Protective Fines or Credible Deterrence?

Penalty Design and Environmental Enforcement Effectiveness

Shusen Qi (Xiamen University, shusenqi@xmu.edu.cn)

Zhiyan Zhan (Xiamen University, zhanzhiyan@stu.xmu.edu.cn)

Abstract: This paper examines whether environmental administrative penalties generate general deterrence by updating peers' beliefs about enforcement. Using a monthly panel of Chinese prefecture-level cities linking publicly disclosed penalty decisions to ambient SO_2, we show that aggregate enforcement intensity is a poor proxy for deterrence. Total penalties have no robust effect on air quality, but disaggregating by instrument reveals sharp heterogeneity: only combined penalties that couple monetary fines with production restrictions reduce SO_2, consistent with complementarity between financial and operational sanctions and with credible escalation. Deterrence is stronger where enforcement is more credible, under stronger official career incentives, lower resource dependence, and stronger environmental adjudication, and is short-lived with limited cross-jurisdiction spillovers.

Keywords: environmental enforcement; administrative penalties; general deterrence; air pollution; policy credibility

JEL codes: Q53; Q58; K32; H11

Acknowledgements: The authors have no relevant or material financial relationships, employment, consulting, ownership interests, or other potential conflicts of interest to disclose.

1. Introduction

Environmental regulation is often evaluated by how frequently authorities inspect firms or how many violations they sanction. Yet in practice, a more fundamental question is whether sanctions are economically capable of deterring pollution at all. Across countries, policymakers repeatedly worry that penalties can be too small, or too weakly implemented, to change firms' behavior, effectively converting environmental rules into priced permission to pollute.[1] This concern is explicit in international guidance that recommends sanctions should, at a minimum, remove the economic benefit of non-compliance, precisely to avoid turning penalties into a predictable operating cost.[2] It also underlies the long-standing deterrence perspective in economics: compliance depends on the expected penalty, i.e., punishment severity multiplied by the probability of effective enforcement (Becker, 1968; Polinsky and Shavell, 2000).

This paper studies a particularly salient manifestation of under-deterrence: cases in which penalties are so low that they do not warn or punish and may instead encourage illegal emissions. In China, this phenomenon is often described as protective fines, situations where sanctions are perceived to be set deliberately below a meaningful deterrence threshold. The key point, however, is not that this pathology is uniquely Chinese; rather, China makes the problem unusually visible. Under decentralized enforcement, substantial local discretion, and pervasive uncertainty about implementation, firms may rationally treat low penalties as a signal that non-compliance is tolerable, thereby undermining the credibility of environmental regulation.

Motivated by this puzzle, we ask: when do environmental penalties generate broad deterrence, and when do they fail? The focus is on general deterrence, the response of non-sanctioned firms who learn from observed punishments, rather than only on the specific deterrence of penalized firms. General deterrence can be a first-order margin for welfare when peer learning is strong: punishment of

[1] Using the universe of US civil Clean Air Act violations by stationary sources, Atkinson (2023) finds that in a substantial share of cases, violations remain profitable even after penalties are paid, implying systematic under-deterrence for large violations. In parallel, European policy debates have increasingly emphasized that sanctions must be effective, proportionate, and dissuasive, culminating in Directive (EU) 2024/1203 on environmental crime, which strengthens the framework for criminal enforcement across member states.

[2] See https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0494.

a small number of firms can induce widespread compliance if it credibly updates beliefs about enforcement risk (Shimshack and Ward, 2008; Gray and Shimshack, 2011; Alm et al., 2009). Yet empirically, evidence remains mixed, in part because enforcement intensity is often measured as a single aggregate that treats penalties as interchangeable.

A central argument of this paper is that penalties are not interchangeable. Different enforcement tools alter firms' incentives through distinct margins, and these margins may be weak on their own but powerful in combination. In particular, monetary fines and production restrictions each have intuitive deterrent potential, yet each can plausibly fail in isolation. Fines may not deter if they are systematically below the gains from violating or if firms expect substantial discretion in penalty calculation and collection, conditions under which "paying the fine" becomes an equilibrium compliance strategy. Production restrictions, in turn, may fail to deter if firms expect them to be short-lived, selectively implemented, or partially evaded through operational substitution, especially in settings where enforcement credibility varies across jurisdictions. Crucially, a combined penalty, a monetary fine coupled with a binding production restriction, can function as a credible escalation that simultaneously imposes immediate operational losses and signals tighter future scrutiny. This escalation logic is consistent with recent evidence that dynamic enforcement regimes can generate large compliance gains when sanctions convey credible follow-through rather than symbolic punishment (Blundell et al., 2020).

China offers a high-value setting to test these ideas for two reasons. First, administrative penalties remain a dominant enforcement instrument, and enforcement credibility plausibly varies with local political incentives and institutional capacity. Second, China's penalty system includes a key feature that helps identification: the timeline from violation detection to public disclosure of penalty decisions is often long and irregular. As a result, disclosure constitutes an observable informational shock to local peers' beliefs about what enforcement "really looks like," helping us isolate deterrence operating through belief updating rather than contemporaneous inspections.

Empirically, we assemble a prefecture-by-month panel from 2015-2022 linking administrative penalty disclosures to high-frequency measures of ambient SO_2 concentration. Our first main finding is that treating penalties as homogeneous yields little robust evidence of pollution reduction. This null

result is informative: enforcement volume is not a sufficient statistic for deterrence when sanctions differ sharply in bindingness. Once we disaggregate penalties by type, a clear pattern emerges: only combined penalties, those integrating monetary fines with production restrictions, predict meaningful reductions in SO_2, whereas fines alone or restrictions alone show no consistent effects. We interpret this as evidence that deterrence requires both a sanction that is economically binding and a credible signal of escalation that shifts peers' expectations about future enforcement.

