The "Gorilla in the Closet": Regulatory Enforcement Under Federalism *

Jenna Anders
Romaine A. Campbell

February 4, 2025

Abstract

How does federal regulatory capacity affect state enforcement outcomes? We provide a model in which a stronger federal regulatory agency can either strengthen or weaken states' negotiating position with their regulated entities. The optimal federal enforcement for the states is one that maximizes state-level negotiated penalties. We apply the model's insights in the context of environmental regulation to test whether the US Environmental Protection Agency (EPA) is too strong or too lenient in two environmental programs: the Clean Air Act and Superfund. First, we show that when EPA's budget was cut in 2011, state-issued penalties for Clean Air Act violations shrank by 15%. Second, we show that firms are more likely to complete remediation projects during Democratic federal administrations. Our remediation analysis identifies the mechanism: while firm cleanup behavior is affected by EPA strength, cleanups conducted by the state are not, providing evidence that the effects operate through firm-state bargaining. We conclude that almost one third of EPA's effect on environmental penalties is through its spillovers to state enforcement outcomes, and that states would benefit from a stronger EPA.

---

*Contact: jennaanders@berkeley.edu. We are grateful for guidance from Nathan Hendren, Larry Katz, and Jesse Shapiro. We also thank David Agrawal, Joe Aldy, Kirill Borusyak, Steve Cicala, John Conlon, Zoe Cullen, Liz Engle, Ed Glaeser, Claudia Goldin, Benny Goldman, Emma Harrington, Oliver Hart, William Hoyt, Sarah Jacobson, Casey McQuillan, Amanda Pallais, Dev Patel, Charlie Rafkin, Melanie Rucinski, Kim Sarnoff, Joseph Shapiro, Brit Sharoni, Hilary Sigman, Reed Walker, Anthony Yu, and participants in Harvard's public economics and environmental economics workshops. The project would not have been possible without generative input and guidance from David Hindin. We enjoyed many helpful and interesting conversations with current and former staff members at US EPA and state environmental agencies, and we are grateful they took time away from their important work to share their expertise with us. We thank MC Hanafee LaPlante and Nikolas Kirk for excellent research assistance. Finally, we gratefully acknowledge financial support from the National Science Foundation and the Chae Family Economics Research Fund. All errors are our own.

"The [Environmental Protection Agency] would ... act as a 'gorilla in the closet' for the cities and states to use to frighten polluters into submission. State regulators had long wished for a federal agency to play this role." - The Guardian: EPA's Formative Years[1]

1 Introduction

In the United States, executive regulatory agencies in the federal government (Department of Labor, Department of Transportation, etc.) often have state-level counterparts that regulate similar domains. This overlapping jurisdiction may seem inefficient: why duplicate efforts? In this paper, we present a model of overlapping jurisdiction and characterize when and how a stronger federal government can increase welfare. The key idea is that the federal government provides a threat point for states to leverage when enforcing their own regulations. We consider the context of environmental protection, a setting where many environmental statutes can be enforced by either the United States Environmental Protection Agency (US EPA, or EPA) or states' environmental agencies.

The first administrator of EPA, William Ruckelshaus, coined a term to describe the federal government's ability to affect state enforcement outcomes: "the gorilla in the closet." His idea was based on the reality that state authorities cannot unilaterally impose whatever penalties they'd like on violators of environmental statutes, but rather must negotiate penalties with firms. EPA as "gorilla" would provide the states with a federal agency to "frighten their polluters into submission" (EPA.gov, 1993).

The idea of the gorilla suggests that EPA is less concerned about, or less influenced by, the regulatory costs to local firms and the impacts on local economies than states. Consistent with this idea, we find that state penalties respond to local economic conditions but EPA penalties do not. With this evident discrepancy in regulator preferences, are current federal institutions helpful to the states, and if so, how? We develop a model of the federal agency as "gorilla" and provide an empirical test for whether EPA is too harsh given the state's objective function.

[1]"The Guardian" was an internal publication at EPA with multiple installments recording the agency's history. This installment, written by Dennis C. Williams, can be found online at https://www.epa.gov/archive/epa/aboutepa/guardian-epas-formative-years-1970-1973.html.

In our model, we place federal enforcement in the context of a bargaining game between the state and the firm over environmental enforcement outcomes, where the state can threaten to hand over the enforcement case to EPA. As in the legal reality of delegated authority in environmental programs, EPA can, at some cost, threaten to sanction the state for unsatisfactory enforcement outcomes (e.g., imposing additional reporting requirements or even taking over enforcement in a state). EPA affects state penalties in two ways. First, its penalty serves as an outside option for the states, providing them with a threat point to use in negotiations with firms. Second, its threat of sanctions on the state expands the set of firm penalty offers the state will credibly reject.

The welfare impact on states of EPA strength is non-monotonic. At low levels of EPA enforcement, higher federal penalties improve the bargaining position of the state and increase penalty offers from firms. At high levels of EPA enforcement, EPA involvement becomes unattractive to the state, and the state accepts lower firm offers rather than send the firm to EPA. This non-monotonic comparative static provides us with an empirical test with normative implications. Specifically, EPA strength is lower than the states' optimal level if and only if increases in EPA strength increase state penalties collected. By observing the effect of changes in EPA strength on state penalties, we can infer whether or not states would benefit from a stronger EPA.

We study variation in EPA strength in two environmental programs: the Clean Air Act (CAA) and Superfund. Our CAA analysis exploits federal budget cuts which reduced EPA's workforce, and our Superfund analysis exploits changes in presidential administrations.[2] The latter also provides additional evidence of the gorilla effect's mechanism: that it operates through firm-state negotiations and not state enforcement preferences.

We first estimate the effect of changes in EPA resources on state-negotiated penalties for violations under the CAA. We exploit US EPA agency budget cuts which led to a 15% reduction in EPA workforce between 2011 and 2016. Federal enforcement is largely conducted through EPA's 10 regional offices, which were differentially affected by the budget cuts, as evidenced by heterogeneous reductions in hiring of new staff (specifically,

[2]Superfund involves a natural control group that CAA lacks for analyzing effects of presidential administrations; meanwhile, only CAA's frequency of outcomes enable analysis of the dramatic but one-time budget cuts.

we use hiring of staff in legal occupations). We exploit these differences across regional EPA offices in a difference-in-differences framework, where the outcome is *state* penalty size, from EPA's database of state-reported CAA penalty data. After the budget cuts, in EPA regions which were more affected, state penalties decreased by more, even though EPA was not itself involved in these cases.[3] Our estimates suggest that the budget cut led reduced state penalty size by about 7.5%.[4] That state penalties shrink suggests that EPA strength is below the states' optimal level.

Our second context is the California Superfund program, which compels companies liable for environmental contamination to clean it up. Unlike in CAA enforcement, there is no centralized federal data repository for state cleanup programs. We extract information from hundreds of documents stored online and in physical records rooms to measure the speed of site clean-up and the estimated costs and environmental details of cleanup projects.

These data span 30 years, allowing us to exploit a different source of variation in EPA strength: political party of the US presidential administration. At the federal level, EPA collects higher penalties during Democratic administrations than during Republican administrations, indicating EPA involvement is more costly for firms during Democratic administrations. The Superfund context also includes a control group: a subset of sites, called "orphan sites," do not have viable responsible parties to conduct cleanups, so the state cleans up these sites itself instead of negotiating with firms. We combine these in a difference-in-differences analysis, using cleanup speed, a measure of firm cooperation in this context, as our outcome.

We show that firm-led cleanups under state oversight are significantly less likely to begin major cleanup projects during Republicans presidencies, when EPA is a more lenient enforcer; meanwhile there is no significant effect among orphan sites. With additional data extracted from cleanup project documentation, we provide suggestive evidence that

[3]We provide evidence that the parallel trends assumption holds: before the budget cuts, (eventual) declines in regional office activity do not predict state outcomes.

[4]The results in the difference-in-differences framework are not statistically significant, likely due to high variance in the outcome and low variance in the treatment variable. We also present a repeated cross-section of state penalties before and after the budget cut which shows a significant decline in state penalty size but, we argue, likely overstates the effect of the budget cut.

firms also choose less expensive cleanup projects under Republican presidencies. Because we see firm cooperation increase and not decrease when EPA becomes harsher, we conclude that EPA is not too harsh for the states' liking. Moreover, as orphan sites' cleanup speed is an outcome of state behavior and not state-firm bargaining, the null result for orphan sites serves as validation of our modeling assumption that changes in state enforcement outcomes are not driven by states' fear of being sanctioned by EPA (which, if operative, would affect outcomes even in the absence of firm-state bargaining).

Together, our results show that federal agency strength matters for state enforcement outcomes. We can use our estimates from the Clean Air Act to ask how much of EPA budget's effect on penalties collected comes from EPA's own enforcement outcomes versus its spillover effects on the states. We find that almost one third of EPA budget's total effect on penalties comes from the "gorilla" effect.

Related literature. We provide insights into the consequences of a federalist government by incorporating the importance of firm bargaining power in regulatory outcomes. Prior literature cast the social benefit of the federal government as either to internalize spillovers across jurisdictions, to prevent a regulatory race-to-the-bottom, or to provide services more efficiently (Buchanan, 1950; Tiebout, 1961; Oates, 1972; Besley and Coate, 2003; Harstad, 2007; Dijkstra and Fredrikssonn, 2010; Calabrese et al., 2012; Chang et al., 2014; Callander and Harstad, 2015; Agrawal et al., 2022; Slatterey, 2022; Tang, 2022; Agrawal et al., 2023).[5] We show that the structure of federal and state policy can affect states' interactions with firms. We also propose a sufficient statistics test for whether the states would benefit from a stronger federal government.

Our paper further relates to work estimating regulator preferences or providing evidence of different preferences across regulatory bodies (Agarwal et al., 2014; Jung and Makowsky, 2014; Lim and Yurukoglu, 2018; Earnhart and Frieson, 2021; Kang and Silveira, 2021; Tang, 2022). Our model adds to a smaller literature studying whether differences in regulator preferences can be beneficial to the regulators (Barro and Gordon, 1983;

[5]Several papers in the environmental federalism literature explicitly consider the role of decentralization in overall pollution levels (Sigman, 2002, 2005; Lipscomb and Mobarak, 2017). The broader federalism literature mostly focuses on the design of policy and not its implementation; we share a focus on regulatory *enforcement* with Woods (2006) and Konisky and Woods (2010).

Rogoff, 1985; Gutiérrez and Philippon, 2019). Our sufficient statistics test brings methods from public economics, which are usually applied to questions of optimal taxation, to this literature and allows us to map reduced form empirical results to welfare implications; we apply these methods to studying interactions between government agencies (Baily, 1978; Chetty, 2006, 2009; Piketty et al., 2014; Allcott et al., 2014; Hendren, 2016, 2021; Farhi and Gabaix, 2020; Kleven, 2021).

Finally, we bring new data to and ask new questions in the environmental enforcement literature. Our dataset of California's Superfund program allows us to identify a novel factor in the efficacy of a large state-run program.[6] The Clean Air Act is a more established setting for studies of enforcement (e.g. Evans, 2016; Evans et al., 2018; Blundell et al., 2020), but there is little empirical analysis on EPA as a "gorilla." We believe the closest papers to ours are Evans and Stafford (2019) and Blundell (2020), which both study formal and informal sanction threats EPA has used to affect state enforcement behavior.[7] Relative to these papers, we propose and identify the role of state bargaining with firms as a mechanism for the effect of EPA oversight, and we propose welfare implications.

The paper proceeds as follows. Section 2 describes the regulatory context. Section 3 provides the model. Section 4 describes the data we use. Sections 5 and 6 present our empirical analyses of state Clean Air Act penalties and California Superfund cleanups, respectively. Section 7 concludes.

2 Qualitative evidence and regulatory context

Do the states view the US EPA as a "gorilla?" There is ample qualitative evidence that they do. We begin by presenting such evidence from federal and state Superfund programs, which oversee cleanup of environmental contamination. In these programs, reg-

[6]Prior Superfund work documents substantial health effects of Superfund cleanups (Currie et al., 2011; Persico et al., 2022) and has mixed evidence on capitalization of benefits (Greenstone and Gallagher, 2008; Gamper-Rabindran and Timmins, 2011; Gamper-Rabindran et al., 2011). Given the program's unusually punitive enforcement, there is also an extensive legal literature on Superfund. Mintz (1988) specifically discusses the role of political leadership at EPA in Superfund implementation in the 1980s.

[7]Evans and Stafford (2019) show that when EPA published a "Watch List" which identified high-priority violators, state CAA enforcement activity increased for facilities at risk of Watch List listing. Blundell (2020) shows that when EPA determined Florida's CAA enforcement on certain facilities to be lacking, compliance improved among these facilities.

ulators and firms (in this case referred to as responsible parties) negotiate agreements for the work that the responsible parties will do to remediate contaminated sites.

Our first qualitative evidence, presented in Figure 1, comes from documentation found in two state environmental agencies' records rooms. In Panel A, an attorney from the Maine Department of Environmental Protection writes to a mining company to alert them that the US EPA is assessing their site as a potential federal Superfund site, and asks whether this information changes the company's position in negotiations with the state. In Panel B, after sending multiple letters with no response, the California EPA's Department of Toxic Substances Control warns a towing company that they will submit the site for screening from EPA if the company continues to be unresponsive.

Additional evidence beyond the Superfund program is found in documentation from the federal government and a survey of state officials. In 1985 and 1990, a researcher surveyed state environmental officials and asked them directly whether they agreed or disagreed that the "threat of EPA intervention strengthens [the state's] position with polluters" (Tobin, 1992). In 1990, 90% of the state officials surveyed felt that their state's enforcement was strengthened by the threat of EPA intervention. A 2003 report by the Government Accountability Office (GAO) on the fiscal state of the Superfund program mentions that, although states have their own cleanup programs, "a few states pointed to the value of maintaining strong EPA enforcement powers because they encourage responsible parties to cooperate with the state" (US General Accounting Office, 2003). Later, in 2019, when EPA undertook efforts to have this role formally recognized (conceivably in response to the Trump administration's efforts to reduce federal enforcement), an EPA administrator testified to Congress about EPA's efforts "working with states on problem facilities, encouraging them to comply with the state or face potential EPA action."

2.1 Regulatory context: Clean Air Act

The Clean Air Act (CAA), passed in 1970 and amended in 1990, is a multifaceted federal statute. We focus on the enforcement of CAA's regulation of stationary sources (also known as point sources) from 2002 to 2020. Under CAA, stationary sources are subject to emissions limits and requirements on pollution control equipment and operating permits.

Both excess emissions and procedural noncompliance constitute violations which can be penalized.

Our analysis uses the size of CAA penalties issued by the states as our main outcome. Penalties are largely determined by the economic benefit of noncompliance (to achieve a deterrence effect) and the gravity of the violation (EPA, 1991). The gravity of the violation is not only limited to the extent of possible emissions exceedance, but can also include other considerations, such as a company's net worth. EPA also explicitly allows for adjustments for "public interest" (to avoid plant closings and bankruptcies) and "litigation risk" (allowing lower penalties when the court case is weaker). In many cases, each individual day of noncompliance constitutes a separate offense, so that the final penalty can depend on the duration of noncompliance.

Like many federal statutes created in the 1970s, the CAA allows for enforcement authority to be delegated the states. States authorized with "primacy" in enforcement are the primary entity responsible for enforcement of the federal law (Norwood, 2015). All fifty states currently have CAA primacy for Title V sources. States conduct over 95% of inspections and issue over 80% of formal enforcement actions for violations.[8]

States can, and do, ask EPA for support on cases "when the weight of the EPA is needed" (Earnhart and Frieson, 2021). This can mean collaborating on an enforcement action, or alternatively asking EPA to handle a case.[9]

CAA enforcement actions brought by the US EPA are largely brought through the 10 regional offices. Regional offices are relatively independent: they have different organizational structures, different priorities, and different enforcement cultures (Fiorino, 1995; Engelberg et al., 2011).

[8] Authors' calculations using data from EPA Enforcement and Compliance History Online (ECHO).

[9] EPA also retains the right to enforce independently. For example, EPA "does not delegate ... the authority to make decisions that are likely to be nationally significant." (https://www.epa.gov/caa-permitting/delegation-clean-air-act-authority) US EPA also has direct jurisdiction over some facilities (e.g., federally-owned facilities).

2.2 Regulatory context: Superfund

The federal Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA) was passed in December of 1980. In this project, features of the federal program are relevant because it provides a threat California enforcers can use in their dealings with firms; our outcomes come from the California counterpart to the federal Superfund program.

The federal Superfund program. The CERCLA Congressional bill created a trust fund (the "Superfund"), funded through appropriations and earmarked corporate taxes, for the US EPA to use for site cleanup and enforcement actions against parties responsible for contamination. Initially, the idea of the Superfund was to allow EPA to move freely with expensive remediation projects, funded through the trust fund, before recovering their costs from liable parties (EPA.gov, 2005b). In 1986, however, the program shifted to an "enforcement first" approach (EPA.gov, 2005a), in which EPA uses its enforcement power to compel companies to conduct the cleanups themselves. Regardless, EPA cannot use the Superfund money for (non-emergency) cleanup actions unless a site is listed on the National Priorities List (NPL). Once a site is on the NPL, CERCLA grants the federal government extensive power.[10]

State Superfund programs. Many states run their own Superfund programs modeled after the federal program, complete with their own Superfund trust funds allocated by the state legislature.[11] Upon discovery of a contaminated site, these states will often attempt to address the problem in a similar fashion to the approach in the federal Superfund program. These programs are not under federal oversight but address environmental issues that the federal Superfund program could also address.

As shown earlier, states exploit this overlapping jurisdiction in their dealings with responsible parties, explicitly threaten to refer sites to the US EPA (Figure 1). In our anal-

[10]EPA generally tries to negotiate agreements with companies wherein the company agrees to remediate the site. However, if the company is unwilling to negotiate, EPA can issue unilateral orders, and if the company does not abide by these orders, EPA can sue them in court for damages and penalties. Liability under CERCLA has unusually broad scope, and courts can order triple damages (fines up to three times the costs incurred by EPA).

[11]In 2001, all states had state laws enabling cleanup enforcement, but only 24 were funded by legislative appropriations (Environmental Law Institute, 2002).

yses, we consider how the strength of this threat affects the speed of cleanups in the California state Superfund program.[12]

Federal non-emergency Superfund-funded response actions must be approved by state governors.[13] Thus, the threat of federal involvement in a contaminated site is moderated to a substantial degree by the state.

2.3 Sanctions

In Section 3, we introduce a model of the gorilla whereby the EPA can, at a cost, threaten to sanction states for unsatisfactory enforcement of environmental statutes. The sanctions available to EPA differ by program.

Clean Air Act sanctions: The Clean Air Act is a delegated environmental program. EPA's methods of imposing costs on states with unsatisfactory implementation of delegated programs include withholding grant funds, increasing oversight of state processes and decisions, and revoking state primacy (U.S. EPA, 1984; Engelberg et al., 2011). EPA can also issue its own enforcement actions against specific facilities (called "overfiling") if it doesn't consider the state to be taking timely action against noncompliance. In each of these sanctions options, policy guidance is to take a "constructive approach" and, before taking action, "give the state a chance to explain and/or correct" problems that might otherwise result in sanctions.[14]

Superfund sanctions: As the state Superfund programs are independent programs and not delegated federal programs, EPA's recourse against lenient states is more limited in this context. The most realistic option for "sanctions" broadly understood is through

[12] As put in a testimony made in a 2002 Congressional hearing on behalf of the U.S. Public Interest Research Group: "The success of... state programs heavily depends on the Federal Superfund program providing a credible deterrent against polluters who refuse to clean up sites under state programs" (US Senate, 2002). Among other reasons, firms may prefer state enforcement to federal because they believe it will involve lower transaction costs and a smaller decrease in property values, or because they believe they will have more influence over the state process (Environmental Law Institute, 1990, p. 59-60).

[13] CERCLA §9611(h)(i). EPA's Superfund Emergency Response and Removal Program, which orchestrates a large share of all EPA cleanup actions (Jenkins et al., 2012), does allow EPA to intervene without governor approval.

[14] McCarthy (1997) collected data on EPA sanctions for inadequate state plans to attain ambient air quality standards and found that between 1990 and 1997, EPA imposed additional offset requirements 14 times and withheld a portion of state highway funds twice. Overfiling is also understood to be rare in practice (US Senate, 1997).

reduced funding of grants supporting state remediation (cleanup) efforts. EPA can also initiate emergency actions, including restricting land use and suing firms, without explicit state consent.

3 Model

The model casts EPA involvement in a case as the states' outside option when negotiating penalties with firms.

The model gives the result that the states can benefit from an EPA which has different preferences from its own, and specifically, an EPA which issues larger penalties than the states do. Even so, it is possible for EPA to be *too* harsh, as states will shy away from involving it in enforcement if it carries too big a stick. The model provides a sufficient statistics test for whether EPA is too harsh or too lenient: we show that if (and only if) a stronger EPA raises state penalties, then EPA is weaker than the states would like.

In reality, EPA has another tool for affecting state penalties: it can threaten to sanction states for unsatisfactory enforcement outcomes. In this section, after setting up the model, we first derive results when EPA's cost of issuing sanctions approaches infinity, essentially depriving it of this tool, to make clear the key intuition in the model. We then reduce EPA's cost of sanctions, show that this introduces a threat to our normative conclusions, and show how our empirical analysis addresses this threat.

3.1 Set-up

There are three actors: EPA, a representative state, and a representative firm. There is a set of environmental violations by firms determined in the base period. The actors interact to decide a penalty size for the enforcement case of a given violation v.[15]

The firm commits a violation if its benefit from violating exceeds a randomly drawn cost. Once the violation is established, the firm offers a penalty to the state, and the state either accepts it or sends the case to the EPA. Let $I_S = 1$ if the state issues the penalty

[15] In the model, we use penalty size as the enforcement outcome. However, this could be thought of as any enforcement outcome that is costly to the firm and beneficial for the state. In our Superfund context, we use cleanup as the enforcement outcome.

(i.e., accepts the firm's offer), and 0 if EPA issues the penalty. Let p denote penalty size; $p = p_{s_v}$ when the state issues the penalty, and $p = p_e$ when EPA issues the penalty. For each case, the state has a uniformly drawn comparative advantage in enforcement $\zeta_v \sim U[\underline{\zeta}_v, \overline{\zeta}_v]$, with $\overline{\zeta}_v > 0$. This comparative advantage could reflect, for example, states' specialization in the environmental issues most relevant to their geographic and demographic characteristics. ζ_v could also represent the state's political cost of involving EPA.

Finally, EPA has the ability to threaten to sanction the state. In reality, EPA has several sanctioning tools at its disposal which vary in the costs the sanctions impose on the state. At the extreme, it can revoke the state's enforcement primacy. Examples of lesser sanctions include imposing additional reporting requirements and requiring federal review of state enforcement actions.[16] Regardless of the sanction, EPA policy is to give the state an opportunity to correct their behavior before EPA imposes sanctions on the state. In our model, EPA must pay a cost to threaten sanctions, even before implementing them. If EPA chooses to threaten the state with sanctions, it chooses a level k of sanctions to threaten, which costs EPA $c(k)$. We assume costs of sanction threats are linear: $c(k) = ck$.

