D.D.D Results: Extended ICIS Sample

	(1)	(2)	(3)	(4)
Panel A: Southern Plant Control Group				
FL·Post	0.348***	0.004*	0.093***	0.043
	(0.026)	(0.002)	(0.017)	(0.031)
FL·Post·Treat	-0.279***			
	(0.032)			
Post·Treat	0.262***			
	(0.021)			
N	22,755	106,435	15,438	7,317
Plant Fixed Effects	Y	Y	Y	Y
Quarter Dummies	Y	Y	Y	Y

Notes: This table reports the results of the southern control plant DID and DDD specifications over the time period of 2007 - 2017 using both the ICIS and AFS data. The quarterly compliance measure from the ICIS data starting in Q4 2014 is a more coarse measure of compliance than the one in the AFS data. Results from this table indicate that even with this less responsive measure plant compliance improved significantly after the 2012 Florida review. The indicator Post takes a value of one if the time period is after the first quarter of 2010. The indicator FL takes a value of one if the plant resides in Florida. In Column (1) the indicator Treat takes a value of one if the plant's parent company has at least one additional non-Florida facility. Column (1) contains the sample of plants from columns (3) and (4). Standard errors are clustered at the plant level. $***$ indicates significance at the 1% level, $**$ at the 5% level, and $*$ at the 10% level.
Sample: Column (2) corresponds to the primary estimation sample of plants with no history of noncompliance prior to the review. Column (3) corresponds to the plants with a history of noncompliance but no priority violations prior to the review. Column (4) corresponds to plants with priority violations prior to the review. All sample sizes are larger than the original due to the additional quarters from the ICIS data.