Our second main finding concerns institutional credibility. The deterrent effect of combined penalties is stronger in jurisdictions where enforcement is more likely to be perceived as credible and sustained, shaped by political incentives, economic structure, and supporting institutions. The results thus link deterrence to the political economy of decentralized regulation, where local governments may face competing objectives and where institutional capacity affects whether penalties are perceived as real constraints or negotiable outcomes.

Third, we document that deterrence is temporally and spatially bounded. The effect of combined penalties decays within months, consistent with short-lived compliance responses to heightened scrutiny, and we find limited spillovers across neighboring jurisdictions. This pattern is consistent with the idea that externalities travel across space, but enforcement authority, and hence firms' perceived enforcement risk, remains largely jurisdiction-bound.

These findings contribute to the broader enforcement literature in three ways. First, we provide new evidence on general deterrence by exploiting a setting in which penalties become salient as public informational shocks, complementing prior work showing that enforcement can spill over through reputation and learning rather than only through sanctions on the punished firm (Shimshack and Ward, 2005, 2008; Gray and Shimshack, 2011). Second, we demonstrate that policy evaluation must move beyond more or less enforcement to which enforcement instrument, consistent with theory emphasizing the joint roles of detection and sanction design in generating deterrence (Polinsky and Shavell, 2000), and with evidence that enforcement regimes with credible escalation can generate markedly larger welfare gains than static, fine-only approaches (Blundell et al., 2020; Gray and Shimshack, 2011). Third, we identify institutional conditions under which enforcement is more effective, helping reconcile mixed empirical findings by showing that deterrence hinges on both

penalty design and credibility, a theme central to modern environmental economics and to empirical work linking regulatory pressure and institutions to observed pollution outcomes (Shimshack and Ward, 2005; Greenstone and Jack, 2015; Shapiro and Walker, 2018).

The results carry direct implications for environmental governance in China and other decentralized systems. First, they underscore that strengthening enforcement is not primarily a question of increasing the number of cases, but of ensuring penalties are economically deterrent, i.e., that they remove the gains from violating and cannot be rationally treated as routine operating expenses, consistent with international compliance-assurance principles. Second, the evidence that only combined penalties matter suggests that regulators should prioritize enforcement tools that bind firms' production decisions, operational constraints backed by meaningful monetary sanctions, rather than relying predominantly on low-level fines that risk becoming protective. Third, because deterrence depends on credibility, reforms that strengthen transparency, standardize penalty implementation, and improve institutional follow-through may yield outsized benefits relative to statutory changes alone. Finally, the short-lived and jurisdiction-bound nature of deterrence highlights a coordination problem: episodic local enforcement may generate only temporary compliance while pollution externalities persist, strengthening the case for sustained oversight and cross-jurisdictional coordination.

2. Institutional Background and Conceptual Framework

2.1. Environmental administrative penalties and disclosure

Although many countries have expanded the environmental policy toolkit, combining market-based instruments, disclosure rules, and financial channels, administrative penalties remain the backbone of routine enforcement in most regulatory systems. In China, they are the workhorse instrument for day-to-day compliance management, operating alongside emissions trading pilots, green finance policies, and mandatory environmental information disclosure. This institutional reliance makes administrative penalties particularly suitable for studying how enforcement shapes behavior outside the set of directly punished firms.

A major inflection point in China's enforcement regime occurred with the revision of the

Environmental Protection Law, adopted on April 24, 2014 and effective January 1, 2015. The revised law strengthened statutory liabilities and elevated local-government responsibility for environmental quality, catalyzing a visible rise in enforcement activity. Figure 1 documents a sharp increase in penalty cases after 2015, consistent with a structural shift in enforcement intensity induced by this reform.

[Insert Figure 1 about here]

Administrative penalties are imposed through a structured yet decentralized process. Municipal environmental authorities typically detect violations through routine inspections, monitoring anomalies, complaint-based investigations, and special enforcement campaigns. After detection, cases proceed through standard administrative steps: case filing, investigation and evidence collection, internal legal review, notice-and-hearing (where required), and the final penalty decision, reflecting general administrative-law principles of due process and procedural regularity.

Two features of this institutional architecture are central for our empirical design. First, penalties vary meaningfully in their bindingness: some cases involve monetary fines only, some involve production restrictions (e.g., suspension or limits), and others bundle both. This heterogeneity is not a detail; it is the primary reason aggregate enforcement intensity can be an uninformative measure of deterrence. Second, the process often includes a substantial and irregular lag between violation detection and public disclosure of the penalty decision.[3] Disclosure therefore plays a distinct informational role: it is the moment when enforcement becomes salient and publicly observable to local peers, enabling belief updating about the regulator's willingness and ability to impose meaningful sanctions (Foulon et al., 2002).

[3] A representative case from Xiamen illustrates this time lag. On November 10, 2020, authorities identified excessive air pollutant emissions of Xiamen Youlaisi Ball Industry Co., Ltd. during a routine inspection. The prior notice of administrative penalty was issued four months later, on March 10, 2021. After the firm's unsuccessful appeal, the final penalty decision was issued on April 16, 2021, and publicly disclosed on May 19, 2021, approximately five months after the initial violation.