3.1.1 Agents' preferences and technology

The firm is a cost minimizer. It draws a compliance cost $\eta_v \sim F$ and violates if its compliance cost exceeds its expected penalty, which is $I_S p_{s_v} + (1 - I_S) p_e$.

The state has strictly concave utility over penalty size.[17] For generic penalty size p, it trades off the environmental benefits of a penalty $b(p)$ (deterrence) with the economic harm that issuing the penalty will incur, $\tau(p)$.[18] We assume $b'(p) > 0$, $b''(p) < 0$ and $\tau'(p) > 0, \tau''(p) > 0$.

Specifically, the utility of the state from penalty size p and sanctions $k(p_{s_v})$ is

$$U_S(p, k) = b(p) - \tau(p) + I_S(\zeta_v - k).$$

[16] https://www.epa.gov/sites/default/files/2014-06/documents/state-oversight-strategy.pdf

[17] We model the regulators as having concave utility over penalty size for ease of exposition. Appendix E.1 shows how to recast the model as regulators seeking to minimize deadweight loss.

[18] For a discussion of deterrence effects, see Section E.2.

Denote the state's preferred penalty

$$p_s^* = \arg\max b(p) - \tau(p).$$

EPA's utility from penalty size a is similar, although it pays a cost of imposing sanctions and it weighs economic harms of enforcement differently than the states so. It also receives ζ_v if the state issues the penalty, and not otherwise.

$$U_E(p,k) = b(p) - \beta\tau(p) - ck + I_S\zeta_v.$$

Denote EPA's preferred penalty

$$p_e^* = \arg\max b(p) - \beta\tau(p).$$

If $\beta < 1$, then EPA will prefer a penalty larger than the states do.

Technology. EPA has a workforce N which allows it to get some share $\sigma(N)$ of its preferred penalty, with $\sigma(N) \in [0,1]$, $\sigma'(N) > 0$, for cases it handles.

EPA strength. When we refer to EPA strength, we are referring to the determinants (N, β) of the penalty EPA issues for cases it handles. A "stronger" EPA is one that would issue a higher penalty, either because of a larger workforce N or because of a lower weight β on economic harm.

3.1.2 Timing

Since states have enforcement primacy in our context, our model allows the state to choose whether it or EPA issues the penalty. If the state handles the case, the state must accept whatever penalty the firm offers, and whatever sanctions EPA threatened. If EPA handles the case, EPA unilaterally issues its own penalty.

The sequence of actions taken about a given violation v is:

0. The firm draws η_v and decides whether to commit a violation. If it does, then...

1. The firm makes a penalty offer p_{s_v} to the state.

2. EPA pays $c \cdot k$ to issue sanction threat k.

3. The state either accepts the firm's offer p_{s_v}, or it rejects the offer and sends the case to EPA.

4. If the state sends the case to EPA, EPA chooses p_e^* and issues penalty $p_e = \sigma(N)p_e^*$.

There are a few features the timing worth elaborating on. First, that the firm moves first in its negotiations with the state represents its bargaining power in this context. We give the firm the bargaining power because the penalty is a cost to the firm and a benefit to the state, so the firm would be more patient than the state in a game with alternating offers. Second, our assumption that EPA chooses its sanctions after the firm's offer reflects the fact that EPA can adjust its sanctions depending on the firm's offer.[19] That EPA threatens sanctions before states accept or reject an offer captures EPA's institutional policy to allow states an opportunity to correct their behavior if EPA believes enforcement is inadequate: official EPA policy dictates that EPA give states every opportunity to improve their programs after warnings like sanction threats. Finally, our assumption that EPA chooses its preferred penalty p_e^* last gives us that EPA does not internalize the effects of its preferred penalties on the state's decision.

3.2 Equilibrium

All proofs are in Appendix E.

Let $u_S(p) = b(p) - \tau(p)$ and $u_E(p) = b(p) - \beta\tau(p)$. That is, u represents only the portion of each agent's utility derived from environmental benefits and economic harm.

Since the state has enforcement primacy (i.e., it decides who handles the case), it will only allow EPA to take the case if the utility it would receive from EPA involvement is higher than the utility it would receive from accepting the firm's offer. The relative utility the state would get from rejecting vs. accepting the firm's offer depends on (1) its utility from EPA's penalty and (2) the threat of sanctions it faces.

Sanction threats increase state penalties by making EPA involvement a more credible threat to the firm—that is, by expanding the set of firm offers the state will credibly reject.

[19]It also gives a more realistic equilibrium outcome than a model where EPA chooses sanctions before firm offers: with this alternative timing, EPA always threatens sanctions, while in reality, sanctions are rarely threatened.

If EPA threatens sanctions, the state will have a good reason to reject penalty offers (since rejecting the offer will allow them to avoid sanctions). Firms know this, and raise penalty offers.

In equilibrium, firm offers ensure both (1) the state has at least as high utility from accepting the offer as rejecting it and (2) EPA does not have an incentive to threaten the state with sanctions. This is true for firm offers $p_{\tilde{s}_v}$ such that

$$\underbrace{u_S(p_{\tilde{s}_v})}_{\text{State utility from firm offer}} \geq \underbrace{u_S(p_e) - \zeta_v}_{\text{Pure outside option value}} + \underbrace{\frac{u_E(p_e) - u_E(p_{\tilde{s}_v}) - \zeta_v}{c}}_{\text{Sanctions threat}} \tag{1}$$

As long as there exists at least one $p_{\tilde{s}_v}$ which satisfies Expression (1), the firm offers $p_{\bar{s}_v} = \min\{p_{\tilde{s}_v}\}$, and the state accepts the firm's offer. For $\zeta_v < 0$ (when EPA has a comparative advantage in enforcement on a given case), it is possible that there does not exist a $\tilde{s}_v$ which satisfies Expression (1). In this situation, regardless of the firm's offer, the state will reject the firm's offer and send the case to EPA (given EPA's sanction threats).

In reality, very few cases get sent by the state to EPA. For this reason, going forward, we suppose $\overline{\zeta}_v > \underline{\zeta}_v > 0$.

In equilibrium, $k = 0$, and $I_S = 1$. The mass of violations are those satisfying $\eta_v > p_{\bar{s}_v}$.

3.2.1 Without Sanctions

To illuminate the intuition behind our proposed test, we start by depriving EPA of sanction power. Now, in Equation 1, firm offers are the minimum $p_{\tilde{s}_v}$ such that $u_S(p_{\tilde{s}_v}) = u_S(p_e) - \zeta_v$.

Our first result is that the state penalties will never exceed the state's preferred penalty p_{s^*}, regardless of EPA's penalty p_e.

Proposition 1. *When EPA cannot sanction the state, equilibrium state-issued penalties are lower than the state's preferred penalty ($p_{\bar{s}_v} < p_s^*$) and EPA's penalty ($p_{\bar{s}_v} \leq p_e$).*

The logic behind the first inequality is illustrated by Figure 2. Recall that the firm offers the lowest penalty possible that gives the state equal utility to EPA involvement. Because

p_s^* maximizes the state's utility, any penalty offer higher than p_s^* has a corresponding penalty offer *lower* than p_s^* which gives the state equal utility. The firm will always choose this lower penalty offer.

That the state can never attain p_s^* comes from our assumption that $\zeta_v > 0$. Even if EPA's penalty is the state's preferred penalty ($p_e = p_s^*$), the firm can take advantage of the fact that the state prefers to handle the case itself, and can offer the state something lower. If $\zeta_v \leq 0$, the first inequality would be weak ($p_{\bar{s}_v} \leq p_s^*$).

The second inequality follows trivially from firm optimization: the firm gets no benefit from offering a penalty higher than p_e^*.

Our second result is that when increasing EPA strength increases state penalties, it is also increasing state welfare, and vice versa.

Proposition 2. *When EPA cannot sanction the state,* $\frac{dU_S}{dp_e} > 0$ *if and only if* $\frac{dp_{\bar{s}}}{d[\sigma(N)p_e^*(\beta)]} > 0$.

Notice in Figure 2 that there exists an EPA penalty such that strengthening EPA (i.e., increasing p_e further) strictly decreases penalty offers from firms, which in turn strictly decreases state welfare. With Proposition 1, we can conclude that an increase in EPA strength improves state welfare if and only if it increases state penalties. Note that the *magnitude* of $\frac{dU_S}{dp_e}$ is not reflected in $\frac{dp_s}{d[\sigma(N)p_e^*(\beta)]}$, only the sign. In fact, the magnitude of $\frac{dU_S}{dp_e}$ is inversely related to the magnitude of $\frac{dp_s}{d[\sigma(N)p_e^*(\beta)]}$: a large change in observed penalty size reflects a smaller change in utility than does a small change in penalty size.

3.2.2 With Sanctions ($c < \infty$)

Statutorily, EPA has an additional tool for affecting state penalties: it has oversight authority of state enforcement, and it can sanction states for unsatisfactory enforcement outcomes. Sanctions are rarely observed in reality. However, even when EPA's sanctions are off the equilibrium path, they can still affect firm penalty offers.[20]

In this section, we allow the reality of EPA sanction power.[21] EPA's sanction power

[20] As noted in Tobin (1992), "Faced with possible sanctions... state agencies can assert [to firms] that they have no choice but to enforce the mandates that the federal government has imposed on them."

[21] We aim to capture the spirit of EPA oversight described in the conclusion of a 1984 policy memo on oversight in delegated programs: *"[This oversight policy] will demonstrate our desire to work with and assist states in a positive manner... while at the same time retain-*

makes an additional assumption necessary to maintain the conclusions of Propositions 1 and 2. We show in the next subsection that it also rationalizes a harsher EPA.

When EPA has sanction power, firm offers are weakly higher than in the case without sanction offers. For low enough sanctioning costs, this can drive firm offers above p_s^*: the state requires a penalty larger than its preferred penalty because accepting only its preferred penalty would induce EPA to impose sanctions. Specifically,

Proposition 3. *For a given ζ_v, state penalties exceed p_s^* if and only if $c < \frac{u_E(p_e)-\zeta_v-u_E(p_s^*)}{u_S(p_s^*)+\zeta_v-u_S(p_e)}$.*

To have $p_{s_v} < p_s^*$ for all β, N, and ζ_v, it is sufficient, but not necessary, for $c > \frac{1-\underline{\zeta}_v}{1+\underline{\zeta}_v}$.

When state penalties can exceed p_s^*, we can no longer conclude that higher state penalties improve state welfare. Proposition 3 clarifies the costs c that are needed such that state penalties will not exceed p_s^*, ensuring that our normative conclusions (Proposition 2) are valid.

Proposition 3 motivates our Superfund analysis. In "orphan sites," states can choose their own enforcement outcomes (cleanup pace), instead of being beholden to negotiations with firms. They will choose their preferred outcome p_s^* unless EPA sanctions are binding. If EPA costs of threatening sanctions are sufficiently low, the state will have to choose $p_{s_v} > p_s^*$ and moreover, will change its choice of p_{s_v} when EPA strength changes. If EPA's sanction costs are sufficiently high, the state will choose p_s^* regardless of EPA's strength. We will show that enforcement outcomes do not change with EPA strength when the state has complete control over enforcement outcomes and does not need to negotiate with firms; thus we can infer that EPA sanctions are not binding.

Our final proposition is that when EPA has sanction power, the state's utility is maximized when EPA and state preferences diverge.

Proposition 4. *For $0 < c < \infty$ and $\forall\ N$, $\arg\max_\beta U_S < 1$.*

ing our commitment to maintain high national environmental standards through appropriate sanctions and independent action, as necessary." Found at https://www. epa.govsites/default/files/2019-12/documents/epa_policy_on_oversight_of_delegated_environmental_ programs_1984.pdf. Our model emphasizes the role of sanctions and not independent action. Independent action is understood to be unusual; moreover, it could be recast as sanctions to the state if states face a utility cost of losing control of an enforcement case.

The intuition: the state's comparative advantage weakens its bargaining position, as the firm can exploit the fact that the state prefers to handle cases itself rather than send it to EPA. EPA willingness to sanction compensates for this, strengthening state bargaining power. While an EPA which underweights economic harm relative to the state (β<1) provides a worse outside option penalty for the state relative to an EPA which shares the state's preferences, it also more readily threatens sanctions; for intermediate values of β, this expands the set of firm offers the state will credibly reject.

3.3 Discussion

3.3.1 EPA vs. states' utility functions

Does EPA's objective function differ from the states'? Appendix Figure A.1 provides descriptive evidence that state penalties are lower when the local unemployment rate is higher, while EPA penalties are not. Under the assumption that a penalty causes more economic harm during bad economic times than good, this suggests that EPA cares less about the economic harms of penalties than the state. In Appendix C, we provide additional empirical evidence that this modeling assumption is realistic.

3.3.2 Implications for optimal policy

Finally, we discuss optimal policy in the context of our model. Let the social planner's objective function be

$$U_{SP} = b(p) - \beta^{SP}\tau(p) + \zeta_v I_S - k - \kappa(N).$$

The social planner has her own weight on the economic harm of enforcement which can differ from EPA's weight. Like the state and EPA, the social planner also receives ζ_v if the state issues the penalty. She must also pay EPA's cost of threatening sanctions k (although $k = 0$ in equilibrium), and she pays a cost to fund EPA's workforce, $\kappa(N)$. We allow her to also change EPA's preferences, which in this section we denote β^{EPA}, at no cost.

Generalizing Proposition 2. Proposition 2 allows us to draw normative conclusions from our empirical effects, but its statement is limited to state welfare. If the state's objective function differs from the social welfare function, what can we say about social welfare? As long as the social planner's optimal penalty is higher than the state's preferred penalty ($\arg\max_a U_{SP} > \arg\max_a U_S$), we can generalize Proposition 2: U_{SP} is increasing in EPA strength if and only if state penalties are increasing in EPA strength.

The assumption $\arg\max_a U_{SP} > \arg\max_a U_S$ may be reasonable if, for example, states' concern about economic harm of enforcement is partially about trans-state movement of industry. A counterexample would be a model where investing in EPA resources is particularly expensive (large $\kappa'(N)$), so that obtaining the state penalties comes at too high of a cost.

Optimal β. Proposition 4 can be extended to the social planner. To the extent that changing β is costless for the social planner, a social planner with $\beta^{SP} \leq 1$ prefers an EPA with $\beta^{EPA} < \beta^{SP}$.

3.3.3 Model extensions

Multiple states with heterogeneous preferences. The effect of changing EPA strength depends on the weighted average of states' preferences. (Only) If some states prefer a stronger EPA and others prefer a weaker EPA, then it is not necessarily true that an increase in average penalty size reflects an increase in average state utility. This is proven and discussed in Appendix E.4.

Firm mobility across state lines. Firm mobility can invalidate the sufficient statistics test, but only if firms are particularly mobile. A state that dislikes a stronger EPA in its own state might still like a stronger EPA in neighboring states, if this prompts firms to move into its own state. However, it would need to be that enough firms move to offset the change in utility from penalties. As noted earlier, smaller changes in state penalties reflect larger changes in state utility, so in cases where state utility is more responsive to EPA strength, firms movement is *less* responsive, making firm mobility less likely to offset the change in utility from penalties on the existing firms.

4 Data

4.1 Clean Air Act

State penalty data. The data for our Clean Air Act analyses come from an EPA database called ICIS-AIR, available from EPA.gov, which includes enforcement and facility data for stationary sources of emissions. We use data from 2001-2019. The formal enforcement action data include penalty size, settlement date, and enforcement agency (state, federal, and local); the data on facilities include facility location, industry, and current operating status. EPA categorizes stationary sources by their emissions potential and only requires that states submit data on formal enforcement actions for major and synthetic minor sources; thus, we limit our data to this subsample.[22] The definition of a "formal enforcement action" (FEA) is somewhat at the states' discretion. For consistency across states, we consider an FEA one that has a non-zero penalty associated with it.[23]

We adjust penalty amounts to 2010 dollars. The raw penalty data include very large outliers. For example, while the 95th percentile penalty issued by EPA in our sample period is $133,526 (in 2010 dollars), the largest penalty in the data is $26 million. EPA often uses median penalty values when presenting summary statistics for this reason. We instead top-code state (EPA) penalties at the 90th percentile of the state (EPA) penalty distribution, and then log-transform the penalties. We show that the results are similar but less (more) precise when values are top-coded at the 95th (85th) percentile.

Violations. Our focus in this section is on average penalty size for state formal enforcement actions. We do not have systematic data on violations for the penalties we observe. We requested such data from several states, but only received these data from one (the Florida Department of Environmental Protection).

[22]Sources are categorized by the quantity of regulated pollutants they emit or have the potential to emit. This sample accounts for about 23% of currently operating facilities registered with EPA. However, in additional data obtained from the Florida Department of Environmental Protection, we see that these facilities account for 58% of violations that resulted in formal enforcement actions (in Florida) between 2000 and 2022.

[23]There also appears to be some misreporting on the zero-penalty margin. For example, one state staff member told us that they define a FEA as an enforcement action that has a penalty associated with it, and yet, over 10% of this state's FEAs appear to have no penalty in US EPA's data.

In the Florida data, the plurality of violations (35%) that resulted in formal enforcement actions were discovered by direct inspections; an additional 20% were discovered by file review. The median violation was resolved within six months of being discovered, although the longest 3% of violations took over 2 years to resolve. We are able to categorize roughly three-quarters of violations into "procedural" and non-procedural violations (our own distinction), and find that 58% are procedural: i.e., they relate to incomplete permitting, late testing, etc., and not to excess emissions.

Emissions. We include facility-reported data on total air emissions from EPA's Toxic Release Inventory (TRI). We use these data as a proxy control for facility size. TRI reporting is only required for a subset of facilities, based on number of employees, chemicals emitted, and emissions quantity. We observe roughly half of penalized facilities in the TRI.[24] Appendix Figure A.2 shows that TRI emissions are strongly correlated with penalty size.

4.2 Hazardous substances (Superfund)

Our data on environmental remediation (cleanup) projects come from the California Department of Toxic Substances Control (DTSC), which is a department within the California Environmental Protection Agency (CalEPA).[25] We rely on the database the DTSC uses to track their cleanup projects internally, "EnviroStor." Project managers overseeing cleanups are required to update EnviroStor with cleanup progress, entering the dates of specific milestones and uploading certain documents that provide additional details. Thus, for each site, EnviroStor includes a history of relevant activities (site assessments, cleanup decisions, results from post-cleanup monitoring, etc.), as well as limited site characteristics (location, acreage, funding source). Our main outcome uses the dates of "remedial actions," which are large cleanup projects meant to either contain or remediate the contaminaton.

Many environmental remediation efforts began in the 1980s, or even earlier. As one

[24]We bin emissions quantity in our regressions, so that facilities without data can still be included in the analyses.

[25]The CalEPA was created in 1991 (DTSC, 2023); before this, the cleanups were handled by the toxic substances control division of the California Department of Health Services.

might expect, some early remediation projects have less extensive coverage in the online database EnviroStor. However, when DTSC project managers update the database with new activities, they are instructed to retroactively log dates and documents of any missing prior major activities. To explore the possibility of sample selection in the early period of our sample, we visited four DTSC records rooms across California (the two Los Angeles offices, the Berkeley office, and the Sacramento office) to view paper records from early sites.[26] We found little evidence of major activities in the paper documents that were not logged online, suggesting limited sample selection.

Sample Restriction. In our main analysis, we limit the sample to sites under DTSC jurisdiction. We further restrict the sample to sites that are over 3 acres. Using acreage as a proxy for site complexity and threat to human health, we argue that the threat of EPA involvement is much less credible for small sites. Indeed, the probability of becoming a Superfund site is three times as large for sites over 3 acres versus sites under (Appendix Figure A.3).[27]

Outcome: Remedial actions dates. Our main outcome uses the date that a remedial action was completed on sites under state oversight. In the model, the probability of a cleanup activity in a given year corresponds to penalty size p: timely remediation is costly for the firm but has environmental benefits.

Remedial action completion dates (i.e., the date the state approved a completed remedial action) are a logged activity in EnviroStor. The dates that remedial action began are not. These dates are, however, usually available within uploaded forms. For every remedial action we observe in the database, we search documents uploaded to EnviroStor for remedial action start dates. In supplemental results, we replace remedial action completion dates with the start dates listed in these forms where available. Because the remedial action start dates are not always obvious even within the uploaded forms, we present these results as a robustness check instead of as the main results.[28]

Remedial alternatives and estimated costs. Before a remediation project begins under DTSC oversight, DTSC requires that the responsible firm(s) propose and assess multiple

[26]We're deeply indebted to the DTSC records staff and project managers for their help with this effort.
[27]We also show in appendix figures robustness to different size cutoffs.
[28]We detail our procedure for identifying start dates in Appendix D.

remedial alternatives, or options for cleanup, that vary in how extensive and how expensive they are. For example, alternatives for remediation of contaminated soil might include no action (required as an alternative for all projects, at zero cost); monitoring of soil and groundwater for a certain length of time; and excavation and disposal of contaminated soil. The firm is required to estimate and report the projected costs of each remedial alternative. Benefits are also considered, but are generally only discussed qualitatively. The firm and DTSC then agree on a single remedial alternative to pursue.

The remedial alternatives are included in report PDFs uploaded on EnviroStor. For all sites over three acres with the relevant documentation (103 sites in total), we access these reports and log the remedial alternatives, their costs, and the chosen alternative. In supplemental analyses, we use the cost of the chosen remedial alternative.

For additional details on our data for the Superfund and the CAA analysis, please see Appendix D.

4.3 Other data

For economic conditions, we use data on state and county unemployment rates from the Bureau of Labor Statistics Local Area Unemployment Statistics, and data on establishment counts and employment totals from the US Census's County Business Patterns dataset. We also use state government expenditures from the US Census's Annual Survey of State and Local Government Finances. Finally, we use data on federal hiring from FedScope, provided by the US Office of Personnel Management.

5 Setting 1: EPA Budget Cuts (The Clean Air Act)

With evidence that EPA cares less about the economic harms of enforcement, we now turn to the empirical test implied by our model. We focus on Clean Air Act enforcement and exploit budget cuts the US EPA faced after the 2011 Budget Control Act.[29]

[29]These budget cuts likely affected many (if not most) of EPA's enforcement programs. We focus on the Clean Air Act because of data availability and context. Unlike the water programs, state penalty data is reliable dating back to at least 2002. Unlike hazardous substance and waste programs (Superfund and the Resource Conservation and Recovery Act (RCRA)), enforcement actions are high frequency and likely

5.1 Empirical strategy

In the years following the 2011 Budget Control Act, EPA's full-time equivalent workforce fell by almost 20%.[30] EPA budget proposals submitted to Congress during this time explicitly reference their efforts to cut the payroll, and also note that the agency is focusing their enforcement efforts on the worst offenses. For example, the Fiscal Year 2012 EPA Budget in Brief (released in February 2011) begins with the sentence, "The [budget] request reflects the tough choices needed for our nation's short- and long-term fiscal health."

As we show in Section 5.2, US EPA enforcement actions for CAA violations fell significantly in the aftermath of the budget cuts, and average penalties increased, suggesting that federal enforcement actions focused only on the worst cases. We exploit additional variation driven by differences across EPA regional offices, which have famously idiosyncratic enforcement preferences and approachs (Engelberg et al., 2011).[31] CAA enforcement decreased more substantially in some regions than in others; we treat this as treatment intensity and run an event study specification interacting treatment intensity with time period indicators. We also show results using, instead, the reduction in new legal staff hires in each region as a treatment intensity variable.