2.2. Theoretical motivation

The empirical enforcement literature has documented important effects of regulatory monitoring and sanctions on compliance and environmental outcomes, but it also highlights why deterrence can be uneven in practice. A large body of work grounded in the economics of crime and law enforcement emphasizes that firms compare the private gains from non-compliance with the expected penalty, determined by sanction severity and the probability of effective enforcement (Becker, 1968; Polinsky and Shavell, 2000). When either component is low, violations can remain privately optimal even under a formally strict legal regime.

Two broad limitations motivate our focus on penalty design and credibility. The first is under-deterrent monetary sanctions. When fines are systematically below the gains from violating, they can function as a predictable price for non-compliance rather than a deterrent (Harrington, 1988). This concern appears explicitly in regulatory guidance, e.g., the OECD's compliance-assurance recommendations stress that sanctions should remove the economic benefit of non-compliance, and it is consistent with evidence that violations can remain profitable even after penalties in some settings (Gray and Shimshack, 2011; Stafford, 2002; Atkinson, 2023).

The second limitation is weak or selective implementation, which undermines deterrence even when statutory penalties are high. Recent work shows that enforcement regimes can be highly dynamic: regulators may escalate scrutiny and impose more severe consequences on repeat offenders, generating large welfare gains relative to static fine schedules (Blundell et al., 2020). But when firms perceive enforcement as negotiable, episodic, or unevenly applied, the perceived probability of effective punishment is low, muting behavioral responses. This insight aligns with a growing literature emphasizing the importance of institutional capacity and credibility for environmental policy effectiveness (Greenstone and Jack, 2015; Shapiro and Walker, 2018).

Most importantly for this paper, sanctions can shape outcomes not only through specific deterrence (the response of punished firms), but also through general deterrence (the response of peers who learn from observed enforcement actions). In settings where firms are uncertain about enforcement severity and follow-through, publicly visible penalties can update beliefs about enforcement risk and induce compliance by firms that were not sanctioned. This channel has been

documented in environmental contexts where regulator reputation and enforcement signals influence compliance more broadly (Shimshack and Ward, 2005, 2008; Rincke and Traxler, 2011).

Insights from social learning and institutional perspectives further clarify why public penalties may matter as information. Firms learn not only from direct payoff consequences but also from salient events in their environment that reveal how authorities behave; they may update expectations and mimic precautionary behavior in response to credible enforcement signals (Bandura, 1977; DiMaggio and Powell, 1983). In the enforcement context, this implies that the deterrent effect of penalties depends on how informative they are about the regulator's willingness to impose meaningful costs and the likelihood that those costs are actually enforced.

Taken together, this literature motivates two core premises that guide our empirical work. First, policy evaluation should move beyond more vs. less enforcement to which enforcement instruments, because different sanctions change incentives through distinct margins and can be complements rather than substitutes. Second, general deterrence should be strongest when penalties are both binding and credible, meaning they raise expected violation costs and are perceived as signals of sustained enforcement rather than symbolic punishment.

2.3. A simple framework for general deterrence through public disclosure

We formalize these ideas with a parsimonious framework rooted in standard deterrence logic. Consider a continuum of risk-neutral firms indexed by j, each choosing violation intensity O_{jt}. Let $\pi_{jt} > 0$ denote the private benefit from non-compliance (e.g., avoided abatement cost). Enforcement is characterized by two primitives: the probability that a violation results in effective punishment, p_{it}, and the severity of punishment conditional on enforcement, f_{it}. A reduced-form decision rule is:

$$O_{jt} = O\left(p_{it}, f_{it}, \pi_{jt}, u_{jt}\right),\ \frac{\partial O}{\partial p} < 0,\ \frac{\partial O}{\partial f} < 0,\ \frac{\partial O}{\partial \pi} > 0$$

where u_{jt} captures other determinants of compliance (technology, financing conditions, governance, and local institutions). In the canonical model, deterrence strengthens when the expected penalty $p_{it}f_{it}$

rises relative to the gains from violating π_{jt} (Becker, 1968; Polinsky and Shavell, 2000).

To capture general deterrence, we allow firms to observe enforcement only imperfectly. Let $\tilde{p}_{it}$ denote firms' perceived probability of effective punishment. Firms update $\tilde{p}_{it}$ when enforcement becomes publicly visible, for example, when penalty decisions are disclosed. We model belief updating as:

$$\tilde{p}_{i,t} = (1 - \lambda)\tilde{p}_{i,t-1} + \lambda \cdot s_{i,t-1}$$

where $s_{i,t\text{-}1}$ summarizes the information content of disclosed enforcement in jurisdiction i (increasing in the frequency, severity, and perceived bindingness of observed penalties), and $\lambda \in (0, 1)$ captures the salience and credibility of the signal. When λ is larger, firms place more weight on newly observed penalties, consistent with stronger institutional credibility and expected follow-through.

This formulation yields two implications directly aligned with our empirical hypotheses. First, penalties that are more binding, particularly combined sanctions that couple fines with operational restrictions, should generate larger $s_{i,t\text{-}1}$ and thus stronger belief updating than fines or restrictions alone. Second, even for the same observed penalty, belief updating (and hence general deterrence) should be stronger where λ is higher, i.e., where enforcement is perceived as credible, sustained, and difficult to bargain away. Through these channels, public penalty disclosures can reduce violation intensity among non-sanctioned firms by increasing perceived expected violation costs, thereby improving ambient environmental outcomes.