Ultimately, we are interested in the effect of EPA strength on state outcomes. Thus, while our variation comes from federal budget cuts (i.e., variation in the budget of the US EPA), our main outcome is the size of penalties issued by *state* environmental agencies.

5.1.1 Specification

To test whether and how lower EPA strength affects state penalties, we run the following specification using data on formal enforcement actions issued by states and by EPA:

exhibit less substantial time trends.

[30]From conversations with EPA staff, we understand that much of this was from additional restrictions imposed on hiring new staff.

[31]Discussions with EPA staff suggest that program leadership in regional offices is a significant determinant of regional office enforcement behavior. We do not have data on the priorities of EPA regional office program staff.

$$y_{j,s,t} = \beta D_{t(j)} \times (\text{Regional decrease})_{s(j)} + \delta_{s(j)} + \gamma_{i(j),t(j)} + \Gamma X_{j,s,t} + \epsilon_j \tag{2}$$

$y_{j,s,t}$ is the log of the penalty size in penalty j, which is issued by state $s(j)$ in year $t(j)$, to a firm in industry $i(j)$. D indicates the penalty was issued in the post period (in 2011 or later). We also present an event study version of the regression, combining years into bins for precision.

In all specifications, we include state fixed effects and industry-by-year fixed effects, where an industry accords to 3-digit NAICS codes.[32] The vector of controls $\mathbf{X}_{j,s,t}$ includes indicators for deciles of the state unemployment rate (lagged 3 months); indicators for deciles of facility emissions (from the TRI); facility type controls; the (log of the) number of penalties previously issued to the same facility; and a control for the total annual state expenditures.[33] To increase power, we include enforcement actions issued by local authorities (sub-state authorities); however, we allow these outcomes to vary with the post-period because our model does not provide a clear interpretation for these.[34]

"Regional decrease" is a continuous variable which encodes treatment intensity: the extent to which each EPA regional office was affected by the federal budget cuts. Ideally, in correspondence with the model, our variation would come from the size of the penalty EPA issues. However, we are unable to hold constant the underlying severity of the violations over time, making it difficult to know how much penalties decreased for a given violation. Thus, in our primary specification, we use instead the extent to which each regional office reduced its hiring of legal staff after the budget cuts.[35] Specifically, we define

[32]While states may adjust their enforcement across industries depending on EPA strength, inclusion of industry fixed effects improves precision by addressing changes in US industry composition over time. It also alleviates concerns that state-issued penalty decreases are driven by economic considerations for industries that suffered in the Great Recession.

[33]Facility type controls are an indicator for major facilities (see Section 4) and an indicator for non-public facilities. We control for the number of prior penalties because penalty size increases with the number of prior violations.

[34]The vast majority of enforcement actions issued by local authorities in the data are in California, Pennsylvania, and Washington state.

[35]We define legal staff as staff listed as having a legal occupation in the public-use OPM federal workforce data. We assign staff to regions based on their state of residence. Because we cannot tell who works in EPA headquarters in DC, we exclude staff in DC, Virginia, and Maryland.

$$\text{Regional decrease}_{\text{alt}} = \frac{\text{New Legal Hires 2008-2010}_j - \text{New Legal Hires 2012-2014}_j}{(\text{N facilities in region}_j)/10{,}000} \quad (3)$$

The interpretation, then, is a reduction of 1 new hire per 10,000 CAA facilities corresponds to a β percent reduction in average *state* penalty size. We note that to scale the coefficient, we count all facilities that appeared in EPA's ICIS-AIR facility database at the time of our data download. Approximately 30% of those facilities were no longer operating at the time of the data download, suggesting that we systematically underestimate the number of employees (or the decrease in new hires) per 10,000 active facilities. The average (penalized) facility in our penalty dataset is in a region with around 30,000 CAA facilities.

The identification assumption is that state outcomes would have trended similarly after the budget cuts if not for the reduction in enforcement efforts proxied by reduced legal hiring.

Note that we do not intend to suggest that the entire effect of the budget cuts operates through the reduction in new legal hires. Rather, we use this as a proxy for the extent to which each region's enforcement program was affected by the budget cuts.

5.2 CAA Results

5.2.1 Descriptives

State penalties. Appendix Figure A.4 shows descriptives for the regression sample. Figure A.4B shows the distribution of state-issued penalties, in level terms, separately for before and after EPA's budget cuts. The distribution is right-skewed, with the majority of state-issued penalties falling under \$10,000 in 2010 dollars. Appendix Figure A.4A shows that the plurality of penalties issued are issued to firms in manufacturing industries. Other over-represented industries include oil and gas and utilities. Many of these industries faced secular declines during our sample period, which motivates our inclusion of industry-by-year fixed effects in our main specification.

EPA budget cuts. Our identifying variation uses budget cuts many federal agencies faced after the Budget Control Act of 2011. Figure 3A shows that the EPA workforce declined over 15% in the years following the budget cuts.[36] Figure 3B shows the reduction in new hires, at the regional office level, during this period. Finally, Clean Air Act formal enforcement actions brought by the US EPA declined in the years after the budget cuts (Figure 3C).[37]

As described in EPA annual fiscal year budget proposals, these budget cuts appear to lead the agency to focus its efforts on the worst offenders; i.e., in EPA enforcement decisions, the marginal enforcement cases are the ones with lower penalty sizes. Moreover, consistent with qualitative evidence of regions operating with substantial independence (Engelberg et al., 2011), this targeting happened within region, and not across regions.[38] Figure 4 shows that EPA regional offices with the largest decreases in enforcement activity also have the largest increases in average penalty size. The expansive discretion of regional offices motivates our use of variation in regional office reactions to budget cuts in our identification strategy.

5.2.2 Estimated effect of reduced resources

We begin by presenting an event study with no treatment intensity dimension: we regress penalty size on year with our baseline controls and state fixed effects. Figure 5 shows that average penalty size is unchanging in the pre-period and then drops in the post period. Averaging the estimated coefficients before versus after the budget cuts, penalties dropped by 31% ($p < 0.01$). This is likely an overestimate, since the Great Recession also reduced average state penalties (Appendix C).

To address this confounding factor, we add treatment intensity variation and present the results of estimating the difference-in-difference specification in Equation (2).

[36]While many federal agencies saw budget cuts after the Budget Control Act, other agencies do not appear to have had the workforce declines that EPA did (Appendix Figure A.5).

[37]Weighting by the number of pre-period enforcement actions issued by state agencies, regions with larger hiring reductions also saw larger declines in the number of enforcement actions issued. However, this is not true in an unweighted regression.

[38]In Engelberg et al. (2011), a 2011 review by EPA Office of Inspector General (OIG), the OIG recommended that EPA centralize its enforcement efforts to effectively allocate its resources nationwide instead of within region. Of the five recommendations that the OIG made in this report, this is the only one that EPA disagreed with.

First, we present an event study plot. Figure 6 estimates Equation (2) with biannual year dummies interacted with treatment intensity. We combine the first two year dummies because of inconsistent penalty reporting from states in the early period of the sample.[39] Before the EPA budget cuts, states in regions that are eventually more and less affected have similar trends in average penalty size. After the EPA budget cuts, states in regions where the US EPA decreased legal hiring more collected smaller penalties, on average, compared to states where the US EPA decreased enforcement less. The results are not statistically significant, likely due to the limited variation in the treatment intensity variable, but tell broadly the same story as the repeated cross-section. We note that unlike the typical use case for an event study figure, we (largely) do not follow the same observations over time; rather, each year contains a new draw of facilities. For this reason, we calculate a p-value for Figure 6 pooling the post-period coefficients and the pre-period coefficients ($p = 0.20$).

Table 1 pools the pre- and post-period and estimates Equation (2), showing that the results are robust to the specification. In the first column, we run a regression with sparse controls: only the fixed effects and controls for the state unemployment rate and facility emissions. In the second column, we add facility type controls. In the final column, we also control for state budget. The coefficients imply that reducing hiring of regional legal staff by 1 person per 10,000 CAA facilities led to a 2.5% reduction in average average penalty size for state-issued formal enforcement actions.[40]

In the sample, the average penalty size is $12,828 (with a standard deviation of $13,593).[41] Thus, a decrease of 2.5% of average penalty size corresponds to a reduction of about $321 per penalty. The median region reduced hiring by 3 people per 10,000 CAA facilities, suggesting that state penalties decreased by about $963 for the median region.

In Appendix Table B.1, we explore changes in the distribution of penalties; we replace our main outcome with dummies indicating a penalty is under $Y dollars. Penalties un-

[39] Appendix Figure A.6A shows that the share of penalties coded as zero in the data drops sharply after 2005. Appendix Figure A.6B shows that before 2005, eventual treatment intensity is not correlated with the share of zeros in the data.

[40] We note that this result does not replicate using reduction in *total* hiring at the regional level. However, regional offices can allocate resources within the office, so we believe that legal staff hiring is a better reflection of the effect of the budget cuts on enforcement specifically.

[41] The median penalty is $6,649.

der $5,000 appear more affected than larger penalties. In Appendix Table B.2, we show that capping penalty size at the 85th or at the 95th percentile does not meaningfully affect the results.

5.3 Robustness

Extensive margin responses. EPA strength may have affected the composition of penalties collected, and not only the size of penalties collected. Our model speaks to the final penalty size paid by individual firms—an intensive margin measure. We explore whether extensive margin effects affects the internal or external validity of our results.

Sample selection would affect the internal validity of our estimates, if the types of violations we observe incurring penalties differ before and after EPA budget cuts.

First, we note that many related concerns would bias our effects towards zero. For example, if states are less likely to issue penalties conditional on violation severity when EPA is weaker, we would be missing some low-severity violations in the post period, *increasing* average penalty size in the post period. However, if states are less willing to pursue the most severe violations—e.g., preferring to directly hand them to EPA—then this would cause us the opposite issue: our effects would be overestimates.

To address concerns of sample selection, we first ask whether treatment is associated with a change in the characteristics of facilities penalized. Table B.3 shows that most facility characteristics are not significantly related to treatment status, suggesting that the types of facilities incurring state penalties is not systematically different after EPA budget cuts. Penalized facilities' number of prior penalties is significantly higher in treated state-years. However, the magnitude of this effect is small (1 additional prior penalty on a mean of 18), and the direction of the effect runs counter to the main results (facilities with more prior violations should have larger, not smaller, penalties).

Robust inference Our main analysis uses what Abadie et al. (2023) refer to as a model-based framework: we take the stance that errors are correlated within state over time, and we cluster by state level. An alternative, design-based approach to inference would suggest it is appropriate to cluster standard errors at the level at which treatment is assigned. Considering treatment to be regional EPA office strength, our treatment varies by regional office. However, given the small number of clusters at this level (10), it's not clear how to appropriately calculate standard errors (Roth et al., 2023).

One common approach with few clusters is to present wild bootstrap clustered standard errors. Canay et al. (2021) (as referenced in Roth et al. (2023)) note that the validity of these standard errors requires assumptions about homogeneity of treatment effects across clusters, which in our setting may not hold. Nevertheless, we present these for completeness in Appendix Table B.4.

Heterogeneity in state preferences. Our sufficient statistics result is derived from a model with a representative state. The result is threatened in a model with multiple states, if some states prefer EPA be stronger while others prefer EPA be more lenient.

We show in Appendix E.4 that if all states prefer a stronger EPA or all states prefer a more lenient EPA, the sufficient statistics result holds. We examine this empirically by separating states by political orientation using the results of the 2004 presidential election (Table B.5). We consider a state blue if John Kerry won the state in the 2004 presidential election, and red if George W. Bush did. We use the difference-in-differences regression specification for each and find that the effect of EPA strength on state penalties is negative for both sets of states, suggesting that EPA is weaker than both sets of states would like. The difference between the coefficients is not significant. The lack of meaningful heterogeneity along political lines bolsters our interpretation of the main analysis, that EPA is weaker than the states would like, even though we are averaging over all states in the analysis. We note also that a larger effect among red states is, through the lens of the model, consistent with our priors, as the model predicts that states with preferences closer to EPA's should see larger effects. We expect that if EPA is too lenient for all states, it would be less so for red states, and these states have larger effects.

We note that the next empirical setting–hazardous substances cleanups–has identification within a single state, so the sufficient statistics test is not threatened in this analysis.

6 Empirical Setting 2: Presidential Administrations (Hazardous Substances Cleanups)

Our second source of variation in EPA's strength uses changes in presidential administrations. EPA administrators are appointed by presidents (and confirmed by Congress) and affect the enforcement culture and capacity of EPA; we provide evidence that a Republican-led EPA is a less harsh enforcer. While the Clean Air Act provides a useful context for studying changes in EPA resources, it is not sufficient for exploring effects of EPA leadership for two reasons. First, since the CAA data only span 2002-2020, they only include two changes in presidential administrations, and one was coincident with the Great Recession (which, as we show in Appendix C, independently affected state penalties). Second, we do not have a control group in CAA enforcement we expect to be less affected by changes in presidential administration.

Instead, we turn to a second environmental program: the Superfund program. Unlike the Clean Air Act, Superfund is not a delegated program, but states often have their own cleanup programs and use the federal program to threaten firms (Figure 1). Our cleanup data date back to the late 1980s and include a convenient control group. We use a difference-in-differences design, where our outcome is the pace of cleanup projects. The first difference is variation in presidential administration, which we argue changes firm expectations about the costs of entering the federal Superfund program. The second difference is in whether the cleanup is funded and orchestrated by a private party or state government.

This design allows to test whether changes in presidential administration, which we argue affects EPA strength through both its preferences and its capacity, affects state outcomes.

6.1 Empirical Strategy

6.1.1 Variation in EPA strength

For variation in EPA strength, we use changes in the party of the US president, since the headquarters and regional EPA administrators, as well as the assistant administrator who oversees waste cleanup, are all political appointees. Appointees are, for politicians, a "vital tool for controlling the bureaucracy" (alongside statutes, budget changes, and oversight investigations and hearings) (Lewis, 2010). While the career staff members at EPA, who are on average left-leaning (Clinton et al., 2012; Spenkuch et al., 2021), may stay at the agency through many administrations, the administrators have substantial sway over the the culture, priorities, and capacity of EPA. Republican-appointed administrators tend to be more closely tied to industry, suggesting they may be more business-friendly in their enforcement. Thus, the strength of the "gorilla" threat should be lower under Republican presidential administrations.

Evidence of the qualitative difference in EPA leadership across administrations can be found in EPA press releases giving backgrounds on appointed regional EPA administrators. We found press releases for eight of George W. Bush's initial ten appointees for regional administrator, and nine of Barack Obama's initial ten. Only one of Bush's regional appointees was noted to have worked in the non-profit sector before their EPA service, while four of the eight had worked in private law or business. In contrast, seven of Obama's appointees had non-profit sector experience mentioned in the press releases, and only one of the nine had worked in private law or business.

For quantitative evidence that EPA involvement is more costly for firms during Democratic presidential administrations, we turn to data from federal enforcement across major EPA programs.[42] Figure 7 shows the total sum of penalties assessed annually (in administrative orders with penalties) by the US EPA for violations of the Clean Air Act, the Clean Water Act (CWA), the Resource and Conservation Recovery Act (RCRA), and Superfund, under different presidential administrations. Across the board, EPA collects higher total penalties in Democratic years. Limiting the data to years after 1992, as early data are less

[42]Source: https://echo.epa.gov/facilities/enforcement-case-search

complete, this difference is statistically significant at the 1% significance level for CWA, RCRA, and Superfund enforcement. In CAA, it is not significant at conventional levels (p=0.107).

6.1.2 Orphan sites control group

We exploit a key feature of the cleanup context to form a time-invariant control group: "orphan" sites. In California, orphan sites are sites with no viable parties to orchestrate the cleanup.[43]

Using orphan sites as a control group allows us to identify the effect of EPA preferences on firm cooperation with state enforcement. For both orphan sites and firm-led cleanups, EPA's leniency is changing with the presidential party in power, and in theory, changes in EPA leniency could affect state-led cleanups (especially to the extent that EPA can sanction the state). However, in state-led cleanups, the state has complete control over the pace of environmental cleanups, so that the relevant difference between orphan and firm-led cleanups is the control the state has over how it handles the environmental issue.

In the model, orphan sites can be thought of as cases where the state has total control over the enforcement outcome, p_s: there is no longer a firm to bargain with. States still get some utility from their own cleanup efforts, so they still have a preferred p_s^*. Thus, in the model, any effect of EPA preferences on orphan sites would operate through the state's fear of sanctions from EPA. We view this as a test of the model's assumption that EPA sanctions are not binding on state behavior.

[43] It can happen that this is because of responsible party recalcitrance; however, it is more commonly because an inability to pay. For example, one major cleanup site in California (Alco Pacific, Inc.) became an orphan site after the DTSC determined that the former owner of the defunct recycling facility owed $500,000 to the IRS and $1 million for a different environmental court judgement, was already in the process of selling his home to pay for these, and had no more than $100,000 in capital assets from his recycling company (source: DTSC regional file room records). The state began cleanup itself while it amassed evidence for a court case, and later sued several other potentially responsible parties to recoup its costs.

6.1.3 Outcome: Cleanup completed in a given year

We do not consider financial penalties in the Superfund context. Penalties do not exist for orphan sites, meaning we would not have a control group if we used penalties as an outcome. Projected cleanup costs, another negotiated outcome of enforcement, are difficult to find for sites, and have very large variation driven mostly by the geological and chemical characteristics of the sites. Instead, we use the probability of cleanup activities occurring in a given year, conditional on site age: i.e., cleanup pace. Our conversations with Superfund program staff, as well as prior academic and policy research on Superfund (Environmental Law Institute, 1990; Sigman, 2001), indicate that the pace of cleanup is a measure of firm cooperation. Cleanup pace has costs and benefits in terms of present discounted value; it is more expensive to complete a project faster, but it also confers environmental benefits.

6.1.4 Empirical Specification

We analyze cleanup site-year data from 1987-2016 using the following empirical specification:

$$y_{i,t} = \delta \text{Rep}_t + \beta \text{Rep}_t \times (\text{firm-led site})_i + \rho(\text{site age})_{i,t} + \mu_i + \epsilon_{i,t}, \tag{4}$$

where $y_{i,t}$ is whether site i had a remedial action in year t; Rep_t is an indicator equal to 1 in Republican presidential administration years and 0 otherwise; $(\text{site age})_{i,t}$ gives the number of years since we first observed the site (divided by 10 to make tables readable); and firm-led site_i is a site-level indicate for whether the site's remedial action is orchestrated and funded by a firm (as opposed to the state); and μ_i are site fixed effects. If a stronger EPA increases firm cooperation, we expect $\beta < 0$.

Orphan sites may be different from firm-led sites. However, this empirical specification identifies the effect of EPA leadership on firm cooperation under the assumption that outcome trends do not change differently under different presidential administrations for firm-led versus orphan sites for reasons besides the bargaining power of the state.

Cox proportional hazard model. Our data are censored: once a site has been cleaned up (which may require additional cleanup actions), its outcomes are on longer observed. Thus, the remaining observable sites will be negatively selected. This becomes problematic if the sample selection differs by treatment status.

We include results using a Cox proportional hazards model. The identification assumption for the Cox model in this context is similar to the linear model—treatment status must not be (differentially by orphan vs. non-orphan status) correlated with anything unobserved which also affects cleanup likelihood (Fisher and Lin, 1999). We note that our setting is not a typical use case for a hazard model: in our setting, treatment status flips every 4-8 years and affects all sites at once. Because of our alternating treatment status, we are not obviously subject to the usual problem in survivorship models: that sites are differentially selected in treatment and control.

6.1.5 Hazardous Substances Results

Appendix Table B.6 gives descriptive statistics for the regression sample and the full sample (relaxing the acreage restriction). Sites in our sample are more likely to have cleanup activities than in the full sample. By construction, they are much larger. Within our regression sample, orphan sites are less likely to have any cleanup activities or be certified during the sample period, and they tend to be smaller on average than firm-led cleanups. We note that our identification strategy does not rely on orphan and firm-led sites being similar, but rather on an assumption that the pace of their cleanups does not differently change with presidential administration for reasons besides EPA strength.

In Table 2, we report the results of the main difference-in-differences regression (Equation 4). Overall (Column 1), sites under DTSC oversight are less likely to have remedial actions in years when US EPA is headed by a Republican appointee. This effect is driven entirely by firm-led sites; i.e., where the cleanup is orchestrated by the firm and not the state. Firm-led sites are three percentage points less likely to have a remedial action in a given year if the EPA is led by a Republican appointee ($b = 0.03$, $p < 0.01$). Orphan sites, which the state cleans, are not significantly more less likely to have remedial actions. Furthermore, as shown in Column 4, this difference is statistically significant ($p < 0.01$).

We next explore whether cleanups are more thorough (and expensive) during Democratic presidential administrations. Here, we have limited data, and lack power. In particular, we only have data on 11 orphan sites. However, we present these results for two reasons: the costs to the firm (and state) more closely align with our model, and it also allows us to test whether cleanup *quality* is affected by EPA preferences.

Table 3, Column 1, shows that firms choose less expensive remedial alternatives during Republican EPA administrations. The difference is large (over one-third of a standard deviation, $p = 0.026$). Column 2 includes the 11 orphan sites for which we were able to obtain these data. Even compared to the state's choices, firms still choose lower cost projects during Republican administrations, but the relationship is smaller (about 1/5 of a standard deviation) and less significant ($p = 0.241$).

Appendix Table B.7 shows that firms do *not* choose less costly remedial actions when compared to the other alternatives they present to the state as options. That is, the ranking of the project they choose among possible projects they present is not significantly lower during Republican administrations. However, we note that this outcome is likely endogenous, in that firms can choose which projects to consider. Indeed, controlling for the cost of the least expensive remedial project, the most extensive clean-up options the firms present are significantly lower-cost than the most extensive clean-ups the state presents (1/10 of a standard devation, $p = 0.025$), despite that firm-led sites tend to be larger and more expensive than orphan sites.

6.1.6 Robustness

Censored data. One concern about our difference-in-differences model is that the resulting estimates are biased by a censoring problem. Sites that remain in the panel for longer are negatively selected: they may be harder to remediate or less of a priority.

First, we point out that sites' older ages are less likely to be under Republican administrations. Appendix Figure A.7 shows how the time patterns of site discovery and presidential administrations interact: sites are most likely to have a Republican EPA when they are young, and the least likely to have a Republican EPA when they are old. If the oldest sites are the most difficult to clean up, then older sites being under Democratic EPAs would bias us towards a positive effect of Republican EPAs.

We also run our regressions using a Cox proportional hazards model (Table B.8). The results are similar to our linear model—firm-led sites are significantly less likely to have remedial actions during Republican administrations.

State policy changes. State policy changes may be correlated with federal policy changes (for reasons besides a concern about federal involvement). We note that our orphan sites outcome should be affected by state policy changes, so that this only presents a threat to identification if state policy changes differentially for firm-led versus orphan sites under Democratic versus Republican US presidencies. In Table B.9, Panel A, we add controls for the political party of the California state governor; these do not meaningfully change our results.