3. Data and Empirical Strategy

3.1. Data and sample construction

We construct a prefecture-by-month panel covering the period 2015-2022. This sample period captures the implementation of China's revised Environmental Protection Law enacted in 2015, while concluding before the introduction of the new Ecological and Environmental Administrative Penalty Rules in 2023, thereby providing a clean identification window unaffected by subsequent major

regulatory changes. The unit of observation is a prefecture-level city i in month t. The empirical setting is motivated by the fact that environmental administrative penalties are publicly disclosed and thus observable not only to regulators and sanctioned firms, but also to local peers. This disclosure feature enables us to examine general deterrence, how local firms adjust behavior when enforcement becomes salient, rather than restricting attention to firm-specific responses.

Our key enforcement data consist of administrative penalty decisions disclosed by environmental authorities. We aggregate penalty events to the city-month level and construct separate measures for different penalty instruments, reflecting that enforcement actions differ in bindingness and informational content. The baseline enforcement measure is $lnPenalty_{it}$, the log number of disclosed penalty cases in city i and month t. To capture heterogeneity in enforcement tools, we further classify penalties into (i) fine-only cases, (ii) production restriction-only cases, and (iii) combined penalties that bundle a monetary fine with a production restriction. The instrument-level decomposition is central because the theoretical and policy discussion surrounding protective fines implies that not all penalties raise expected violation costs in the same way.

Our primary outcome is ambient sulfur dioxide concentration (SO_2), measured at high frequency and aggregated to the monthly city level. We focus on SO_2 because it is a canonical industrial pollutant with substantial variation across space and time, and it plausibly responds in the short run to changes in compliance behavior and production intensity. We use the logarithmic transformation $lnSO_2$ to reduce the influence of outliers and allow coefficients to be interpreted approximately as elasticities.

We control for meteorological conditions that affect pollution dispersion and formation, including temperature, precipitation, wind speed, and other standard weather covariates where available. These controls are important because month-to-month air quality varies strongly with local atmospheric conditions (Auffhammer et al., 2009; Deryugina et al., 2019). We also include time-varying socioeconomic covariates at the city level, such as population, economic activity, and industrial structure, when available, to reduce concerns that enforcement intensity proxies for local economic cycles (Chen et al., 2013). To reduce sensitivity to potential measurement error and extreme observations that could unduly influence the estimates, we winsorize all continuous variables at the

1st and 99th percentiles. Variable definitions and summary statistics are reported in Table 1.

[Insert Table 1 about here]

3.2. Empirical strategy

We begin with a two-way fixed effects model in which current pollution is regressed on lagged penalty intensity:

$$lnSO_{2_{it}} = \beta \cdot lnPenalty_{i,t-1} + X'_{it}\gamma + \mu_i + \tau_t + \varepsilon_{it}$$

Here, μ_i are city fixed effects absorbing time-invariant differences across cities (i.e. geography, industrial composition, historical regulation), and τ_t are time fixed effects (e.g., year-by-month) absorbing nationwide shocks, macroeconomic conditions, and seasonality common to all cities. X_{it} includes meteorological and other time-varying controls. We lag penalties by one month to reflect adjustment frictions and to reduce mechanical simultaneity between measured SO_2 and enforcement disclosures.

To examine heterogeneous deterrence across enforcement instruments, we estimate:

$$lnSO_{2_{it}} = \beta_1 \cdot lnCombined_{i,t-1} + \beta_2 \cdot lnFineOnly_{i,t-1} + \beta_3 \cdot lnRestrictOnly_{i,t-1} + X'_{it}\gamma + \mu_i + \tau_t + \varepsilon_{it}$$

This decomposition maps directly to the conceptual framework: if deterrence requires both binding costs and credible escalation, we expect $\beta_1 < 0$ and larger in magnitude than β_2 and β_3. Standard errors are clustered at the city level to allow serial correlation within cities.

4. Baseline Results

This section presents the baseline empirical evidence in three steps. We first show that aggregate

enforcement intensity, measured by the total number of disclosed penalty cases, has no robust association with ambient SO_2, but that this masks sharp heterogeneity across instruments: pollution reductions are driven exclusively by combined penalties that bundle monetary fines with production restrictions. We then strengthen causal interpretation using quasi-experimental evidence around the 2015 enforcement reform and a battery of robustness and inference checks, including controls for concurrent policies, richer fixed effects, alternative treatment definitions, and permutation placebos. Finally, we examine a mechanism consistent with protective fines, showing that fine-only penalties can be ineffective, and in low-fine jurisdictions may even coincide with higher pollution, whereas combined sanctions remain deterrent.

4.1. Aggregate enforcement is not a sufficient statistic

We begin with the baseline two-way fixed-effects specification in equation (1), relating monthly ambient SO_2 to lagged penalty intensity. Table 2, Columns (1)-(3), shows that the coefficient on total penalties is economically small and statistically indistinguishable from zero across progressively richer specifications. For example, with meteorological and socioeconomic controls, the estimate remains close to zero (Column (3): β = -0.003, se = 0.006). This result indicates that case counts alone are an uninformative proxy for deterrence, consistent with the idea that enforcement must be evaluated not only by frequency but by the expected sanction (Becker, 1968; Polinsky and Shavell, 2000). This aggregate null is not surprising if penalties differ substantially in bindingness and informational content. When firms learn from disclosed enforcement, what matters is whether observed penalties credibly signal a meaningful increase in expected violation costs.

[Insert Table 2 about here]

4.2. Instrument heterogeneity: Only combined penalties reduce SO_2

Table 2, Column (4), decomposes penalties into three mutually exclusive categories: monetary fines only, production restrictions only, and combined penalties. This decomposition reveals sharp heterogeneity: only combined penalties are associated with a statistically significant reduction in SO_2.

The estimated elasticity is -0.023 (se = 0.009, $p < 0.01$). In a log-log specification, doubling combined-penalty disclosures in month t-1 reduces SO_2 in month t by approximately 2.3%, and a one-standard-deviation increase in combined penalties corresponds to a 0.022 standard deviation reduction in SO_2.

The absence of effects for fine-only and restriction-only penalties is consistent with the complementarity logic. Fine-only penalties may fail to deter when fines are systematically below the gains from violating or are expected to be negotiable (priced permission). Standalone production restrictions may fail when they are perceived as temporary, selectively implemented, or partially circumventable. In contrast, combined penalties plausibly represent credible escalation, they impose immediate operational losses and signal stronger follow-through, raising both the perceived severity and the perceived certainty of enforcement. This mechanism aligns with recent evidence that dynamic enforcement and escalation can generate large compliance gains relative to static fine-only regimes (Blundell et al., 2020), and with broader evidence that monitoring and enforcement effectiveness depends on instrument design (Gray and Shimshack, 2011).

Figure 2 provides complementary evidence by tracing the dynamic effects of enforcement over subsequent months and directly comparing aggregate penalties to instrument-specific penalties. Two patterns stand out. First, the dynamic response to aggregate penalties remains weak throughout the post-disclosure horizon, consistent with the null effect in Table 2 when penalties are pooled. Second, and crucially, the deterrent response is concentrated in the combined-penalty series: SO_2 declines in the months immediately following combined-penalty disclosures, while the fine-only and restriction-only series remain close to zero. Taken together, Table 2 and Figure 2 imply that what matters for general deterrence is not the number of disclosed cases, but whether disclosed enforcement conveys a binding and credible increase in expected violation costs.

[Insert Figure 2 about here]

The dynamic evidence also suggests that deterrence is transient rather than structural. The pollution reduction effect is statistically significant but short-lived, persisting for only about three months

following public disclosure, consistent with a model in which firms temporarily intensify compliance when enforcement salience is high and then revert as attention and scrutiny fade. This motivates the dedicated temporal analysis in Section 6, where we more formally characterize duration using spike-based designs.

4.3. Identification: Quasi-experimental evidence around the 2015 law

A natural concern is endogeneity: enforcement may respond to pollution shocks, and therefore the baseline associations could understate or misstate causal effects. To strengthen identification, we exploit the implementation of the revised Environmental Protection Law in January 2015 as a quasi-natural experiment and implement a difference-in-differences design based on pre-determined variation in enforcement capacity. Specifically, we classify cities by whether environmental agency staffing per institution in 2013 exceeds the sample median (two years prior to the shock), and estimate equation (3) in a 24-month window around the reform (January 2014-December 2015). This pre-determined measure captures intrinsic differences in regulatory capacity across cities, reflecting their organizational readiness to implement the new law's stringent provisions (Wang et al., 2017).

$$lnSO_{2it} = \beta \cdot HighStaffperAgency_i \times Post_t + X_{it}'\gamma + \mu_i + \tau_t + \varepsilon_{it}$$

where *High Staff per Agency*$_i$ indicates cities with above-median environmental agency staff per institution in 2013, and *Post*$_t$ marks the period following the law's implementation in January 2015. The coefficient of interest, β, captures the differential change in SO_2 concentrations in high-staff cities relative to low-staff cities following the law implementation. Table 3 shows that the interaction *High Staff per Agency* × *Post* is negative and statistically significant at the 10% level (β = -0.062, se = 0.035). This implies that more administratively prepared cities experienced a 6.2% greater reduction in SO_2 following the reform, consistent with the idea that capacity strengthens credible enforcement and thus general deterrence.

[Insert Table 3 about here]

We further use an event-study specification to examine pre-treatment dynamics and assess whether the data are consistent with the identifying assumption. To ensure statistical power and present a parsimonious visualization of temporal patterns, we group periods beyond four months before and after the law implementation. Specifically, all periods preceding four months before the shock are consolidated into a single "-4" category, while all periods following four months after the shock are consolidated into a "4" category. As presented in Figure 3, the estimated pre-reform coefficients are statistically indistinguishable from zero, supporting the identifying assumption, while the post-reform coefficients indicate a reduction effect in better-staffed cities.

[Insert Figure 3 about here]

4.4. Robustness

We next assess whether the combined-penalty effect is robust to alternative samples, concurrent policy controls, richer fixed effects, and more demanding inference procedures, with results presented in the Online Appendix. Across all exercises, the central result is stable: combined penalties continue to predict lower SO_2, while single-dimensional penalties remain weak.

4.4.1. Sample sensitivity

We first test whether the baseline estimates are driven by particular time periods or by a subset of cities with atypical governance structures or measurement quality. In Online Appendix, Table OA1, Columns (1)-(2), restrict the sample to 2015-2019, excluding the COVID-19 period during which mobility restrictions, supply-chain disruptions, and demand shocks may mechanically alter emissions and ambient concentrations (Dang and Trinh, 2021). The combined-penalty coefficient remains negative and statistically significant (-0.028, se = 0.010), while the monetary-only and restriction-only coefficients remain statistically insignificant. This stability is important because it indicates that the main effect is not an artefact of pandemic-era noise or policy responses that could correlate with both enforcement and pollution.