Data decisions. Appendix Figure A.8 shows robustness to varying the acreage threshold for our site sample. The difference in clean-up probability across presidential administrations for firm-led sites remains highly significant. The magnitude of the interaction term (comparing the effect of presidential party for firm-led and orphan sites) shrinks somewhat at a threshold of 5 acres, and then remains stable as the sample is further restricted. Appendix Table B.10 shows two additional robustness checks. Panel A outcomes replace the dates cleanup projects were approved by DTSC as complete with the dates found in documents that indicate when the projects *started*. Panel B includes as outcomes removal actions in addition to remedial actions. Removal actions are less expensive projects than remedial actions, reducing the firms' incentive to avoid beginning these projects, so we expect a smaller treatment effect with this outcome. However, these projects are more frequent and address a potential concern that our orphan sites treatment effect is low only because of a low base value.

7 Conclusion

In this paper, we document several novel empirical findings. In two different settings, we show that characteristics of EPA which affect its enforcement behavior also affect enforcement outcomes for state environmental agencies. This is consistent with a characterization, originally proposed at EPA, of a federal agency as a "gorilla in the closet" for the states in regulatory enforcement. We model the "gorilla" as affecting firm offers in a bargaining game between the state and its regulated entities, and we show how changing the resources or preferences of the federal agency can change firm offers to the state. State primacy (i.e., that states can choose EPA involvement in enforcement) provides us with a sufficient statistics characterization of the states' welfare effects from increasing federal strength, and our empirics reveal that EPA is currently less strong than what would maximize state welfare.

How much of EPA's total effect on environmental penalties is through its effects on state penalties? EPA's own penalties collected within our sample of facilities fell by about $4.8 million.[44] Our estimates suggest that states lost $1.5 million annually after EPA's budget cuts. This suggests that almost one-third of EPA's role in CAA enforcement is through the "gorilla" effect. Work that characterizes the benefit of federal regulation and enforcement, environmental or otherwise, should not neglect this spillover onto the states.

Finally, these results are increasingly relevant for today's most pressing environmental issues, and in today's political climate. As the U.S. federal government struggles to pass major federal climate legislation, state and local governments are "stepping up" to address demand (Astor, 2022). There are myriad reasons federal climate legislation may be desirable. Our results suggest that states' efforts will be less effective without backup from a federal authority.

[44]To account for the lag in the effect of the budget cuts on penalties, we calculate this using the average annual federal penalty sum from 2014 through 2016, compared to the average annual federal penalty sum 2008-2010.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge, "When Should You Adjust Standard Errors for Clustering?," *The Quarterly Journal of Economics*, 2023, *138* (1), 1–35.

Agarwal, Sumit, David Lucca, Amit Seru, and Francesco Trebbi, "Inconsistent regulators: Evidence from banking," *The Quarterly Journal of Economics*, 2014, *129* (2), 889–938.

Agrawal, David R., William H. Hoyt, and John D. Wilson, "Local Policy Choice: Theory and Empirics," *Journal of Economic Literature*, 2022, *60* (4), 1378–1455.

_ , _ , and Tidiane Ly, "A New Approach to Evaluating the Welfare Effects of Decentralized Policies," 2023. https://papers.ssrn.com/sol3/papers.cfm?abstract$_id$ = 4009872.

Allcott, Hunt, Sendhil Mullainathan, and Dmitry Taubinsky, "Energy policy with externalities and internalities," *Journal of Public Economics*, 2014, *112*, 72–88.

Astor, Maggie, "As Federal Climate-Fighting Tools Are Taken Away, Cities and States Step Up," *The New York Times*, July 2022.

Baily, Martin, "Some Aspects of Optimal Unemployment Insurance," *Journal of Public Economics*, 1978, *10* (3), 379–402.

Barro, Robert J. and David B. Gordon, "Rules, Discretion, and Reputation in a Model of Monetary Policy," *Journal of Monetary Economics*, 1983, *12* (1), 101–121.

Becker, Gary S., "Crime and Punishment: An Economic Approach," *Journal of Political Economy*, 1968, *76* (2), 169–217.

Besley, Timothy and Stephen Coate, "Centralized Versus Decentralized Provision of Local Public Goods: A Political Economy Approach," *Journal of Public Economics*, 2003, *87*, 2611–2637.

Blundell, Wesley, "When Threats Become Credible: A Natural Experiment of Environmental Enforcement from Florida," *Journal of Environmental Economics and Management*, 2020, *101*, 102288.

_ , Gautam Gowrisankaran, and Ashley Langer, "Escalation of Scrutiny: The Gains from Dynamic Enforcement of Environmental Regulations," *American Economic Review*, 2020, *110* (8), 2558–2585.

Buchanan, James M., "Federalism and Fisal Equity," *American Economic Review*, 1950, *40* (4), 583–599.

Calabrese, Stephen, Dennis Epple, and Richard Romano, "Inefficiencies from Metropolitan Political and Fiscal Decentralization: Failures of Tiebout Competition," *Review of Economic Studies*, 2012, *79* (3), 1081–1111.

Callander, Steven and Bård Harstad, "Experimentation in Federal Systems," *The Quarterly Journal of Economics*, 2015, *130* (2), 951–1002.

Canay, Ivan A., Andres Santos, and Azeem M. Shaikh, "The Wild Bootstrap with a "Small" Number of "Large" Clusters," *Review of Economics and Statistics*, 2021, *103* (2), 346–363.

Chang, Howard F., Leah G. Traub, and Hilary Sigman, "Endogenous Decentralization in Federal Environmental Policies," *International Review of Law and Economics*, 2014, *37*, 39–50.

Chay, Kenneth Y. and Michael Greenstone, "The Impact of Air Pollution on Infant Mortality: Evidence from Geographic Variation in Pollution Shocks Induced by a Recession," *The Quarterly Journal of Economics*, 2003, *118* (3), 1121–1167.

Chetty, Raj, "A General Formula for the Optimal Level of Social Insurance," *Journal of Public Economics*, 2006, *90* (10-11), 1879–1901.

_ , "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annual Review of Economics*, 2009, *1* (1), 451–488.

Clinton, Joshua D., Anthony Bertelli, Christian R. Grose, David E. Lewis, and David C. Nixon, "Separated Powers in the United States: The Ideology of Agencies, Presidents, and Congress," *American Journal of Political Science*, 2012, *56* (2), 341–354.

Currie, Janet, Michael Greenstone, and Enrico Moretti, "Superfund cleanups and infant health," *American Economic Review*, 2011, *101* (3), 435–41.

Dijkstra, Bouwe R. and Per G. Fredrikssonn, "Regulatory Environmental Federalism," *Annual Review of Resource Economics*, 2010, *2*, 319–339.

DTSC, CA, "Our History," February 2023. https://dtsc.ca.gov/about-dtsc/our-history/.

Earnhart, Dietrich and Lana Frieson, "Enforcement Federalism: Comparing the Effectiveness of Federal Punishment versus State Punishment," *Environmental and Resource Economics*, 2021, *78* (227-255).

Engelberg, Dan, Kathlene Butler, Charles Brunton, Anthony Chirigotis, Julie Hamann, Tiffine Johnson-Davis, and Genevieve Soule, "EPA Must Improve Oversight of State Enforcement," Report 12-P-0113, U.S. Environmental Protection Agency Office of Inspector General December 2011.

Environmental Law Institute, "Enhancing State Superfund Capabilities: Nine-State Study," Technical Report, U.S. Environmental Protection Agency 1990.

_ , "An Analysis of State Superfund Programs: 50-State Study, 2001 Update," Technical Report 2002.

EPA, U.S., "Clean Air Act Stationary Source Civil Penalty Policy," Available online 1991. https://www.epa.gov/sites/default/files/documents/penpol.pdf.

EPA.gov, "The Guardian: EPA's Formative Years, 1970-1973," 1993. https://www.epa.gov/archive/epa/aboutepa/guardian-epas-formative-years-1970-1973.html.

_ , "Interview with Henry Longest," Superfund 25th Anniversary Oral History Project August 2005.

_ , "Interview with Lois Gibbs," Superfund 25th Anniversary Oral History Project August 2005.

Evans, Mary F., "The Clean Air Act Watch List: An Enforcement and Compliance Natural Experiment," *Journal of the Association of Environmental and Resource Economists*, September 2016, *3* (3), 525–778.

_ **and Sarah L. Stafford**, "The Clean Air Act Watch List and Federal Oversight of State Enforcement Efforts," *Journal of Environmental Economics and Management*, 2019, *93*, 170–184.

_ **, Scott M. Gilpatric, and Jay P. Shimshack**, "Enforcement Spillovers: Lessons from Strategic Interactions in Regulation and Product Markets," *The Journal of Law & Economics*, November 2018, *61* (4), 555–769.

Farhi, Emmanuel and Xavier Gabaix, "Optimal Taxation with Behavioral Agents," *American Economic Review*, 2020, *110* (1), 298–336.

Feng, Kuishuang, Steven J. Davis, Laixing Sun, and Klaus Hubacek, "Drivers of the US CO_2 Emissions 1997-2013," *Nature Communications*, 2015, *6* (7714).

Finkelstein, Amy, Matthew J. Notowidigdo, Frank Schilbach, and Jonathan Zhang, "Lives vs. Livelihoods: The Impact of the Great Recession on Mortality and Welfare," 2023. http://users.nber.org/ñotom/research/FNSZ_Great_Recession_july2023.pdf.

Fiorino, Daniel J., *Making Environmental Policy*, Berkeley: University of California Press, 1995.

Fisher, Lloyd D. and D. Y. Lin, "Time-Dependent Covariates in the Cox Proportional-Hazards Regression Model," *Annual Review of Public Health*, May 1999, *20*, 145–157.

Gamper-Rabindran, Shanti and Christopher Timmins, "Hazardous waste cleanup, neighborhood gentrification, and environmental justice: Evidence from restricted access census block data," *American Economic Review*, 2011, *101* (3), 620–24.

_ **, Ralph Mastromonaco, and Christopher Timmins**, "Valuing the benefits of superfund site remediation: Three approaches to measuring localized externalities," Technical Report, National Bureau of Economic Research 2011.

Greenstone, Michael and Justin Gallagher, "Does hazardous waste matter? Evidence from the housing market and the superfund program," *The Quarterly Journal of Economics*, 2008, *123* (3), 951–1003.

Gutiérrez, Germán and Thomas Philippon, "How EU Markets Became More Competitive Than US Markets: A Study of Institutional Drift," 2019. http://germangutierrezg.com/GutierrezPhilippon_Europe_2019.pdf.

Harstad, Bård, "Harmonization and Side Payments in Political Cooperation," *The American Economic Review*, 2007, *97* (3), 871–889.

Hendren, Nathaniel, "The Policy Elasticity," *Tax Policy and the Economy*, 2016, *30*, 51–89.

—, "Measuring Ex Ante Welfare in Insurance Markets," *Review of Economic Studies*, 2021, *88*, 1193–1223.

Jenkins, Robin, Heather Klemick, Elizabeth Kopits, and Alex Marten, "U.S. Emergency Response and Removal: Superfund's Overlooked Cleanup Program," *Review of Environmental Economics and Policy*, 2012, *6* (2), 278–297.

Jung, Juergen and Michael D. Makowsky, "The Determinants of Federal and State Enforcement of Workplace Safety Regulations: OSHA Inspections 1990-2010," *J Regul Econ*, 2014, *45* (1), 1–33.

Kang, Karam and Bernardo S. Silveira, "Understanding Disparities in Punishment: Regulator Preferences and Expertise," *Journal of Political Economy*, 2021, *129* (10), 2947–2992.

Kleven, Henrik, "Sufficient Statistics Revisited," *Annual Review of Economics*, 2021, *13*, 515–538.

Konisky, David M. and Neal D. Woods, "Exporting Air Pollution? Regulatory Enforcement and Environmental Free Riding in the United States," *Political Research Quarterly*, December 2010, *63* (4), 771–782.

Lewis, David E., *The Politics of Presidential Appointments: Political Control and Bureaucratic Performance*, Princeton University Press, 2010.

Lim, Claire S. H. and Ali Yurukoglu, "Dynamic Natural Monopoly Regulation: Time Inconsistency, Moral Hazard, and Political Environments," *Journal of Political Economy*, 2018, *126* (1), 263–312.

Lipscomb, Molly and Ahmed Mushfiq Mobarak, "Decentralization and Pollution Spillovers: Evidence from Re-drawing of County Borders in Brazil," *Review of Economic Studies*, 2017, *84*, 464–502.

McCarthy, James E., "Highway Fund Sanctions for Clean Air Act Violations," Technical Report, Congressional Research Service 1997.

Mintz, Joel A., "Agencies, Congress and Regulatory Enforcement: A Review of EPA's Hazardous Waste Enforcement Effort, 1970-1987," *Environmental Law*, 1988, *18* (4), 683–777.

Norwood, Kellen, "Federal Preemption of State and Local Law," in William Scheiderich, ed., *Municipal Law Deskbook*, American Bar Association, 2015.

Oates, Wallace E., *Fiscal Federalism*, Harcourt Brace Jovanovich, 1972.

Persico, Claudia, David Figlio, and Jeffrey Roth, "The Developmental Consequences of Superfund Sites," *Journal of Labor Economics*, 2022, *38* (4), 1055–1097.

Piketty, Thomas, Emmanuel Saez, and Stefanie Stantcheva, "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, 2014, *6* (1), 230–271.

Rogoff, Kenneth, "The Optimal Degree of Commitment to an Intermediate Monetary Target," *The Quarterly Journal of Economics*, 1985, *100* (4), 1169–1189.

Roth, Jonathan, Pedro H.C. Sant'Anna, Alyssa Bilinski, and John Poe, "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature," *Journal of Econometrics*, 2023, *235* (2), 2218–2244.

Sigman, Hilary, "The Pace of Progress at Superfund Sites: Policy Goals and Interest Group Influence," *The Journal of Law & Economics*, April 2001, *44* (1), 315–343.

_ , "Interational Spillovers and Water Quality in Rivers: Do Countries Free Ride?," *American Economic Review*, September 2002, *92* (4), 1152–1159.

_ , "Transboundary Spillovers and Decentralization of Environmental Policies," *Journal of Environmental Economics and Management*, 2005, *50* (2005), 82–101.

Slatterey, Cailin, "Bidding for Firms: Subsidy Competition in the US," December 2022. Working Paper.

Spenkuch, Jorg L., Edoardo Teso, and Guo Xo, "Ideology and Performance in Public Organizations," *NBER Working Paper No. 28673*, 2021.

Tang, Johnny, "Regulatory Competition in the US Life Insurance Industry," November 2022. Working Paper.

Tiebout, Charles M., "An Economic Theory of Fiscal Decentralization," in Universities-National Bureau Committee for Economic Research, ed., *Public Finances: Needs, Sources, and Utilization*, Princeton University Press, 1961, pp. 79–96.

Tobin, Richard J., "Environmental Protection and the New Federalism: A Longitudinal Analysis of State Perceptions," *Publius: The Journal of Federalism*, Winter 1992, *22*, 93–107.

U.S. EPA, "EPA Policy on Oversight of Delegated Environmental Programs," April 1984. https://www.epa.gov/sites/default/files/2019-12/documents/epa_policy_on_oversight_of_delegated_environmental_programs_1984.pdf.

US General Accounting Office, "Superfund Program: Current Status and Future Fiscal Challenges," Technical Report, Government Accountability Office 2003.

US Senate, "Enforcement of Environmental Laws," 1997. Senate Hearing 105-173.

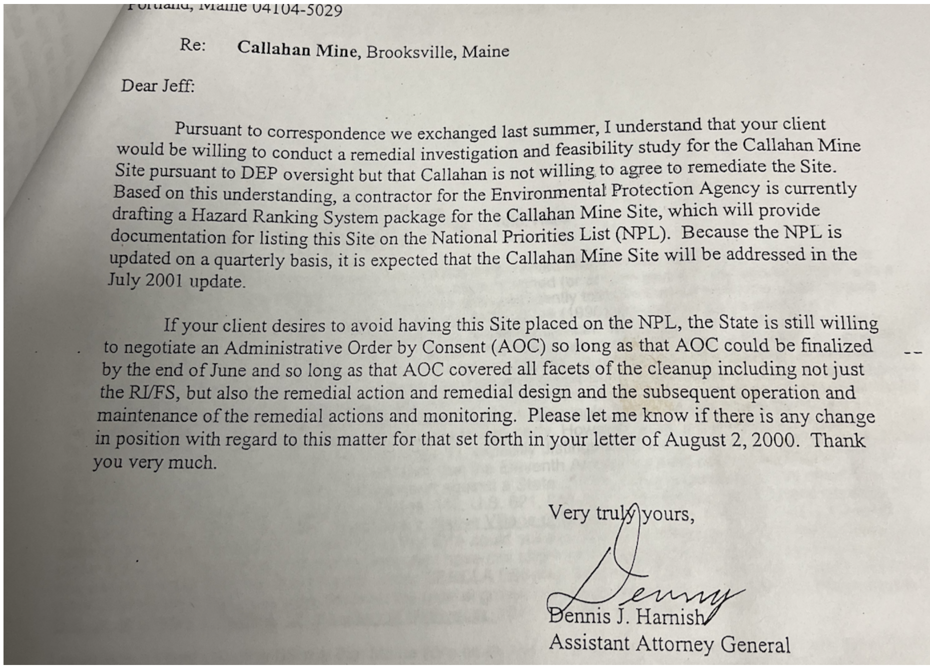
— , "Oversight of the Environmental Protection Agency's Management of the Superfund Program," 2002. Senate Hearing 107-996.

Woods, Neal D., "Interstate Competition and Environmental Regulation: A Test of the Race-to-the-Bottom Thesis," *Social Science Quarterly*, 2006, *87* (1), 174–189.

8 Figures

Figure 1: The Gorilla in Action

(A) Callahan Mine (Maine)



Portland, Maine 04104-5029

Re: **Callahan Mine**, Brooksville, Maine

Dear Jeff:

Pursuant to correspondence we exchanged last summer, I understand that your client would be willing to conduct a remedial investigation and feasibility study for the Callahan Mine Site pursuant to DEP oversight but that Callahan is not willing to agree to remediate the Site. Based on this understanding, a contractor for the Environmental Protection Agency is currently drafting a Hazard Ranking System package for the Callahan Mine Site, which will provide documentation for listing this Site on the National Priorities List (NPL). Because the NPL is updated on a quarterly basis, it is expected that the Callahan Mine Site will be addressed in the July 2001 update.

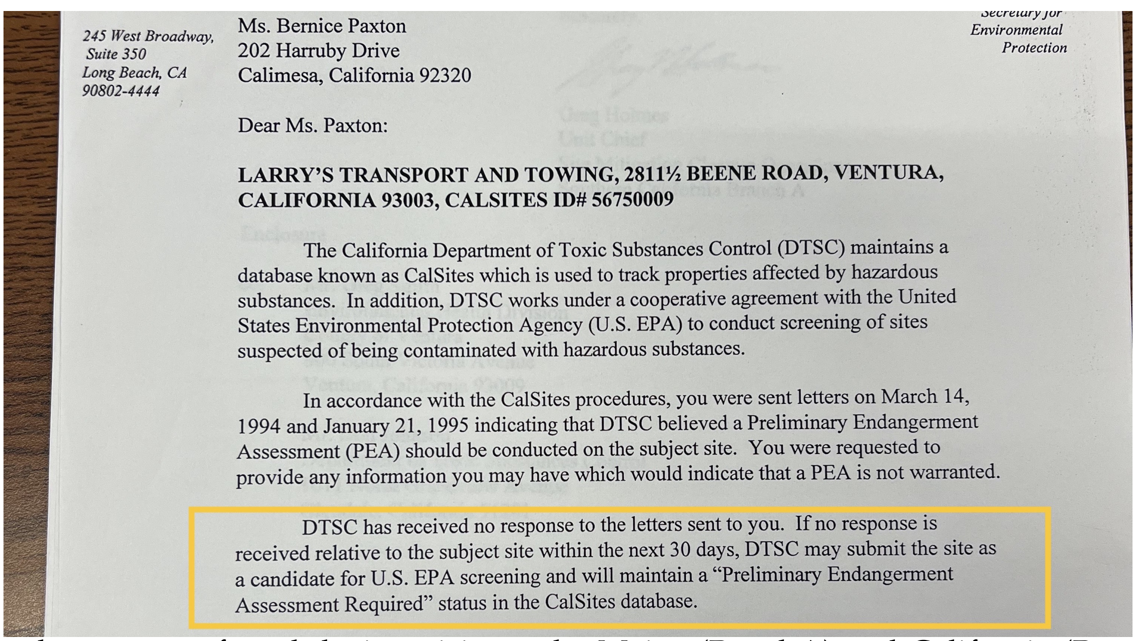
If your client desires to avoid having this Site placed on the NPL, the State is still willing to negotiate an Administrative Order by Consent (AOC) so long as that AOC could be finalized by the end of June and so long as that AOC covered all facets of the cleanup including not just the RI/FS, but also the remedial action and remedial design and the subsequent operation and maintenance of the remedial actions and monitoring. Please let me know if there is any change in position with regard to this matter for that set forth in your letter of August 2, 2000. Thank you very much.

Very truly yours,

Denny

Dennis J. Harnish
Assistant Attorney General

(B) Larry's Truck and Towing (California)



245 West Broadway, Suite 350
Long Beach, CA 90802-4444

Secretary for Environmental Protection

Ms. Bernice Paxton
202 Harruby Drive
Calimesa, California 92320

Dear Ms. Paxton:

LARRY'S TRANSPORT AND TOWING, 2811½ BEENE ROAD, VENTURA, CALIFORNIA 93003, CALSITES ID# 56750009

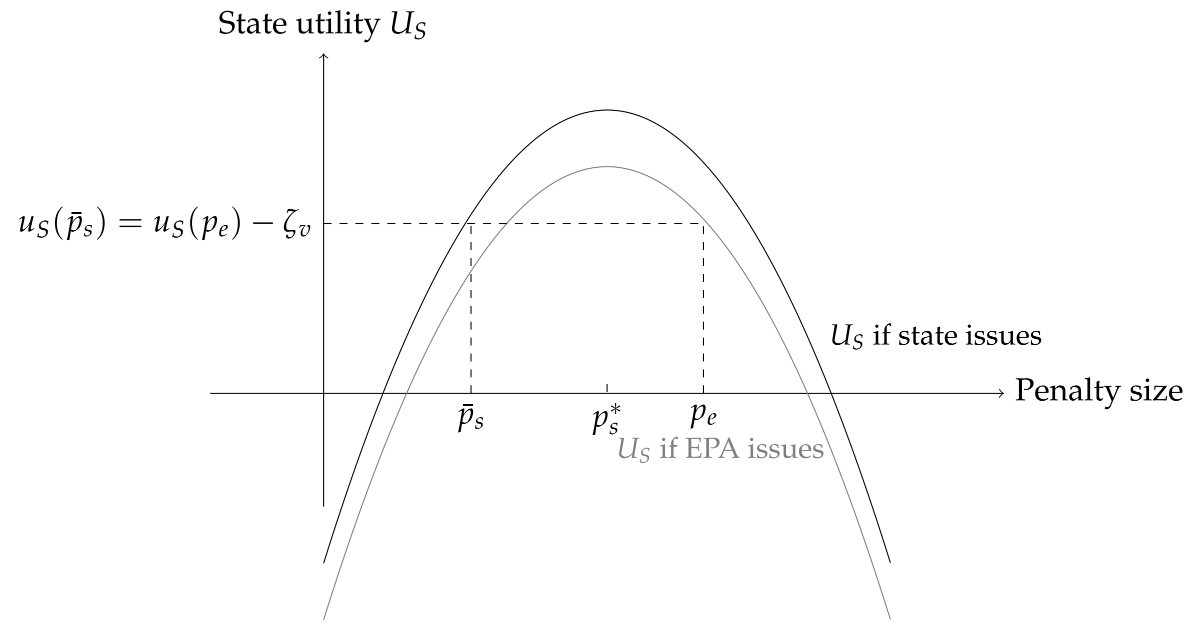
The California Department of Toxic Substances Control (DTSC) maintains a database known as CalSites which is used to track properties affected by hazardous substances. In addition, DTSC works under a cooperative agreement with the United States Environmental Protection Agency (U.S. EPA) to conduct screening of sites suspected of being contaminated with hazardous substances.