Next, we test whether the findings are driven exclusively by major urban centers by excluding

provincial capitals from the sample (Table OA1, Columns (3)-(4)). Provincial capitals typically possess exceptionally strong administrative capacity, higher monitoring density, and greater policy salience compared to average municipalities (Chen et al., 2024). A natural concern is that the baseline deterrence effect might be confined to these highly resourceful and heavily scrutinized jurisdictions. However, upon excluding them, the combined-penalty estimate remains stable and statistically significant (β = -0.023, se = 0.009). This restriction mitigates the concern that the observed outcomes are merely an artifact of elite governance structures. The persistence of the combined-penalty effect in non-capital cities confirms that credible enforcement functions as a salient information shock across a broad spectrum of local institutional environments, proving that general deterrence is not limited to cities with the highest administrative capacity.

4.4.2. Controlling for concurrent policies

A second concern is that post-2015 China experienced multiple environmental policy initiatives that may be correlated with enforcement patterns and pollution outcomes. To address this, we explicitly control for two major concurrent policy channels emphasized in your design: ultra-low emission standards and central environmental inspections. The Ministry of Ecology and Environment mandated ultra-low emission transformations for coal-fired power plants, particularly units with capacity above 300 MW, requiring plants to reach ultralow emission levels for soot, SO_2, and NOx on a region-specific timeline. Following Wu and Cao (2021), we construct a region-staggered policy indicator *ULES* that equals one after 2017 for eastern provinces, after 2018 for central provinces, and after 2020 for western provinces, and zero otherwise.

We also control for the central government's inspection campaign, which may strengthen local enforcement by increasing oversight from higher-level authorities. Prior evidence show that central supervision can improve local environmental enforcement and substantially reduce industrial emissions in China (Agarwal et al., 2023; Karplus and Wu, 2023; Zhang et al., 2018). Beginning in January 2016, inspection teams started to conduct rotating inspections of provincial jurisdictions, first piloted in Hebei and then expanded through multiple rounds in July 2016, November 2016, April 2017, and August 2017, each lasting one to two months; by August 2017 the initial wave had covered

all provinces in a staggered manner. The central government subsequently conducted "look-back" inspections in 2018 (covering 20 provinces), and a second inspection wave began in a staggered manner from July 2019. To capture both the immediate crackdown and its short-run aftermath, we construct a province-level dummy, *Central Inspection*, that equals one during the inspection period and for the subsequent six months in each province (and zero otherwise). This treatment window is motivated by evidence that inspection impacts can be large but short-lived, and it allows us to absorb inspection-driven compliance responses that might otherwise be attributed to local penalty disclosures.

In Online Appendix, Table OA2 shows that after adding these controls, the combined-penalty coefficient remains negative and significant, with only modest attenuation in magnitude. These controls matter for identification because they capture nationwide or regionally targeted policy shocks that plausibly affect emissions directly (e.g., technology upgrades under ultra-low emission standards) and that may also induce enforcement activity (e.g., inspections that trigger investigations and penalty disclosure). The fact that the combined-penalty effect survives these controls supports the claim that our results are not simply picking up broader environmental governance cycles, but reflect an additional deterrence margin operating through penalty design and credibility.

4.4.3. Province-by-year shocks

To further rule out unobserved policy and economic shocks at the provincial level, such as province-wide enforcement campaigns, macroeconomic fluctuations, or regionally coordinated industrial policies, we replace year fixed effects with province-by-year fixed effects (de Chaisemartin and D'Haultfoeuille, 2020). This specification effectively identifies the effect of penalties using within-province, within-year variation across cities, absorbing any province-wide annual changes that could simultaneously affect enforcement and pollution.

As shown in Table OA3, the combined-penalty coefficient remains negative and statistically significant (β = -0.015, se = 0.006). Although the magnitude declines relative to the baseline, the persistence of the effect under this demanding specification is informative: it suggests the deterrence effect is not driven by province-wide shocks or broad time trends but is present even when comparing

cities facing the same province-year environment.

4.4.4. Inference robustness

We next evaluate whether statistical significance depends on a particular clustering choice or on asymptotic approximations. Because enforcement and pollution can exhibit serial correlation and because environmental policies may generate common shocks across cities in the same month, we compute standard errors using two-way clustering (city × year-month). Table OA4 shows that the combined-penalty coefficient remains statistically significant under this more conservative inference procedure.

We also implement a permutation-based placebo test that repeatedly assigns the treatment timing at random and re-estimates the main specification (500 permutations), following the logic of randomization inference (Hagemann, 2019). Figure OA1 demonstrates that the placebo distribution is centered near zero, while the true estimated effect lies in the tail of this distribution. This result provides an additional layer of credibility: if the main effect were driven by generic seasonal patterns, slow-moving city characteristics, or generic correlations between enforcement and pollution, the placebo estimates would more frequently resemble the true estimate. Instead, the real timing appears empirically special, consistent with the interpretation that penalty disclosures convey salient information that induces behavioral response.

4.4.5. Semi-continuous treatment measurement

Penalty variables are semi-continuous: many city-months have zero cases, while some months exhibit spikes. This data-generating process can raise concerns that log transformations or count measures are unduly influenced by a small number of high-enforcement months. We therefore test alternative operationalizations. First, we use binary indicators that equal one if a city experiences any penalty of a given type in month t-1. The combined-penalty indicator remains negative and significant (Table OA5, Column (2): β = -0.033, se = 0.014). This shows that the main result is not driven solely by the intensive margin of large case counts; rather, the presence of combined penalties itself, consistent with an escalation signal, predicts lower subsequent pollution.