In accordance with the CalSites procedures, you were sent letters on March 14, 1994 and January 21, 1995 indicating that DTSC believed a Preliminary Endangerment Assessment (PEA) should be conducted on the subject site. You were requested to provide any information you may have which would indicate that a PEA is not warranted.

DTSC has received no response to the letters sent to you. If no response is received relative to the subject site within the next 30 days, DTSC may submit the site as a candidate for U.S. EPA screening and will maintain a "Preliminary Endangerment Assessment Required" status in the CalSites database.

These letters were found during visits to the Maine (Panel A) and California (Panel B) environmental agency records rooms. We include them as examples of states invoking the "gorilla" in their dealings with companies.

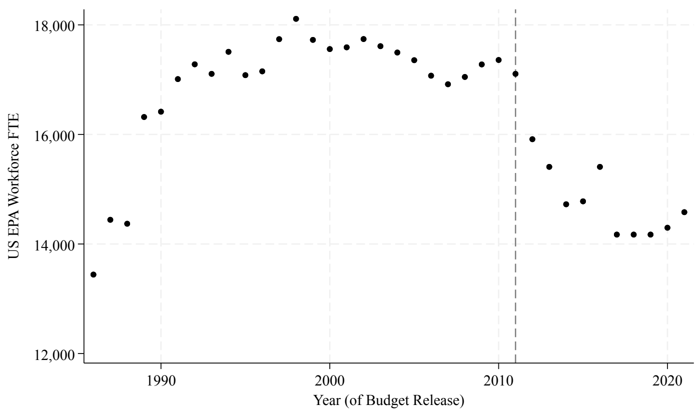
Figure 2: Equilibrium Determination of Firm's Penalty



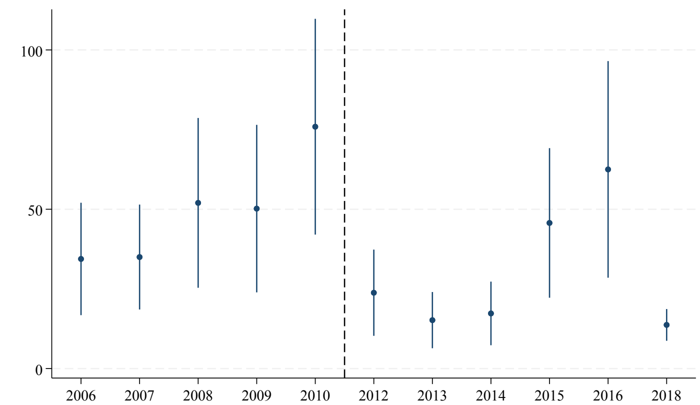
The figure illustrates how, when EPA does not have sanction power, a harsher EPA can decrease equilibrium penalties collected by states. The black (gray) curve shows the state's utility over penalty size when the state (EPA) issues the penalty. EPA penalty p_e (i.e., the penalty the case would receive if the state sent the case to EPA) is marked. The dashed horizontal line indicates the utility level the state receives from sending the case to EPA; where it intersects the state's utility curve if the state handles the case provides the value for the equilibrium firm offer $\bar{p}_s$.

Figure 3: US EPA Budget Cuts

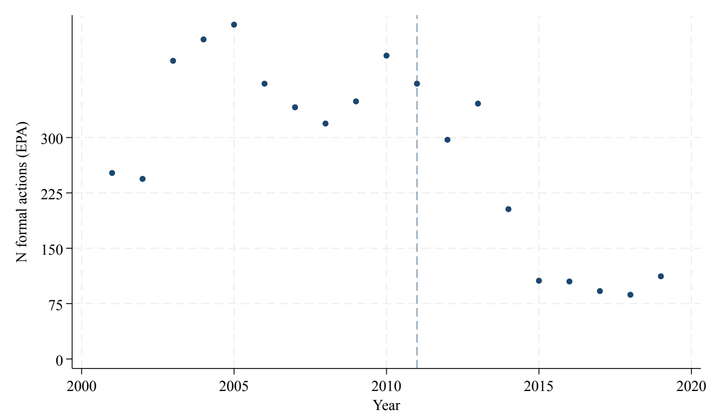
(A) Workforce FTE By Year



(B) EPA New Hires by Region by Year

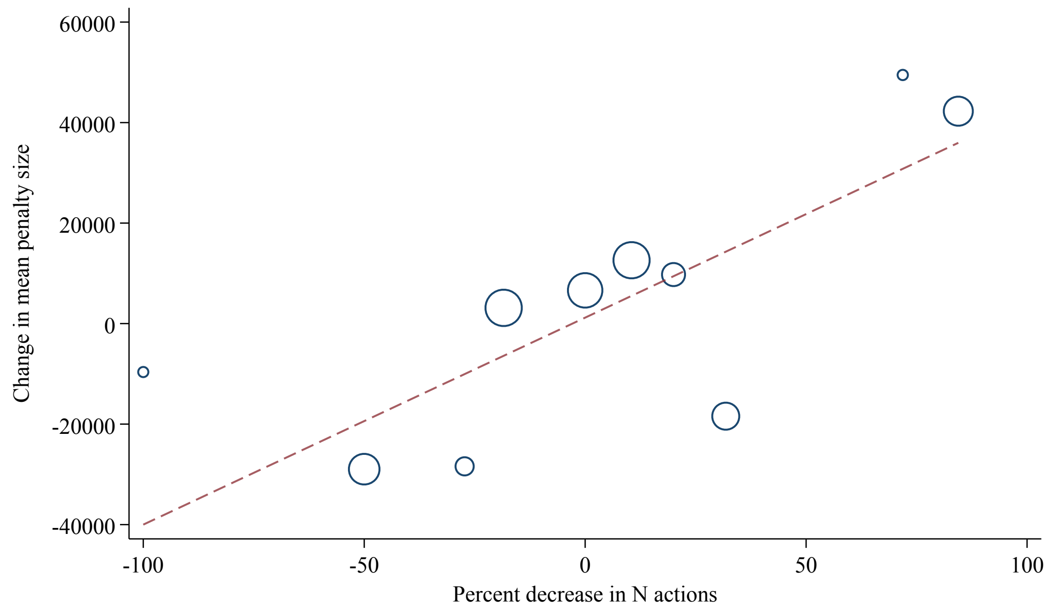


(C) Total US EPA Formal Actions under CAA



Panel A shows the number of workforce FTE budgeted for the US EPA by year. Source: https://www.epa.gov/planandbudget/budget. Panel B shows the number of formal enforcement actions the US EPA settled for Clean Air Act violations in each year. Refer to Appendix D for details on data construction.

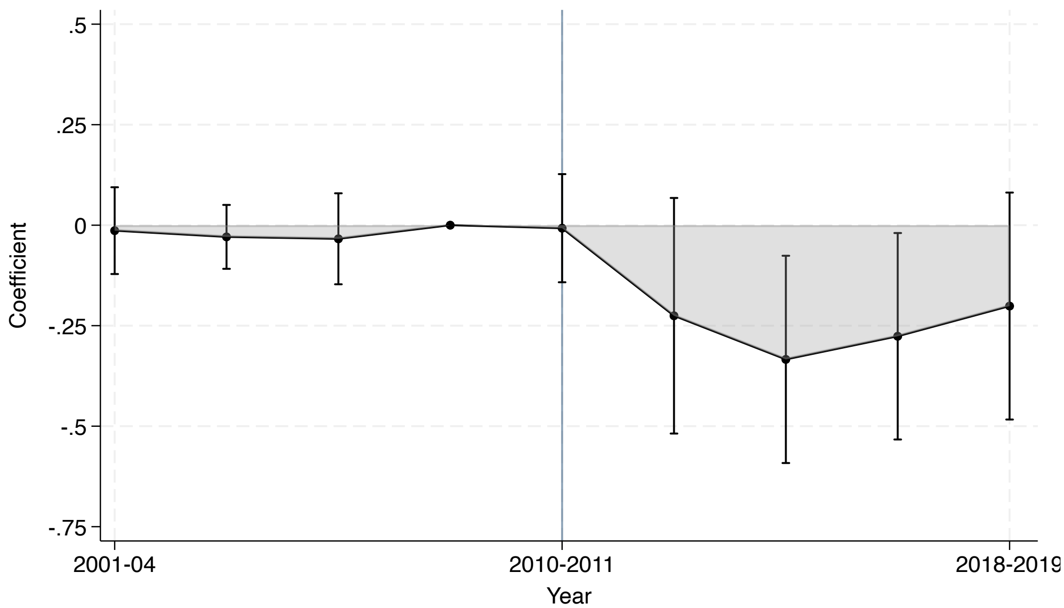
Figure 4: By US EPA Regional Office: Δ N Actions vs. Δ Average Penalty



This figure shows that EPA regional offices with the largest decreases in enforcement actions (in percent terms) also had the largest increases in average penalty size. The x-axis shows the percent *decrease* in number of formal enforcement actions, so that higher values correspond to larger decreases. The x-axis value is calculated the percent difference in the number of formal enforcement actions between 2008-2010 and 2012-2014. The regions are weighted by the number of CAA facilities located in their region across the sample period.

Figure 5: Clean Air Act Results

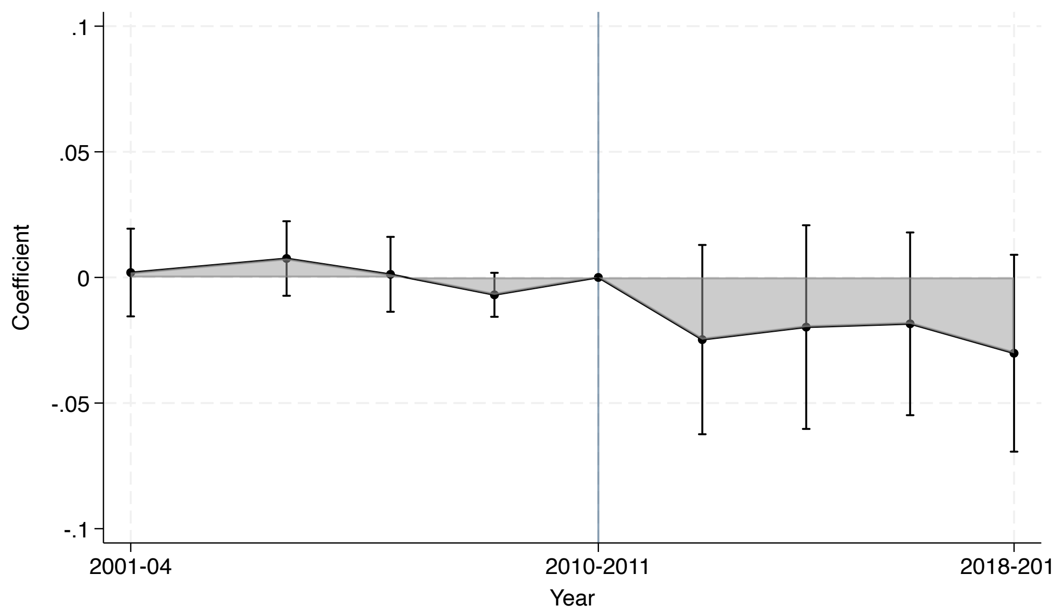
State Penalty Size over Time



The figure plots coefficients from regressing penalty size on year bin dummies, with the full set of controls (see Section 5.1.1).

Figure 6: Clean Air Act Results

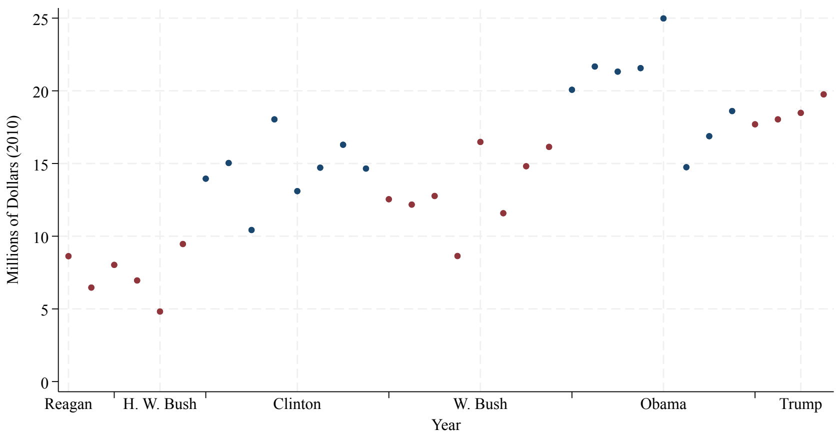
Difference-in-Differences



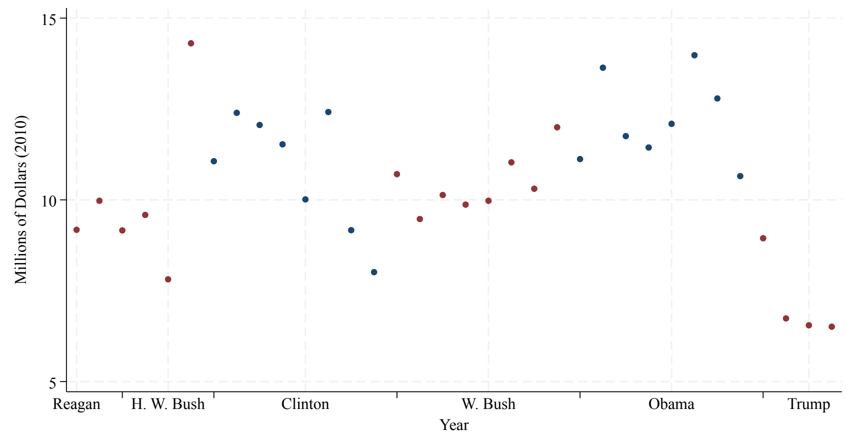
The figure shows the β coefficients from estimating Equation 2 as an event study, binning penalties into two-year bins. The outcome is penalty size. Whiskers show the 95% confidence intervals.

Figure 7: Total US EPA Penalties Collected by Presidential Administration

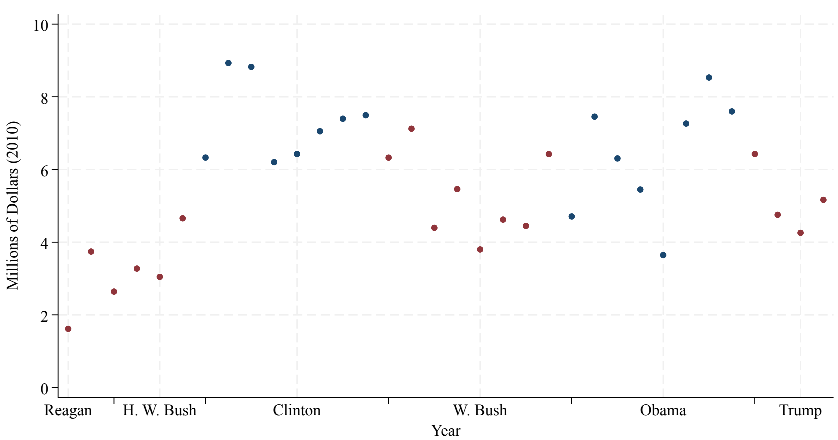
(A) Clean Air Act



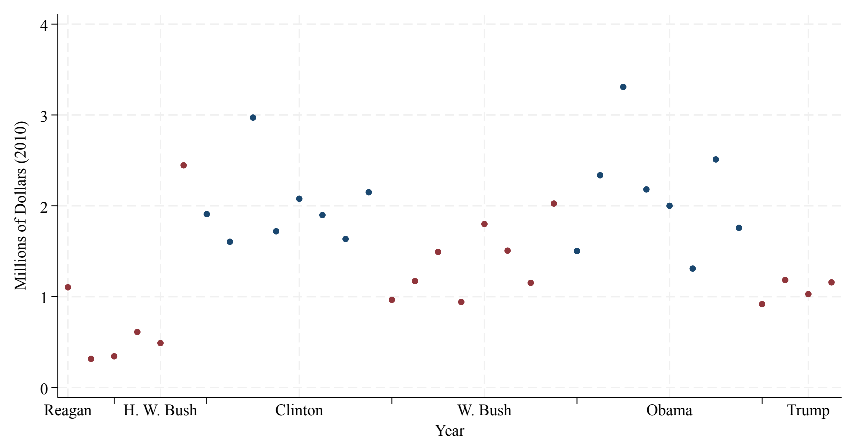
(B) Clean Water Act



(C) Resource Conservation and Recovery Act



(D) Superfund



The figure shows the total dollar amount (in millions, adjusted to 2010 dollars) of penalties issued by the US EPA for violations under each of the major environmental statutes, by presidential administration. The data were downloaded from EPA "Enforcement Case Search" tool. Only penalties assessed in "Final Order with Penalties" cases are included. Cases are assigned to the year they were settled, and penalty amounts winsorized at the 5th and 95th percentile. Cases that were settled more than three years after they were filed are dropped from the data, as are cases missing a filing or settlement date. Early data is less complete; EPA reports that data quality before November 2000 has not been assessed.

9 Tables

Table 1: State Penalty Size (CAA)

	Baseline Controls	+ Facility Controls	+ State Budget Control
Post × Regional decrease	-0.026	-0.030	-0.030
	(0.023)	(0.022)	(0.021)
Major facility		0.168***	0.169***
		(0.050)	(0.050)
Non-public facility		0.047	0.046
		(0.076)	(0.075)
Adjusted action count		0.171***	0.172***
		(0.040)	(0.040)
State budget (per capita)/1k			-0.018
			(0.088)
State FE	X	X	X
Year FE	X	X	X
Industry X Year FE	X	X	X
Penalty Mean	8.79	8.79	8.79
Penalty SD	1.30	1.30	1.30
Obs	23,753	23,753	23,722
R^2	0.26	0.28	0.28

The table shows the results from estimating Equation 2. All columns include state and 3-digit NAICS industry-by-year fixed effects. The post period begins in 2011, and "Regional decrease" gives the treatment intensity, as measured by the regional decrease in federal enforcement actions after the budget cuts. Baseline controls are controls for emissions bins (a proxy for facility size), state unemployment rate bins, and an indicator for sub-state jurisdiction interacted with a post-period indicator. The third column adjusts for state-year budgets, as measured by total state expenditures, from the Census Annual Survey of State and Local Government Finances. Standard errors are clustered by state. Data are from 2001-2019. *, **, *** indicate coefficients are significant at the 10%, 5%, and 1% significance level, respectively.

Table 2: California State Superfund: Probability Remedial Action is Completed

	$\mathbf{1}[\text{Rep Adm}]_t$ Only			Interaction
	Both	Firm-led	Orphan	Both
$\mathbf{1}[\text{Rep Adm}]_t$	-0.027***	-0.032***	0.001	0.004
	(0.006)	(0.007)	(0.016)	(0.013)
$\mathbf{1}[\text{Rep Adm}]_t$ × Firm-led				-0.037***
				(0.014)
Site FE	X	X	X	X
Site Age	X	X	X	X
Mean During Dem Adms	0.05	0.05	0.02	0.05
N Sites	245	209	36	245
N Site-Years	4313	3655	658	4313

The table shows the probability that a cleanup site under California DTSC jurisdiction experienced a remedial action in a given year, using the linear probability model defined in Equation 4. The data are at the site-year level. Sites can experience multiple remedial actions and are removed from the panel when they are certified as no longer requiring additional remediation. The first and final columns use all sites at least 3 acres in size; the second and third columns use subsamples of firm-led and orphan (state-led) cleanups, respectively. All columns have site fixed effects. Standard errors are clustered by site. Data are from 1987-2016. *, **, *** indicate coefficients are significant at the 10%, 5%, and 1% significance level, respectively.

Table 3: Projected Costs of Chosen Remedial Alternatives

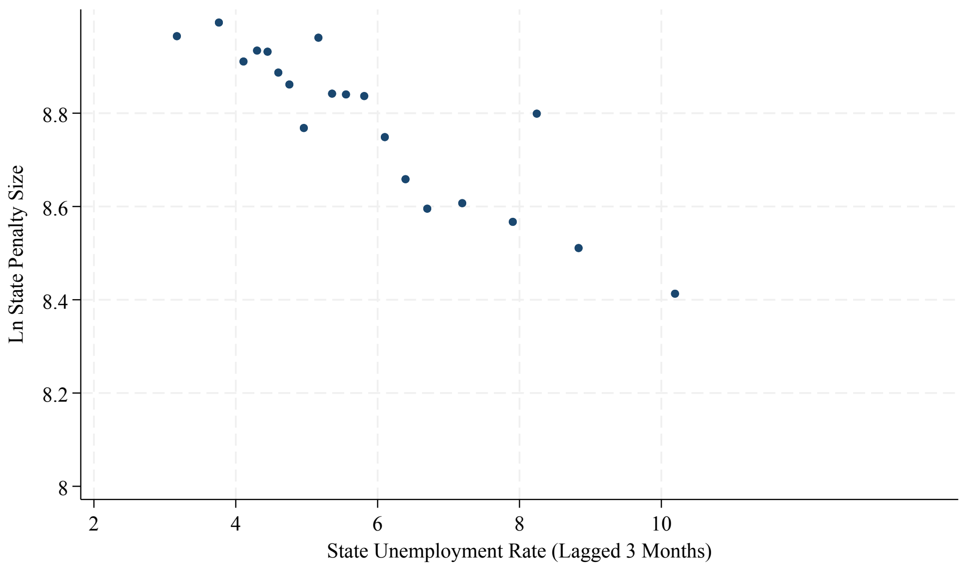
	Firm-led sites	All sites
1[Rep Adm]_*t*	-4.996**	-1.889
	(2.202)	(1.419)
Firm-led		6.615***
		(2.452)
1[Rep Adm]_*t* × Firm-led		-3.108
		(2.636)
N Sites	87	100
Mean During Dem Adms	9.07	8.39
SD During Dem Adms	15.42	14.78
R^2	0.03	0.06

The table shows that firms choose less expensive remedial actions for cleanups under California DTSC jurisdiction during Republican presidencies. The first column includes only firm-led cleanups, and the second column includes all cleanups. The outcome is the project cost of the chosen remedial alternative. The data are at the site level and only include sites under 3 acres where we were able to find information on projected costs. Sites are assigned to years based on when the document with the remedial alternatives was finalized. Standard errors are clustered at the site level. *, **, *** indicate coefficients are significant at the 10%, 5%, and 1% significance level, respectively.

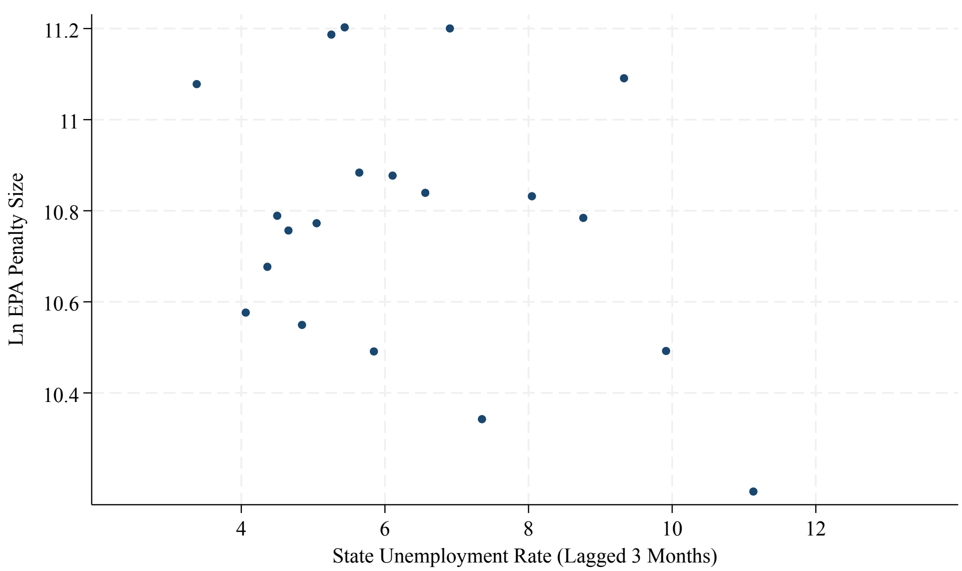
A Appendix Figures

Figure A.1: Standardized Penalties vs. States Unemployment Rates

(A) State Penalties

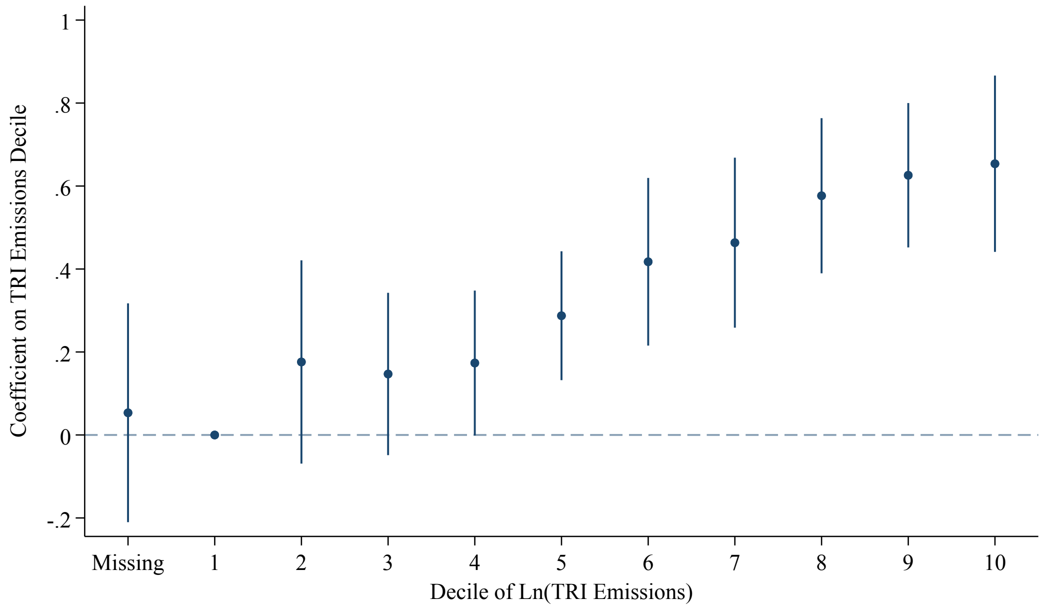


(B) US EPA Penalties



The figure uses data from (log) penalties in formal enforcement actions by states and by EPA for violations of the Clean Air Act by major and synthetic minor stationary sources (see Section 4 for more information). Panel A shows log state-issued penalty size against state unemployment rates, lagged 3 months. Panel B repeats this for EPA-issued penalties. Scatterplots are residualized on state fixed effects and year fixed effects.

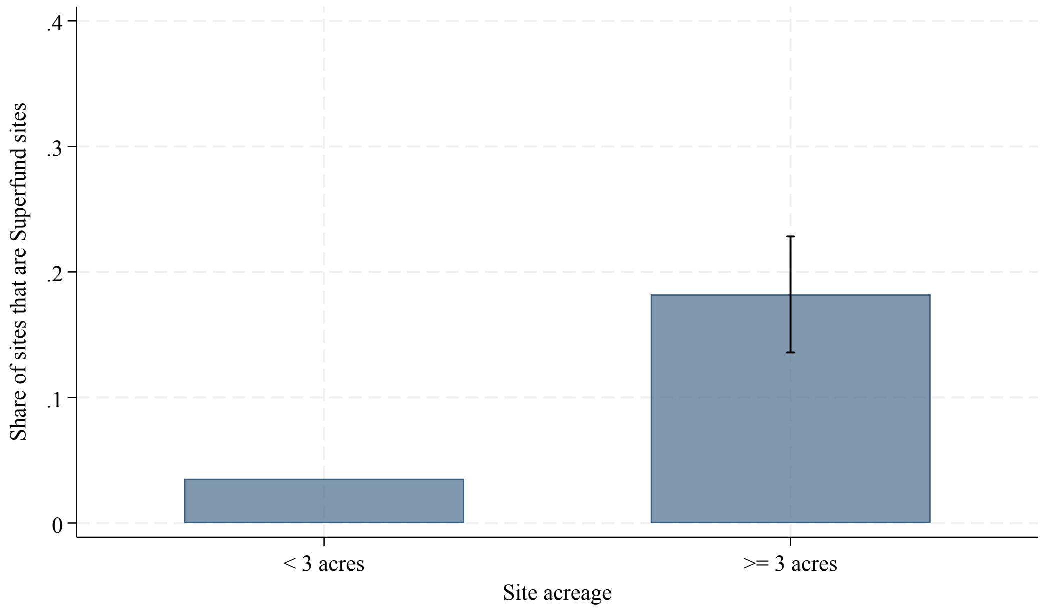
Figure A.2: CAA: Log Penalty Size vs. TRI Emissions



TRI emissions deciles are calculated using the log of total air emissions reported in EPA's Toxic Release Inventory (TRI). The coefficients are on dummies for the TRI deciles (and a dummy for missing TRI data), where outcome is log penalty size. The regression includes state fixed effects and industry-by-year fixed effects.

Figure A.3: Probability of NPL listing

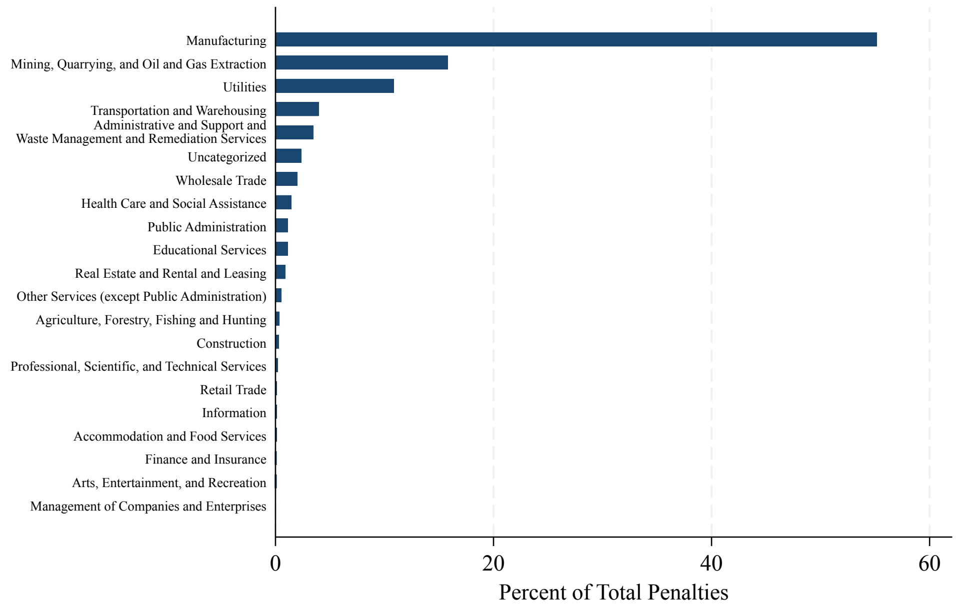
By Site Acreage



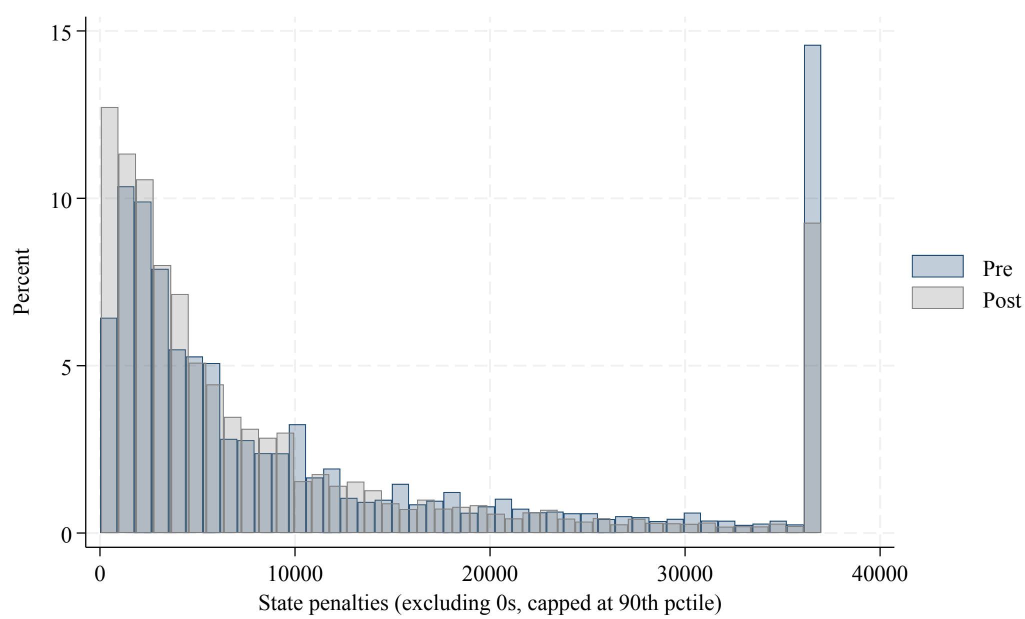
The figure shows that CA DTSC sites over 3 acres are substantially more likely to be listed as federal Superfund sites (i.e., on the National Priorities List) by EPA.

Figure A.4: Clean Air Act Descriptives

(A) Distribution of Industries in CAA Penalty Data

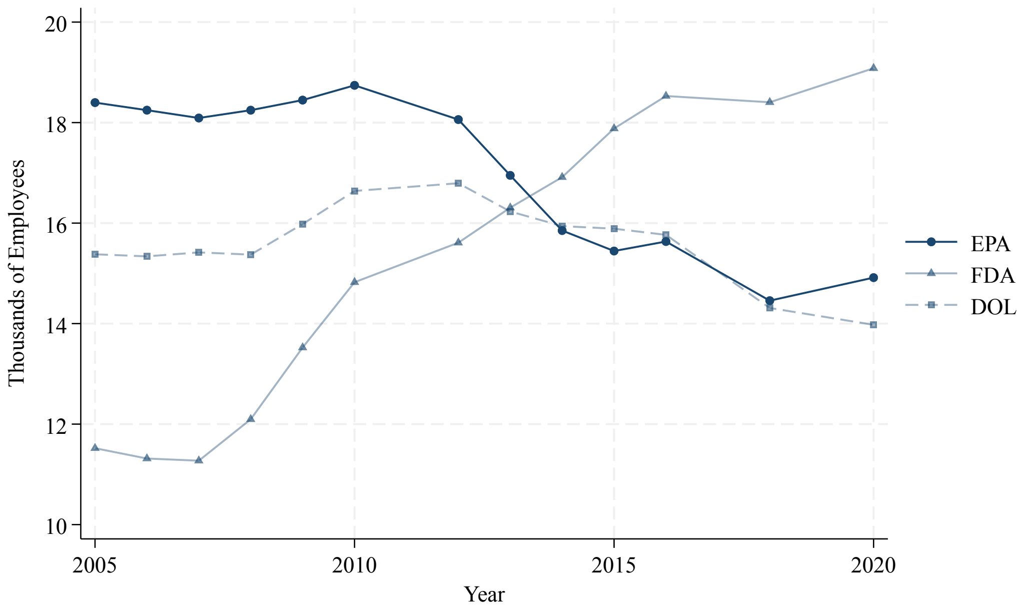


(B) State Penalty Distribution



Panel A shows the share of state-issued CAA penalties which fall into each 2-digit NAICS industry group. Panel B shows the distribution of state penalties before versus after the 2011 EPA budget cuts.

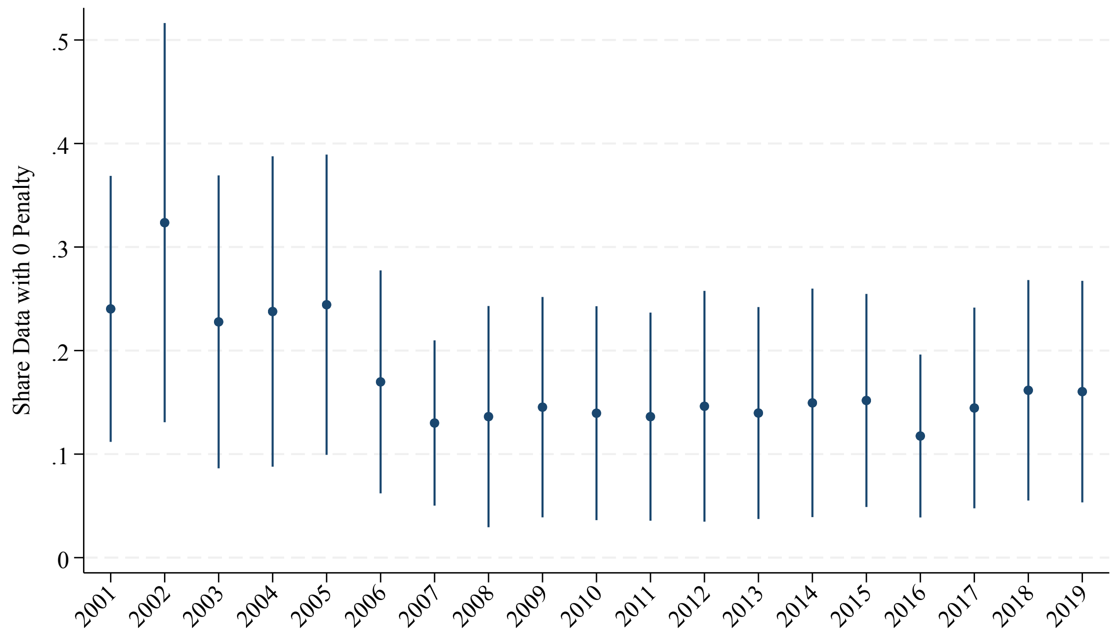
Figure A.5: Workforce Declines: EPA vs. Similar Agencies



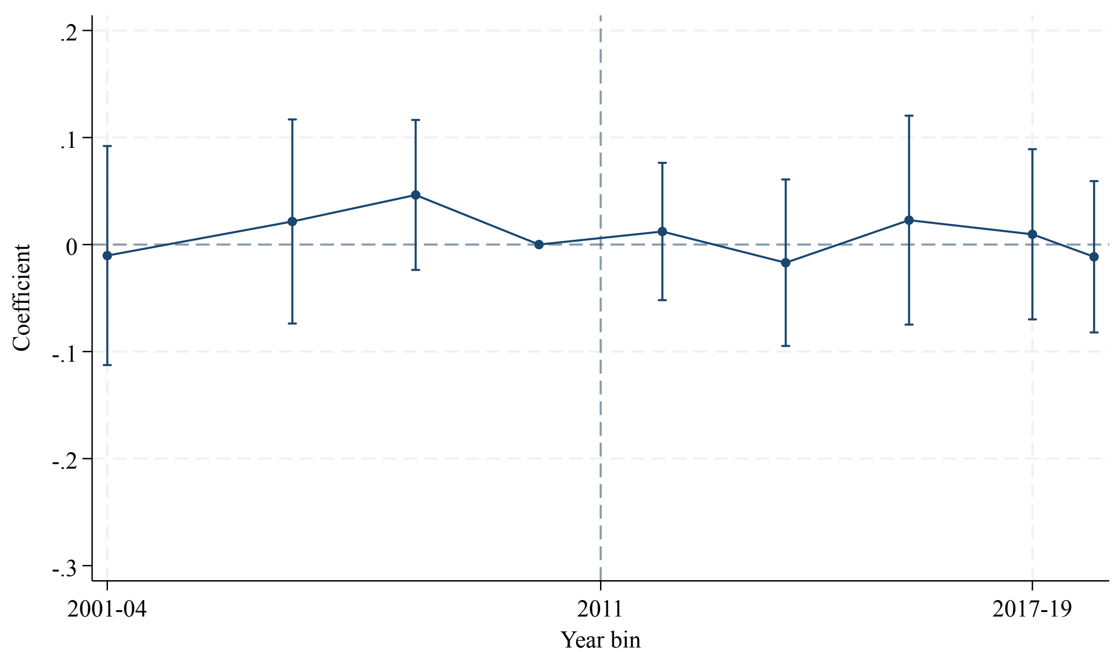
The figure plots total employment in EPA during the sample period against employment in two similarly-sized agencies: the Food and Drug Administration (FDA) and the Department of Labor (DOL). Source: US Office of Personnel Management (Fedscope). Accessed here: https://www.fedscope.opm.gov/

Figure A.6: Inconsistent Penalty Reporting in the Early Sample Period

(A) Share Zero Penalties by Year

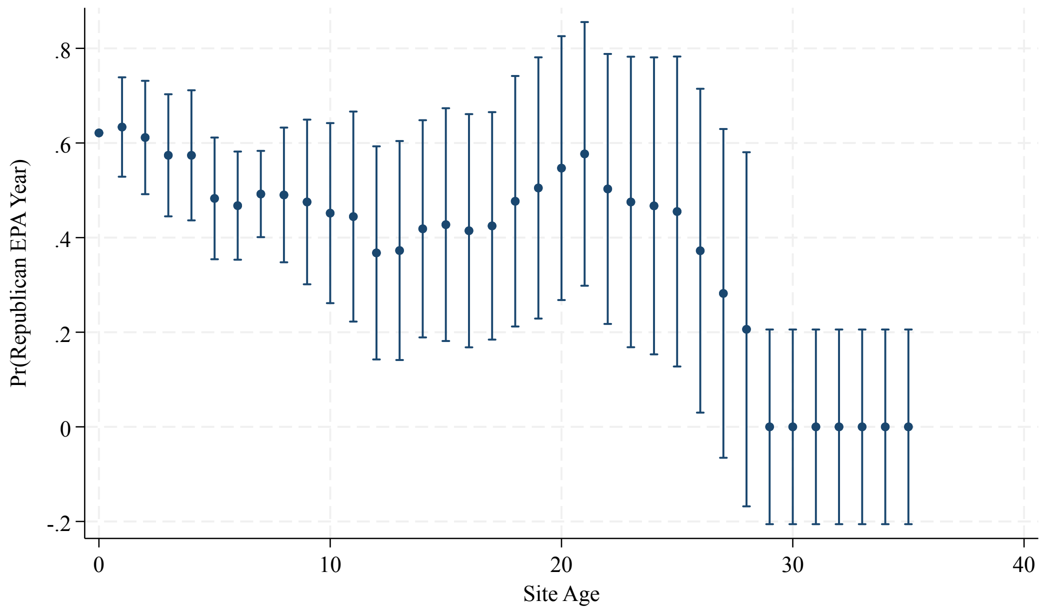


(B) Outcome: Indicator for Penalty Coded as Zero



Panel A shows the share of state-issued formal enforcement actions in ICIS-AIR which have a penalty coded as zero, by year. Zero penalties become much less common starting around 2005-2006. Panel B replicates Figure 6, replacing the outcome with an indicator for a penalty coded as 0.

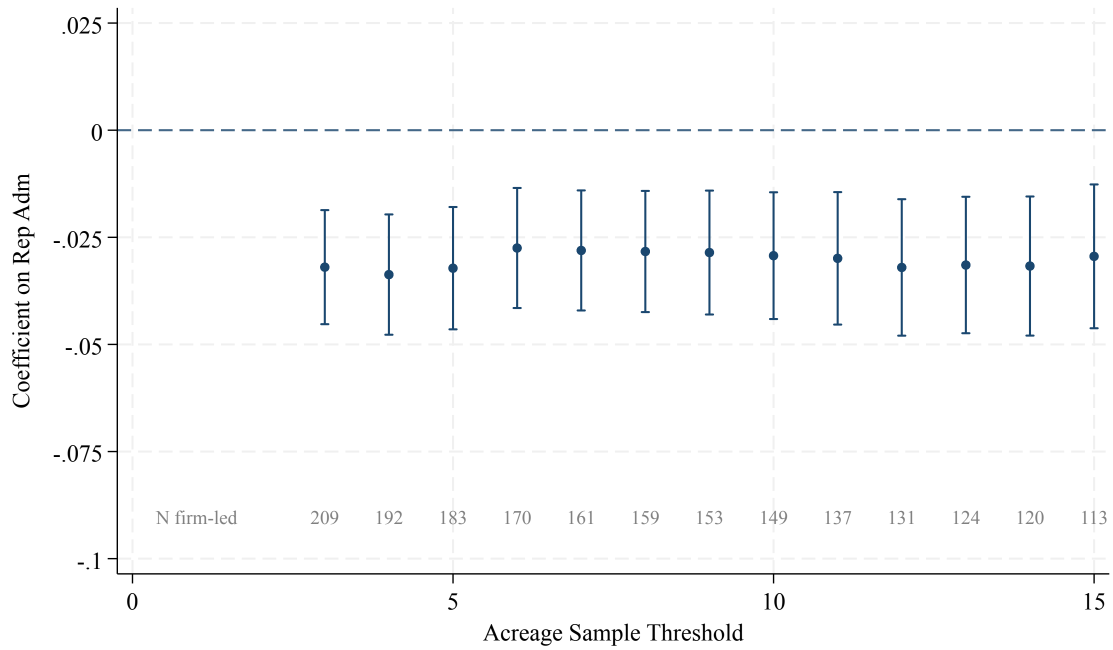
Figure A.7: Probability of Republican EPA by Site Age



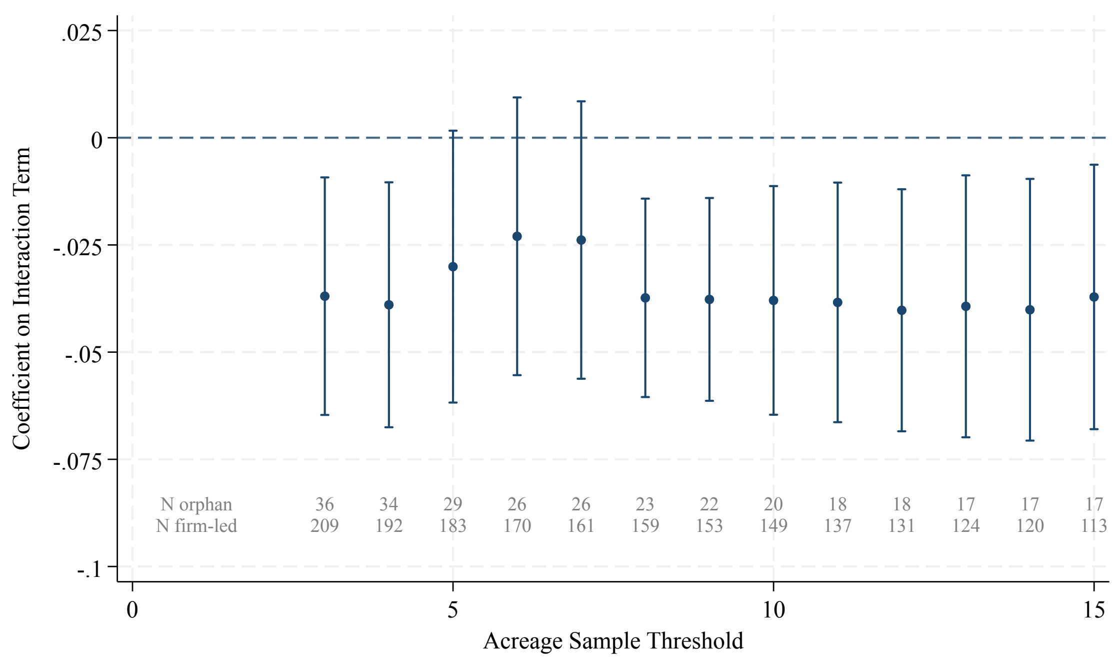
Using the site-year panel, the figure shows the coefficients from a single regression of an indicator for a Republican EPA in a given year on fixed effects for site age in that year. The figure shows that the oldest ages of a site are least likely to be under a Republican EPA. Standard errors are clustered at the year level.