Second, we restrict the sample to city-months with any penalties and re-estimate the specification to focus on variation in penalty composition among treated observations. Table OA5, Column (4), confirms that even within the penalty-experiencing sample, the deterrence effect is driven by combined penalties rather than fines or restrictions alone. This exercise directly reinforces the core interpretation of the paper: the policy-relevant margin is not whether enforcement occurs at all, but whether enforcement takes a form that is binding and credible enough to affect behavior. Together, these semi-continuous checks support the conclusion that combined penalties are not simply a proxy for more enforcement, but represent a distinct enforcement instrument whose informational content and economic bite generate general deterrence.

4.5. Protective fines as a mechanism extension: When monetary-only penalties backfire

A distinctive and policy-relevant result emerges once we examine the extensive margin and heterogeneity in fine severity. In Table OA5, the dummy for fine-only penalties is positive (Column (2): 0.026, se = 0.012), suggesting that when fines are used alone, they may sometimes correlate with worse pollution outcomes. This motivates a mechanism consistent with protective fines: if fines are set far below compliance costs, they can reduce the perceived marginal cost of violating relative to abatement, effectively legitimizing non-compliance.

We test this hypothesis by partitioning city-month observations by average fine magnitude (75th percentile split). Table 4 shows that in low-fine jurisdictions, fine-only penalties significantly increase SO_2 ($\beta = 0.034$, se = 0.012, $p < 0.01$), whereas in high-fine jurisdictions the monetary-only coefficient is near zero. This pattern provides direct support for the protective-fines mechanism and reinforces the broader claim of the paper: penalty design matters. Sanctions must be calibrated to remove the gains from violating; otherwise they may fail to deter and can even generate perverse incentives (Blundell et al., 2020; Shimshack and Ward, 2022).

[Insert Table 4 about here]

5. Institutional Credibility: When Penalties Work Better

A central implication of the belief-updating framework is that general deterrence depends not only on the formal severity of penalties but also on their credibility, which shapes firms' perceived probability of effective enforcement p (Duflo et al., 2013). This interpretation is consistent with recent evidence that institutional arrangements strengthening higher-level control over local environmental regulators can increase punishment activity and improve environmental quality by reducing distorted local incentives and limiting local interference (Kong and Liu, 2024). To test this prediction, we exploit cross-city institutional heterogeneity: we compare penalty effects across cities where local governments face stronger versus weaker incentives and constraints to implement regulation rigorously. Specifically, we study heterogeneity by officials' career advancement incentives, local resource dependence, and the presence of environmental courts, and show that combined penalties reduce pollution primarily in settings where enforcement is perceived as more certain and credible.

5.1. Career incentive

A natural source of credibility heterogeneity in China is the incentive environment faced by local officials. For an extended period, China's cadre evaluation system, long dominated by economic performance criteria, has been widely recognized as a fundamental institutional driver of environmental degradation. Under the political tournament framework, local officials' promotion prospects depended largely on relative economic growth within their jurisdictions (Li and Zhou, 2005). This incentive structure encouraged a focus on short-term development objectives, often at the expense of environmental protection. In this setting, local governments faced a strategic trade-off: by exercising regulatory forbearance, they could ease compliance burdens on firms to support economic growth, thereby improving leaders' career prospects (Ghanem and Zhang, 2014; He et al., 2020). Such discretion weakens the credibility of regulation by lowering firms' expected likelihood of meaningful punishment and, in turn, attenuates deterrence.

Recognizing these distortions, the central government has progressively incorporated explicit environmental targets into the cadre evaluation system, to realign local incentives by making environmental performance consequential for career advancement. The State Council of China issued the Decision on Implementing the Scientific Outlook on Development and Further Strengthening

Environmental Protection in December 2006. This policy explicitly mandated that local cadres' performance in pollution reduction must serve as a core criterion for their appointment, promotion, and disciplinary actions. This institutional shift was further solidified in 2013 when the Organization Department of the Chinese Communist Party Central Committee issued the Notice on Improving the Performance Evaluation of Local Party and Government Leadership Teams and Leading Officials. This directive explicitly required an increased weight for ecological and environmental metrics in the assessment of local officials. Furthermore, the notice stipulated for the first time that performance evaluations within key ecological function zones must prioritize environmental protection (Qi et al., 2024).

Emerging evidence suggests that this shift has begun to alter officials' behavior. Wu and Cao (2021) show that improvements in environmental quality significantly increase local officials' promotion probabilities. Complementing this observational evidence, Buntaine et al. (2024) provide experimental support for an accountability mechanism: when pollution violations are made public, regulators increase enforcement effort, consistent with heightened concern about pollution-induced public unrest and its career consequences. Together, these institutional changes imply that cities differ in the extent to which officials have incentives to implement environmental regulation rigorously, and therefore differ in the credibility with which penalties are perceived by firms.

This evolving incentive framework provides a direct lens through which to examine how enforcement credibility conditions the deterrent effect of environmental penalties. When officials' career concerns remain closely tied to economic outcomes, they retain both motive and discretion to reduce the effective certainty or severity of punishment, lowering firms' perceived probability of enforcement p and undermining deterrence. By contrast, when promotion incentives are more closely linked to environmental outcomes, officials have stronger incentives to implement penalties consistently and rigorously. Such environments raise firms' perceived certainty of punishment, the key channel through which deterrence operates in standard frameworks (Becker, 1968).