Figure A.8: Superfund: Robustness to Acreage Restriction

(A) Main Effect (Table 2, Column 2)



(B) Interaction Effect (Table 2, Column 4)



The figure shows the coefficients from Table 2 varying the acreage cutoff for the regression sample. The number of orphan and firm-led sites in each regression are displayed in gray at the bottom of each figure.

B Appendix Tables

Table B.1: Clean Air Act: Effects Across Penalty Distribution

	Pr penalty size is...			
	$< 1k$	$< 5k$	$< 10k$	$< 20k$
Post × Legal Hiring Reduction	0.00809*	0.00745	0.00198	0.00200
	(0.00418)	(0.00597)	(0.00486)	(0.00361)
Share Penalties	0.10	0.48	0.66	0.80
Obs	34,720	34,720	34,720	34,720
R^2	0.15	0.20	0.20	0.18

Using the regression specified in Equation 2, this table shows effects on penalty sizes throughout the penalty distribution. States in EPA regions with larger enforcement decreases have overall decreases in penalty size; however, this decrease is concentrated among medium-sized penalties. *, **, *** indicate coefficients are significant at the 10%, 5%, and 1% significance level, respectively.

Table B.2: Clean Air Act: Sensitivity to Capping

	Capped at 85th pctile	Capped at 95th pctile
Post $\times$ Legal Hiring Reduction	-0.024	-0.024
	(0.016)	(0.018)
State FE	X	X
Year FE	X	X
Industry X Year FE	X	X
Penalty Mean	8.53	8.66
Penalty SD	1.20	1.40
Obs	34,688	34,688
R^2	0.27	0.28

This table replicates Column 1 of Table 1, capping the outcome at the 85th and 95th percentile, respectively. *, **, *** indicate coefficients are significant at the 10%, 5%, and 1% significance level, respectively.

Table B.3: Clean Air Act: Testing for Sample Selection

	Major Source	Privately Owned	Ln(Emissions)	Prior Penalties
Post $\times$ Legal Hiring Reduction	-0.000	0.001*	0.026	-0.999***
	(0.003)	(0.001)	(0.027)	(0.299)
State FE	X	X	X	X
Industry X Year FE	X	X	X	X
Outcome Mean	0.77	0.97	9.86	17.70
Outcome SD	0.42	0.18	3.47	44.57
Obs	41,319	41,319	17,575	41,319

This table uses the regression specified in Equation 2, replacing the outcome with characteristics of penalized facilities.

Table B.4: CAA: Wild Cluster Bootstrap p-values

	(1) Baseline Controls	(2) + Facility Controls	(3) + State Budget Control
Post $\times$ Legal Hiring Reduction	-0.0262	-0.0234	-0.0246
p-value	0.224	0.196	0.198

The table shows wild cluster bootstrap p-values for the main coefficient in Table 1, clustered at the region level.

Table B.5: State Penalty Size (CAA)

	Baseline Controls	+ Facility Controls	+ State Budget Control
Panel A. Blue States			
Post × Regional decrease	-0.021	-0.018	-0.017
	(0.013)	(0.014)	(0.014)
Penalty Mean	8.34	8.34	8.34
Penalty SD	1.31	1.31	1.31
Obs	19,052	19,052	19,020
R^2	0.22	0.23	0.23
Panel B. Red States			
Post × Regional decrease	-0.036	-0.040	-0.040
	(0.030)	(0.028)	(0.026)
Penalty Mean	8.91	8.91	8.91
Penalty SD	1.20	1.20	1.20
Obs	15,503	15,503	15,503
R^2	0.30	0.33	0.33

The table replicates Table 1, separately for states John Kerry won in 2004 (Panel A) and states George W. Bush won in 2004 (Panel B). All columns include state and 3-digit NAICS industry-by-year fixed effects. The post period begins in 2012, and "Regional decrease" gives the treatment intensity, as measured by the regional decrease in federal enforcement actions after the budget cuts. Baseline controls are controls for emissions bins (a proxy for facility size) and state unemployment rate bins. Facility controls are controls for the facility type (indicator for major facility, indicator for non-public facility) and the log of the number of previous penalties observed in the data. The third column adjusts for state-year budgets, as measured by total state expenditures, from the Census Annual Survey of State and Local Government Finances. Standard errors are clustered by state. Data are from 2001-2019. *, **, *** indicate coefficients are significant at the 10%, 5%, and 1% significance level, respectively.

Table B.6: Site Cleanup Descriptives

	All Sites	Regression Sample		
		All	Firm-led	Orphan
Share Ever Remedial Action	0.21	0.41	0.43	0.28
Share Ever Removal Action	0.46	0.52	0.53	0.47
Share Ever Certified	0.45	0.46	0.48	0.36
Median Acreage	3.3	16.3	16.9	11.0
Share 2+ Media Affected	0.57	0.58	0.59	0.53
Count	492	248	212	36

The table shows descriptive statistics for the DTSC cleanup regression sample, as it compares to all "State Response" sites in the data (see Appendix D for more information on the sample selection). The regression sample is limited to sites that cover at least 3 acres. The sum of firm-led sites and orphan sites exceeds the total number of sites in the regression sample because two sites switch from firm-led to orphan during the sample period. The first three rows show the share of sites that had the specified cleanup activities during our sample period (1987-2016). The last row shows the share of sites that have contamination in more than one media (soil, groundwater, etc.).

Table B.7: Projected Costs of Chosen Remedial Alternatives

	Firm-led sites	All sites
1[Rep Adm]_*t*	-0.001	-0.220
	(0.191)	(0.338)
Firm-led		-0.072
		(0.332)
1[Rep Adm]_*t* × Firm-led		0.218
		(0.391)
N Sites	87	100
Mean During Dem Adms	2.59	2.57
SD During Dem Adms	1.01	1.00
R^2	0.35	0.35

The outcome in this table is the rank of the cost (among all considered alternatives) of the remedial action chosen for site cleanup, with higher ranked alternatives being more expensive. The first column includes only firm-led cleanups, and the second column includes all cleanups. The data are at the site level and only include sites over 3 acres where we were able to find information on projected costs. Sites are assigned to years based on when the document with the remedial alternatives was finalized. Standard errors are clustered at the site level. *, **, *** indicate coefficients are significant at the 10%, 5%, and 1% significance level, respectively.

Table B.8: Hazard Ratios from Cox Hazard Model

	All	Firm-led	Orphan	Interaction
1[Rep Adm]_t	0.375***	0.300***	1.492	1.615
	(0.0825)	(0.0722)	(0.946)	(1.020)
Orphan funded				0.397*
				(0.188)
1[Rep Adm]_t $\times$ Firm-led				0.186**
				(0.126)
N Sites	268	232	36	268

The table shows hazard ratios from a Cox hazard model. Standard errors are clustered by site. Data are from 1987-2016 and limited to sites at least 3 acres in size. *, **, *** indicate coefficients are significant at the 10%, 5%, and 1% significance level, respectively.

Table B.9: Probability of Cleanup: Robustness

State Political Environment

	$\mathbf{1}[\text{Rep}]_t$ Only			Interaction
	Both	Firm-led	Orphan	Both
$\mathbf{1}[\text{Rep Adm}]_t$	-0.028***	-0.033***	0.001	0.004
	(0.006)	(0.007)	(0.016)	(0.013)
$\mathbf{1}[\text{Rep Adm}]_t \times$ Firm-led				-0.037***
				(0.014)
Site FE	X	X	X	X
Site Age	X	X	X	X
Mean During Dem Adms	0.05	0.05	0.02	0.05
N Sites	245	209	36	245
N Site-Years	4313	3655	658	4313

The table replicates Table 2, adding a control for the political party of the California governor in each year. *, **, *** indicate coefficients are significant at the 10%, 5%, and 1% significance level, respectively.

Table B.10: Probability of Cleanup: Robustness

Data Decisions

	All	Firm-led	Orphan	All
Panel A. Using Project Start Dates				
1[Rep Adm]_*t*	-0.027***	-0.030***	-0.008	-0.005
	(0.006)	(0.006)	(0.019)	(0.017)
(Years since discovery)/10	0.008**	0.010**	0.001	0.008**
	(0.004)	(0.004)	(0.008)	(0.004)
1[Rep Adm]_*t* × Firm-led				-0.026
				(0.018)
Site FE	X	X	X	X
Mean During Dem Adms	0.04	0.05	0.03	0.04
N Sites	245	209	36	245
R^2	0.08	0.08	0.07	0.08
Panel B. Including Removal Actions				
1[Rep Adm]_*t*	-0.019***	-0.023***	0.015	0.014
	(0.007)	(0.007)	(0.019)	(0.019)
(Years since discovery)/10	-0.043***	-0.043***	-0.041**	-0.043***
	(0.005)	(0.005)	(0.016)	(0.005)
1[Rep Adm]_*t* × Firm-led				-0.038*
				(0.020)
Site FE	X	X	X	X
Mean During Dem Adms	0.06	0.06	0.05	0.06
N Sites	269	233	36	269
R^2	0.08	0.08	0.08	0.08

The table shows robustness analyses for Table 2. Panel A uses manually extracted project start dates instead of project metadata from the database. Panel B includes removal actions as cleanup actions in the outcome, which are typically smaller in scope and more numerous than remedial actions. *, **, *** indicate coefficients are significant at the 10%, 5%, and 1% significance level, respectively.

C Additional analysis: Support for assumption that $\beta \neq 1$

Our model We begin by asking whether EPA appears to have different preferences than the states. In this section, we provide evidence that average state penalty size responds to local economic conditions, but EPA penalty size does not, suggesting that states and EPA have different objective functions. In the model, this would indicate that $\beta \neq 1$, raising the question of whether EPA is too harsh or too lax for the states.

C.1 Empirical Strategy

We present two analyses using Clean Air Act penalties data to compare the role of economic conditions in state versus federal enforcement decisions. In both analyses, we show that state enforcement outcomes are more related to economic factors than are federal enforcement outcomes.

Our first analysis uses penalties issued in 2005-2006 and 2010-2011 to show that after the Great Recession, industries more affected by the recession faced lower average penalties from states but not from EPA. We use the following regression specification, separately for EPA penalties and for state penalties:

$$\begin{aligned} \text{Ln(Penalty Size)}_{j,i,s,t} = {} & \alpha \mathbb{1}(\text{Low-growth industry})_{i,s} \\ & + \beta \mathbb{1}(\text{Low-growth industry})_{i,s} \times D_t + \delta_s I_S + \gamma_t + \rho_i + \epsilon_{j,s,t} \quad \text{(C.1)} \end{aligned}$$

where $\text{Ln(Penalty Size)}_{j,i,s,t}$ is the natural log of the size of penalty j in industry i, state s, and year t. D indicates the penalty was issued after the Great Recession (2010-2011), I_S indicates the penalty was issued by the state, and $(\text{Low-growth industry})_{i,s}$ indicates the facility was in an industry with a bottom-quantile growth rate from 2005 to 2009. We calculate leave-out industry growth rates as

$$\text{industry growth}_{i,s} = \frac{\left[\left(\sum_{s' \neq s} x_{i,s',2009}\right) - x_{i,s,2009}\right] - \left[\left(\sum_{s' \neq s} x_{i,s',2005}\right) - x_{i,s,2005}\right]}{\sum_{s' \neq s} x_{i,s',2005} - x_{i,s,2005}},$$

where x is the total number of establishments or total employment, depending on the specification, in 3-digit NAICS industry i in state s in the specified year. Because we use leave-out growth rates, the ranking of industries differs slightly across states. Because of the limited number of EPA penalties issued, state fixed effects are only included in the regressions using state-issued penalties. In additional specifications, we add controls for facility emissions.

C.2 Results

We document that penalties issued by states appear to be more sensitive to economic conditions than are penalties issued by US EPA. Prior research has shown that air pollution is lower in bad economic times (Chay and Greenstone, 2003; Feng et al., 2015; Finkelstein et al., 2023), suggesting that our estimates of states' lower penalties after the recession may be driven by less severe violations. However, emphasizing the *difference* between state and EPA enforcement in good versus bad times belies this concern.[45]

First, we compare the relationship between state unemployment rates (lagged by three months) and penalty size. Figure A.1 shows this relationship for state-issued penalties and for EPA-issued penalties. There is a distinct downward-sloping relationship for states, and no such relationship for EPA-issued penalties. Of course, this could be driven by lower state enforcement capacity during bad economic times.

We turn to our main analysis of whether penalties responded to the economic shock of the Great Recession. Appendix Figure A.4A shows the distribution of industries represented in the penalty data. More than half of penalties are issued to manufacturing facilities, which make up almost the entirety (over 98%) of industries in the bottom quartile of industry growth between 2005 and 2009.

Table C.1 shows the results of estimating Equation C.1. We see that in industries hit hardest by the Great Recession, state penalties decrease after the Great Recession, while EPA penalties do not. This is true whether industry decline is measured using the number of establishments or total employment, although the effect is larger using the number of

[45]We also note that a plurality (if not majority) of CAA violations that result in formal enforcement actions are not necessarily for excess emissions but rather for procedural noncompliance such as inappropriate equipment and processes. Violations of, for example, abatement equipment requirements, should not be less likely in bad economic times.

establishments. Using establishment counts as our measure of industry growth, penalties are 25% lower in hard-hit industries relative to other industries and relative to penalties before the Great Recession. The results are also robust to including state-by-year fixed effects instead of state and year fixed effects and to using midpoints of employment ranges where employment data is suppressed. Altogether, it appears that states treat suffering industries differently from other industries, but EPA does not; however, we are not able to reject that the coefficients are the same.[46]

As we noted, the comparison of state penalties to EPA penalties should address most concerns that systematically varying emissions are driving our results. However, we do include two additional analyses related to this concern. First, Appendix Table C.2 uses the sample of all TRI-reporting major and synthetic minor facilities (not only those with penalties) and estimates effects of the recession on reported emissions. For this analysis, we are able to use facility fixed effects. Our employment measure of industry growth is significant related to emissions, but the magnitude of the effect is small (less than 5% of a standard deviation). In Appendix Table C.3, we replicate Table C.1 including a control for facility emissions reported in the TRI. While reported emissions are strongly related to penalty size, this control has minimal effect on the coefficient of interest, suggesting that changes in emissions are not driving the results.

Together, these results show that EPA's penalties do not respond in the same way as state penalties do to economic conditions. Under the assumption that penalties do more damage to firm profits or employment in bad economic times, this suggests EPA down-weights economic harms of enforcement relative to state preferences. In this case, EPA might prefer higher penalties than states do. Is EPA *too* harsh for the states, or could a harsher gorilla help the states more? To answer this question, we next turn to implementing the model's empirical test in two contexts.

[46]One concern with this analysis is that selection into state versus EPA enforcement could differ in suffering industries. If states are less likely to handle specifically the worst violations in suffering industries, we might see this pattern. In this case, states may not be concerned about economic harms from enforcement but may face political constraints on their enforcement. While we cannot test for selection specifically on violation severity, we see that facilities in low-growth industries are not significantly more likely to face EPA enforcement (versus state enforcement) after the Great Recession relative to other industries.

C.3 Tables

Table C.1: State vs. EPA Penalty Patterns

	Establishments		Employment	
	Ln(State penalty)	Ln(EPA penalty)	Ln(State Penalty)	Ln(EPA penalty)
1[Low-growth] × 1[2010-2013]	-0.25***	0.10	-0.15*	0.09
	(0.08)	(0.39)	(0.08)	(0.36)
Issuer FE	X	X	X	X
Year FE	X	X	X	X
Industry FE	X	X	X	X
Mean Ln(Penalty)	8.8	10	8.8	10
SD Ln(Penalty)	1.3	1.8	1.3	1.8
Obs	4,754	379	4,754	379
R^2	0.22	0.09	0.22	0.09

The outcome uses data from the two years before (2006-2007) and the two years after (2010-2011) the Great Recession. The specification is given in Equation C.1. The first set of columns measures industry growth using the number of establishments, and the second set of columns uses total employment. Within each set, the first column shows that penalties issued by states are lower after the Great Recession in industries hit harder by the Great Recession, relative to other industries and relative to penalties issued before the Great Recession, and the second column shows that penalties issued by EPA do not exhibit this pattern.

1[Low-growth] indicates the facility is in an industry with bottom-quartile growth from 2005 to 2009, as measured by the percent change in the number of establishments (total employment) nationwide, excluding establishments (employment) in the state issuing the penalty. This variable is calculated using data from U.S. Census's County Business Patterns. 1[2010-2011] is an indicator for the penalty being issued in 2010-2011. *, *** indicates the coefficient is significant at the 10% and 1% significance levels, respectively.

Table C.2: Great Recession: Log Emissions as Outcome

	Establishments	Employment
1[Low-growth] × 1[2010-2011]	0.03	-0.14***
	(0.04)	(0.03)
Facility FE	X	X
Year FE	X	X
Mean Ln Emissions	8.60	8.60
SD Ln Emissions	3.11	3.11
Obs	36,954	36,954
R^2	0.93	0.93

The table uses the sample of all TRI-reporting major and synthetic minor facilities in 2006-2007 and 2010-2011. We estimate Equation C.1, replacing the outcome with the log of reported emissions, and replacing state and industry fixed effects with facility fixed effects. Standard errors clustered at the state level are in parentheses. *** indicates statistical significance at the 1% significance level.

Table C.3: Great Recession: Emissions Controls

	Establishments		Employment	
	Ln(State penalty)	Ln(EPA penalty)	Ln(State Penalty)	Ln(EPA penalty)
1[Low-growth] × 1[2010-2011]	-0.26***	0.12	-0.14*	0.12
	(0.08)	(0.53)	(0.08)	(0.34)
Ln(TRI Air Emissions)	0.06***	0.17**	0.06***	0.17**
	(0.01)	(0.06)	(0.01)	(0.06)
State FE	X		X	
Year FE	X	X	X	X
Industry FE	X	X	X	X
Obs	4,754	379	4,754	379
R^2	.24	.14	.24	.14

The table replicates Table C.1, adding controls for TRI-reported emissions. Standard errors clustered at the state level are in parentheses. *, **, *** indicate statistical significance at the 10%, 5%, and 1% significance level, respectively.

D Data Appendix

D.1 Superfund

The data we use from the California Department of Toxic Substances Control (DTSC) come largely from their online database, called EnviroStor. EnviroStor is used internally to track cleanup projects, and much of it is accessible online so that interested members of the public can learn more about hazardous waste sites in California.

D.1.1 EnviroStor Sample

EnviroStor does not contain every site in Califronia with hazardous substance contamination. Sites that are not under DTSC jurisdiction are not included. For example, most petroleum contamination (which is often the result of leaking underground storage tanks) falls under the jurisdiction of the California State Water Resources Control Board.

At the same time, EnviroStor contains more sites than are relevant to this project. We apply sample criteria which exclude the the following:

Site types. The DTSC runs several programs which evaluate sites for potential contamination—these are largely military bases and sites proposed for acquisition or development by school districts. We limit the sample to "State Response" sites. This excludes "Evaluation" sites, which were largely historical or current programs that assess public property (schools and military sites) to check for contamination, as well as sites where contamination is (or was at one point) suspected but not confirmed. This is a large share of sites in the database (41.7%), but a much smaller share of documented activities (11.4%). We also exclude sites under the "Cal-Mortgage" program, a loan program for non-profit and public entities which requires environmental review (1% of all sites, and 0.1% of all activities).

The other large class of cleanup projects that are excluded from the sample are voluntary cleanups.

Referred sites. We exclude sites that were referred to other agencies (such as regional water boards) or other California environmental programs (such as the Resource Conservation and Recovery Act, or RCRA), since DTSC does not generally track activities at these sites. This criterion excludes an additional 21% of sites in EnviroStor.

No Action Required. Finally, we exclude sites with an EnviroStor "status" of "No Action Required" or "No Evidence of Release." These are sites where contamination was not found in levels high enough to require cleanup. While 9% of the entire EnviroStor sample has this status, the vast majority are covered by the site types we exclude (required evaluations for public entities); this criterion only excludes an addition 1% of the sample beyond the above criteria.

D.1.2 Remedial Action Dates

We find remedial action dates from several sources. End dates come from the dates of "Remedial Action Completion Reports" in EnviroStor, which correspond to the date that the DTSC approved the remedial action completion.

A robustness analysis uses *start* dates as our outcome. These require examining the contents of the documents uploaded into EnviroStor. The source we prioritize for remedial action (RA) start dates is certification forms. Certification forms are uploaded after all required remediation activities—including, occasionally, monitoring for a period of time—are complete at a site. These forms are standardized and include a field for project start and end dates. There are two reasons we are not be able to get RA start dates for all projects from certification forms. The most common reason is that the site is not yet certified. The second reason is that the certification form does not specify the RA start date — it might instead give the date the DTSC ordered the cleanup, or the date the site assessment began. We are able to get start dates from certification forms for 59 of 148 remedial actions in the sample.

For the remainder of remedial actions during our sample period, we next turn to additional reports uploaded to EnviroStor. Where remedial action completion reports follow design and implementation plans (which describe the plan for the remedial action implementation) within less than a year, we consider the design and implementation plan date the start date. Otherwise, we turn to the Remedial Action Completion Report itself, a technical document often hundreds of pages in length. This report often includes the dates of the project implementation, usually in its introduction. If not in the introduction, dates can sometimes be found in dated documents in appendices—for example, in daily field reports or in date-stamped photographs of project implementation.

Where we are unable to find dates in certification forms or in technical reports, we use the RA Completion Report date. We use this for 11 of the 148 remedial actions in the data.

D.1.3 Judgement calls

We note two additional judgement calls in these data. First, two site (The "Wickes Forest Industries" sites and the "McNamara and Peepe Lumber Mill" site) began as firm-led sites and became orphan sites later, when the firms responsible for the contamination declared bankruptcy. Instead of using the funding source noted in EnviroStor, which gives the current funding source for the project, we consider these a firm-led sites until the first state work order (an indication of a response action which is state-funded) was issued.

Second, some activities uploaded into EnviroStor as Remedial Action Completion Reports use "removal action" terminology in the reports. (Activities are categorized based on their expected costs, where remedial actions are more expensive and involved than removal actions.) It's unclear whether these activities should be considered remedial actions or removal actions. We consider these remedial actions, deferring to the judgement of the project manager who uploaded the document in EnviroStor, but we show robustness to including all remedial *and* removal actions in our results.

D.2 Clean Air Act

D.2.1 Penalties

We adjust penalties to 2010 dollars using the BLS Consumer Price Index.

Some facilities have multiple enforcement actions on a given settlement date. We collapse the data so that a facility has at most one observation on any given date. When a facility has multiple enforcement actions on a given date, if the enforcement actions are associated with different penalty amounts, we sum the penalties within the date. If, however, the enforcement actions are associated with the same penalty amount (less than

5% of all facility-dates), we consider this a duplicate entry, and we keep only the first enforcement action.

Sometimes, EPA and a state will bring a case jointly. In the model, we consider joint state-EPA cases to be instances when the state "rejects" the penalty and reports it to the EPA. To exclude joint state-EPA cases in our data, we exclude any state enforcement actions that happen during the same month as an EPA enforcment action.

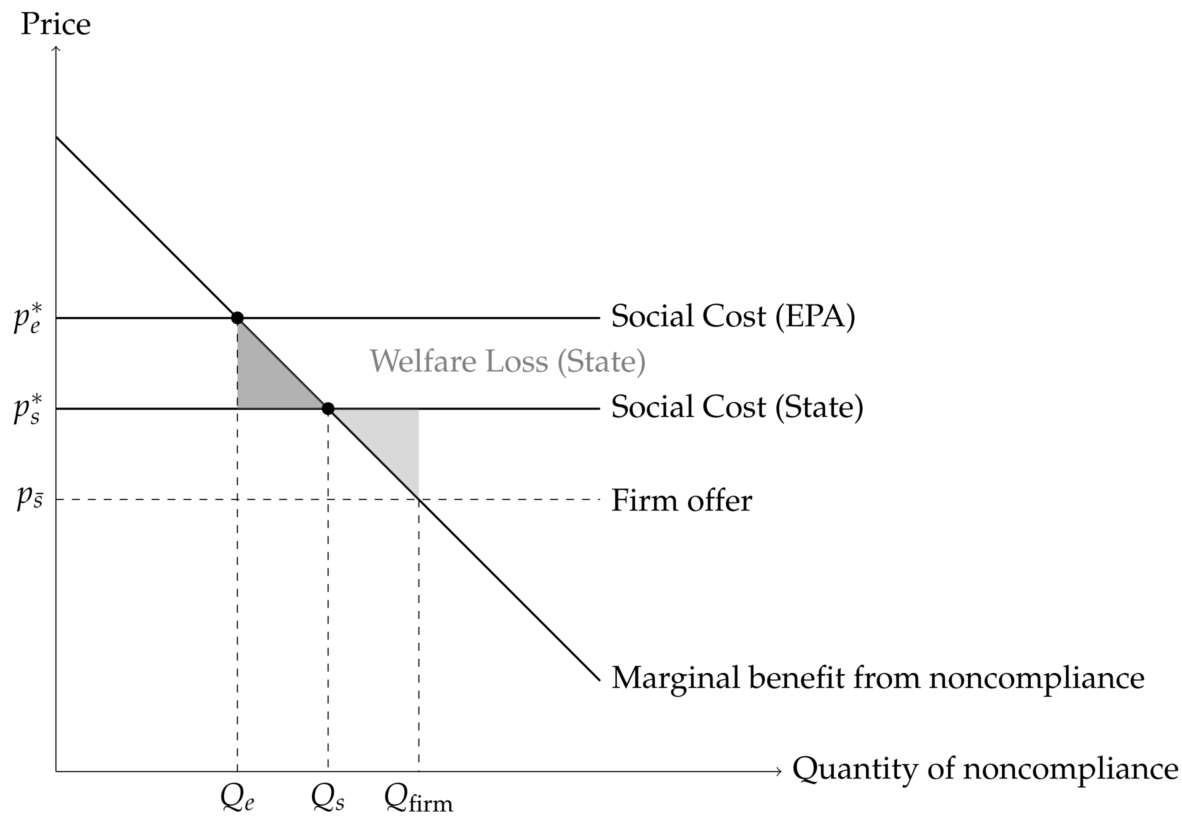
D.2.2 Toxic Release Inventory

We use EPA's "Basic Plus" files, and specifically, the "total air emissions" entry. For facilities ever observed in the TRI, we impute emissions in years when the facility is missing from the data using the last-reported year. We do not replace 0's (which EPA uses to indicate missing values as well as true 0s) in this process. In our main specification, we bin log TRI emissions into deciles, add a group for missing or zero TRI emissions, and include fixed effects for the emissions deciles and the missing TRI data control.

E Model Appendix

E.1 Recasting the model in terms of deadweight loss

We model the state and the EPA as having concave utility over penalty size because it allows us to easily visualize the sufficient statistics result. However, we can recast our welfare problem in a more traditional supply and demand (for noncompliance) diagram.[47] Demand is given by firms' compliance costs: higher compliance costs lead to higher benefits of noncompliance. Supply (marginal cost) represents the marginal social cost of noncompliance. The state and the EPA disagree on the marginal social cost of noncompliance: EPA considers the social cost of noncompliance to be higher than the state does.



The top shaded triangle represents the deadweight loss the state faces under EPA's penalty (p_e^*). The lost trades (from Q_e to Q_s) are firms with higher compliance costs than

[47] The diagram represents solely the marginal costs and benefits of *noncompliance*, independent of output, i.e., assuming firms can change noncompliance without changing output.

the state's social cost of noncompliance, which are now paying the compliance cost to avoid EPA's penalty.

The firm can offer a lower penalty ($p_{\bar{s}}$). The welfare loss from this penalty (the lower gray triangle, representing firms which are now noncompliant despite benefiting less than the social cost) gives the state equivalent welfare loss to EPA's penalty, making the state indifferent between the two.

One important difference between the model in the paper and this model is that in this model, the magnitude of the welfare effects require less information: we need only know the optimal penalties of the regulators and the distribution of compliance costs.

E.2 Deterrence $b(p)$

We model the state and the EPA as receiving benefits from issuing penalties, with the idea that these benefits come from deterring future violations. However, in the game as written in Section 3, there are no explicit deterrence effects: the firm's decision to violate occurs at the beginning of the game, before the state and EPA make their decisions. In this section, we show how the regulators' benefits from penalties $b(p)$ can be microfounded by deterrence effects without meaningfully affecting our main results.

To capture deterrence effects, we add a final period to the game where the firm decides whether to violate for a second time. We have the regulators incur the economic harms of enforcement when they issue a penalty, but do not allow them to receive any benefits of enforcement until this final period. In the final period, they receive utility for any violations avoided.

The sequence of actions taken about a given violation v is now:

0. The firm draws compliance cost $\eta_v \sim F$ and decides whether to commit a violation. It commits a violation if η_v exceeds the equilibrium penalty. If it does, then...

1. The firm makes a penalty offer p_{s_v} to the state.

2. EPA pays $c \cdot k$ to issue sanction threat k.

3. The state either accepts the firm's offer p_{s_v} (choosing $I_S = 1$), or it rejects the offer and sends the case to EPA (choosing $I_S = 0$). If the state accepts the firm's offer, both regulators incur the costs of their economic harm of enforcement from p_{s_v}

4. If the state sends the case to EPA, EPA chooses p_e^* and issues penalty $p_e = \sigma(N)p_e^*$. Both regulators incur the economic harms of enforcement from p_e.

5. **The firm draws a new (independent) compliance cost and violates a second time iff this compliance cost η_w exceeds the penalty that has been issued,** $I_S p_{s_v} + (1 - I_S)p_e$**. The regulators realize their environmental benefits.**

The regulators get utility from averting the second violation. In the final period, the probability that the penalty issued prevents the second violation is $F(p_{s_v})$ if the state issued the penalty, and $F(p_e)$ if the EPA issued the penalty. By the properties of cumulative distribution functions, $F(p) \geq F(p')$ iff $p > p'$, so higher penalties have a higher likelihood of deterring the second violation. Thus, the regulators get higher benefits from higher penalties. That this benefit function would be concave (an assumption in Section 3) could be microfounded with a specific distribution of compliance costs. Below, we introduce variation in the severity of averting violations, which provides alternative assumptions that give a concave benefit function.

The sufficient statistics test. This deterrence model is effectively a rewriting of the main model; in equilibrium the state issues all penalties, and the proofs of the propositions go through.

One-shot vs. repeated game. We do not use a repeated game to model deterrence effects. There would be several differences in a model with a repeated game.[48] Most importantly, EPA would internalize the effect of its penalty on equilibrium state penalties, and so would choose the p_e that maximizes state penalties. In practice, we do not believe that EPA always chooses the optimal p_e for the states.

[48]For example, the regulators would need to incur a cost from penalizing the second violation, and, anticipating this, might prefer a higher penalty for the first violation to accomplish more deterrence and avoid those harms. This has flavors of Becker (1968), but note that our model has complete enforcement, so optimal penalties in our model differ from optimal penalties in Becker's.

E.2.1 Heterogeneous severity of violations

One potential implication of deterrence effects is that they could change the nature of the violations that do occur. In this section, we formalize this concern and show that under reasonable assumptions, this does not threaten the sufficient statistics result (Proposition 2).

There is currently no notion of severity of violations in our model. Each violation is equally beneficial to prevent, or put another way, deterrence effects as we've modeled them above do not affect the regulators' desired penalties for a violation that still does occur. We now relax this assumption.

Let violation v have severity μ_v, which corresponds to the environmental harm the violation causes. η_v and μ_v are jointly distributed $G(\eta_v, \mu_v)$. We assume that $\text{Corr}(\eta_v, \mu_v) > 0$, so that more severe violations are costlier for the firm to avoid. Now penalties affect not only the mass of violations, but also the composition. If a stronger EPA results in higher equilibrium state penalties, only more severe violations will occur (and vice versa). If avoiding more severe violations is more beneficial for regulators (i.e., $\mu_v > \mu_{v'} \implies b(p;\mu_v) > b(p;\mu_{v'})$), then the remaining violations will face higher penalties.

The sufficient statistics test. Under these assumptions, the sufficient statistics result remains valid, since this selection effect operates in the same direction as the sufficient statistics test. If a stronger EPA provides a better outside option for the state, the state will be able to negotiate a higher penalty for a given violation, which in turn means the violations that occur in equilibrium are more severe, and vice versa. It may be that the state prefers a stronger EPA only in the absence of this selection effect, since the selection effect will also increase EPA's desired penalty. However, this does not affect the relationship between EPA's desired penalty and the state's realized penalties.

E.3 Proofs for Section 3

Equilibrium

First, note that when the state has a comparative advantage in enforcement ($\zeta_v >$

0), for any EPA penalty p_e, the state will always accept some offer less than p_e (since with $\zeta_v > 0$, $u_S(p_e) + \zeta_v > u_S(p_e)$). Thus, being a cost-minimizer, the firm prefers state enforcement to EPA enforcement for any case with $\zeta_v > 0$. The firm will make the lowest offer possible that the state will accept (i.e., $I_S = 1$ in equilibrium).

Note also that the optimal sanctions threat for EPA to make, should it make any, is $k^* = u_S(p_{\bar{s}_v}) + \zeta_v - u_S(p_e)$. This is the threat that will convince the state to reject the case. If EPA offers less than this, the state will not reject the case, and the EPA will see no benefit to threatening sanctions at all.[49] Since sanction threats are costly, EPA will not choose a higher k than necessary, either. Thus, EPA chooses between $k = 0$ and $k^* = u_S(p_{\bar{s}_v}) + \zeta_v - u_S(p_e)$.

Imagine the firm offered less than $p_{\bar{s}_v} = \min\{p_{\tilde{s}_v}\}$, where $p_{\tilde{s}_v}$ is defined in Equation 1. A lower p_{s_v} reduces k^*, and increases the EPA's benefit of handling the case ($u_E(p_e) - u_E(p_{\bar{s}_v}) - \zeta_v$), such that $u_E(p_e) - u_E(p_{\bar{s}_v}) - \zeta_v > c * k^*$. In this case, EPA will threaten sanctions, the state will reject the case (by the definition of k^*), and the firm will face the EPA's penalty. This is costly for the firm, so the firm will not offer less than $p_{\bar{s}_v} = \min\{p_{\tilde{s}_v}\}$. The EPA will choose $k = 0$.

Proof of Proposition 1

Figure 2 builds intuition in a case when the EPA cannot sanction the states.

For an algebraic proof: First, note that when EPA cannot sanction the state, the firm will offer the minimum value that makes the state indifferent between issuing the penalty itself (i.e., accepting the firm's offer) and allowing the EPA to issue the penalty: $u_S(p_{\tilde{s}_v}) + \zeta_v = u_S(p_e)$.

From here, we suppress v subscripts, and we prove Proposition 1 in cases.

Case 1: assume that $p_e \leq p_s^*$. In this case, $\lim_{h \to 0^-} \frac{u_S(p_e+h) - u_S(p_e)}{h} > 0$. By continuity of u_S, $\exists\, \epsilon < 0$ such that $u_S(p_e + \epsilon) = u_S(p_e) - \zeta$. (There may also be an $\epsilon > 0$ such that the statement holds, but the cost-minimizing firm will prefer the offer with $\epsilon < 0$). Thus we have $p_s = p_e + \epsilon < p_e \leq p_s^*$.

Case 2: assume that $p_e > p_s^*$. By our assumptions on $b(p)$ and $\tau(p)$, there exists an

[49] That EPA sees no benefit of sanctions unless the sanctions threat convinces the state to reject the case is a feature of the timing. If EPA chose its sanction threat before the firm made its offer (or if this was a repeated game), EPA would face a continuous problem of optimal sanction choice.

$x < p_s^*$ such that $u(x) = u(p_e)$. From here, we are in Case 1.

Proof of Proposition 2

Suppose the EPA cannot sanction, so that $p_{\bar{s}}$ satisfies $u_S(p_{\bar{s}}) + \zeta = u_E(p_e)$. Then Proposition 2 follows easily from Proposition 1: we use that, by the assumptions on b and τ, u_S is increasing in penalty size p_s if and only if $p_s < p_s^*$. From Proposition 1, we know that $p_s < p_s^*$. Thus, if EPA strength increases firm offers, it also increases state welfare.

Proof of Proposition 3

Suppose that $c > \frac{u_E(p_e)-\zeta-u_E(p_s^*)}{u_S(p_s^*)+\zeta-u_S(p_e)}$ and $c < \infty$. Then we have that $c[u_S(p_s^*) + \zeta - u_S(p_e)] > u_E(e) - \zeta - u_E(p_s^*)$, i.e., that $ck > u_E(p_e) - \zeta - u_E(p_s^*)$. Thus, if the firm offers p_s^*, EPA's cost of imposing sufficient sanctions on the state that the state would reject the firm's offer outweighs EPA's benefit of getting the case. Because EPA's benefit of getting the case is strictly decreasing in firm offers (as long as $p_s < p_e$), there is no $p_s > p_s^*$ such that EPA will want to sanction the state, and so there is no reason for the firm to make an offer that exceeds p_s^*.

By a similar logic, if $c < \frac{u_E(p_e)-\zeta-u_E(p_s^*)}{u_S(p_s^*)+\zeta-u_S(p_e)}$, EPA would sanction the state if it accepted p_s^*, and so the state would reject an offer of p_s^*, subjecting the firm to EPA's (higher) penalty. Thus the firm will offer above p_s^*.

Proof of Proposition 4

Proof. First, we note that it is sufficient to prove this for $\sigma(N) = 1$.

Next, we show that firm offers are lower than the state's preferred penalty when $\beta = 1$: Let $\beta = 1$, so that $p_e = p_s^*$. Write $u = u_S = u_E$. The firm's offer is such that $(1 + c)[u(p_s) - u(p_s^*) + \zeta] = 0$. The p_s that satisfies this condition is lower than p_s^*.

Finally, we show that when $\beta = 1$, firm offers are increasing in $(1 - \beta)$. We fully differentiate Equation 1 and rearrange terms to get

$$\frac{dp_{\bar{s}}}{dp_e} = \frac{\tau(p_e) - \tau(\bar{p}_s) - \frac{dp_e}{d(1-\beta)}[u'_S(p_e) + \frac{1}{c}u'_E(p_e)]}{\frac{dp_e}{d(1-\beta)}[u'_S(p_{\bar{s}}) + \frac{1}{c}u'_E(p_{\bar{s}})]}$$

When $\beta = 1$ and $\sigma(N) = 1$, we have that $u'_S(p_e) = u'_E(p_e) = 0$. We are left with

$$\frac{dp_{\bar{s}}}{dp_e}\Big|_{\beta=1} = \frac{\tau(p_e) - \tau(p_{\bar{s}})}{\frac{dp_e}{d(1-\beta)}[u'_S(p_{\bar{s}}) + \frac{1}{c}u'_E(p_{\bar{s}})]}$$

Proposition 1 gives us that $p_{\bar{s}} < p_e$, so that $\tau(p_e) - \tau(p_s) > 0$ and $u'_S(p_{\bar{s}}) = u'_E(p_{\bar{s}}) > 0$. Thus $\frac{dp_{\bar{s}}}{dp_e}\big|_{\beta=1} > 0$.

□

E.4 Model Extensions

States with heterogeneous preferences

Consider equilibrium firm offers in a case where EPA cannot sanction the state.

Equilibrium firm offers $p_{\bar{s}}$ satisfy

$$u_S(p_{\bar{s}}) = u_S(p_e) - \zeta_v, \tag{E.1}$$

where u_S is the state's utility from the penalty, and p_e is the penalty the EPA would issue should it get the case.

When EPA's penalty p_e changes, the effect on equilibrium state penalties is

$$\frac{dp_{\bar{s}}}{dp_e} = \frac{u'_S(p_e)}{u'_S(p_{\bar{s}})} \tag{E.2}$$

To introduce heterogeneity by state, index states by i. States vary in their preferred penalty $p^*_{s_i}$. The average treatment effect over a set of states i, which we write as $\gamma(\{i\})$, is:

$$\gamma(\{i\}) = E[Y_i(1) - Y_i(0)] \tag{E.3}$$

$$= E[Y_i(1) - Y_i(0)|p^*_{s_i} > p_e]\Pr(p^*_{s_i} > p_e) + E[Y_i(1) - Y_i(0)|p^*_{s_i} \leq p_e]\Pr(p^*_{s_i} \leq p_e), \tag{E.4}$$

where the second line uses the law of total expectation. Consider only a marginal change dp_e, such that $Y_i(1) - Y_i(0) = \frac{dp_{\bar{s}}}{dp_e}$.

Now we have

$$\gamma(\{i\}) = \underbrace{E\left[\frac{u'_{S_i}(p_e)}{u'_{S_i}(p_{\bar{s}_i})}|p^*_{s_i} > p_e\right]}_{(+)}\Pr(p^*_{s_i} > p_e) + \underbrace{E\left[\frac{u'_{S_i}(p_e)}{u'_{S_i}(p_{\bar{s}_i})}|p^*_{s_i} \leq p_e\right]}_{(-)}\Pr(p^*_{s_i} \leq p_e), \quad \text{(E.5)}$$

where the brackets indicate the terms' signs under the model.

Proposition E.1. *If EPA is too strong or too lenient for all states, the sufficient statistics result in Proposition 2 holds, even when states differ in their optimal penalty.*

Proof. This can be easily seen in cases, by setting $Pr(p^*_{s_i} > p_e) = 0$ and then $Pr(p^*_{s_i} \leq p_e) = 0$ in the above expression. Then, $\gamma(\{i\})$ is positive if and only if all states benefit from an increase in EPA strength ($\Pr(p^*_{s_i} > p_e) = 1$). □

However,

Proposition E.2. *If EPA is too strong for some states and too lenient for others, positive γ does not necessarily imply higher welfare under a utilitarian welfare function. That is, the sufficient statistics result does not necessarily hold.*

Proof. We prove this by contradiction. Consider a model with two states, i and j, where EPA is more lenient than i would like and more harsh than j would like (Assumption E.1 below), and furthermore, i benefits more than j is hurt from a stronger EPA (Assumption E.2).

Assumption E.1. $p^*_{s_j} < p_e < p^*_{s_i}$

Assumption E.2. $|u'_i(p_e)| > |u'_j(p_e)|$

Because on average, the states benefit from a stronger EPA, the sufficient statistics result would imply that $\gamma(i, j) > 0$.

It suffices to show a single counterexample. The simplest counterexample sets $\zeta_i = \zeta_j = 0$. With this, $p^*_{s_i} > p_e \implies \min\{p_{\bar{s}_i}|u(p_{\bar{s}_i}) = u(p_e)\} = p_e$, and $p^*_{s_j} > p_e \implies \min\{p_{\bar{s}_j}|u(p_{\bar{s}_j}) = u(p_e)\} < p_e$. Thus, $p_{\bar{s}_i} > p_{\bar{s}_j}$.

By concavity of u, then, we have that

$$u_i'(p_{s_i}) < u_j'(p_{s_j}). \tag{E.6}$$

And by firm optimization, both are positive, as the firm will not offer a penalty on the decreasing portion of u when one on the increasing portion will do. Plugging Equation E.1 into Assumption E.2 (and setting $\zeta_i = \zeta_j = 0$), $|u_i'(p_{s_i})\frac{dp_{s_i}}{dp_e}| > |u_j'(p_{s_j})\frac{dp_{s_j}}{dp_e}|$. Together with Equation E.6, this implies that

$$|\frac{dp_{s_i}}{dp_e}| < |\frac{dp_{s_j}}{dp_e}| \tag{E.7}$$

Note also that under Assumption E.1 the sufficient statistics result implies that $\frac{dp_{s_j}}{dp_e} < 0 < \frac{dp_{s_i}}{dp_e}$.

Thus,

$$\gamma(i,j) = 0.5 * \frac{dp_{s_i}}{dp_e} + 0.5 * \frac{dp_{s_j}}{dp_e} < 0. \tag{E.8}$$

The last inequality is a contradiction. □