We exploit this institutional heterogeneity by examining variation in city leaders' career advancement incentives. Following Wang et al. (2020), we measure officials' ex ante promotion likelihood using predetermined characteristics, starting age, education background, administrative

rank, and related traits, that jointly shape promotion prospects under China's personnel rules. This measure captures differences in incentives to exert effort in policy implementation, including environmental enforcement.

The results in Table 5 reveal pronounced heterogeneity consistent with the credibility channel. In cities where officials face above-median promotion likelihood, environmental penalties, particularly combined penalties that integrate monetary fines with production restrictions, significantly reduce ambient SO_2. This pattern is consistent with a high-certainty enforcement environment: officials with stronger career incentives are more likely to enforce penalties thoroughly and to sustain follow-up, increasing firms' perceived probability of punishment and strengthening general deterrence. In such cities, political accountability enhances enforcement credibility, inducing firms to adjust behavior ex ante rather than merely reacting to realized sanctions (Kahn et al., 2015).

In contrast, in cities where officials face below-median promotion likelihood, penalties are associated with higher pollution concentrations, especially for production-restriction-only penalties. This counterintuitive pattern is consistent with a breakdown of deterrence under weak credibility: when restrictions are perceived as temporary, negotiable, or selectively enforced, firms may treat them as short-lived interruptions and respond strategically, for example, by compensatory increases in production intensity once restrictions lapse, yielding little net improvement in ambient pollution. Taken together, these findings underscore that the effectiveness of environmental penalties hinges not merely on their formal existence, but on the certainty with which they are implemented and perceived by regulated firms.

[Insert Table 5 about here]

5.2. Resource dependence

Resource dependence provides a second, complementary lens through which to evaluate enforcement credibility. Cities reliant on mineral or fossil-fuel extraction often exhibit classic symptoms of the resource curse. An initial comparative advantage in natural resources can attract capital, labor, and policy support toward extractive and heavy industries, crowding out diversification and innovation

and locking local economies into pollution-intensive development paths (Sachs and Warner, 2001). Over time, this generates strong economic and institutional path dependence: fiscal revenue, employment, and social stability become anchored to "black" industries, creating a governance environment in which environmental externalities are systematically tolerated.

This path dependence is reinforced by a dual lock-in mechanism. Technologically, large sunk investments in mining, energy, and heavy industrial infrastructure imply high adjustment costs, reducing the feasibility of rapid upgrading. Institutionally, a symbiotic relationship can emerge among local governments, state-owned enterprises, and affiliated stakeholders, forming a vested-interest coalition resistant to stringent regulation and structural transformation. Even when aggregate growth is strong, endogenous incentives for a green transition may remain weak.

These structural features have direct implications for the effectiveness of environmental enforcement. In resource-dependent cities, the political and economic costs of strict enforcement against major polluters, often the largest taxpayers and employers, are especially salient. Local officials thus face an acute conflict of interest and stronger incentives to exercise discretion in implementation. Empirically, such discretion can manifest as selective inspections, delayed case processing, reduced penalty severity, or informal negotiation (van Rooij et al., 2017). Importantly, these practices primarily undermine the certainty of punishment, reducing the perceived probability of effective enforcement p, rather than changing statutory severity on paper.

From the perspective of regulated firms, this enforcement environment sharply lowers perceived enforcement probability p in the compliance calculus. When violations are likely to be tolerated, postponed, or resolved informally, the expected cost of non-compliance declines, weakening the general deterrent effect of formal penalties. In addition, deterrence can be amplified by informal reputational costs. In economically diversified cities, an environmental penalty may trigger reputational damage in capital markets, labor markets, and product markets. In contrast, in mono-industrial resource-dependent cities where high-emission production is the norm, social and reputational sanctions are weaker: widespread non-compliance is normalized, reducing the marginal reputational loss associated with any single penalty and attenuating its informational signal.

To empirically assess this mechanism, we exploit the official classification in the National Plan

for the Sustainable Development of Resource-based Cities (2013-2020) issued by the State Council. Using this list, we identify 102 prefecture-level cities as resource-dependent, with all remaining cities serving as the comparison group. We then estimate subgroup regressions that allow the penalty-pollution relationship to differ across the two sets of cities.

The results, reported in Table 6, reveal a stark contrast. In non-resource-dependent cities, combined penalties significantly reduce SO_2, indicating a strong general deterrent effect. In these cities, more diversified industrial structures and lower fiscal dependence on major polluters plausibly enhance enforcement credibility, raising firms' perceived enforcement probability p. By contrast, in resource-dependent cities, the estimated effects of environmental penalties are economically small and statistically insignificant. This pattern is consistent with the view that entrenched political-economic constraints dilute enforcement certainty, eroding the deterrent power of penalties despite their formal existence.

[Insert Table 6 about here]

Overall, the resource-dependence heterogeneity underscores a central implication of the conceptual framework: the same enforcement instrument can translate into very different behavioral responses depending on local political-economic context. Where resource dependence weakens enforcement credibility, firms rationally discount the probability of effective punishment, lowering p and substantially undermining general deterrence.

5.3. Environmental adjudication

A third institutional condition shaping enforcement credibility is the strength of the judicial pillar of environmental governance. The establishment of specialized environmental courts constitutes an important institutional reform that can increase the certainty and credibility of sanctions by strengthening the likelihood that administrative penalties are upheld and become binding in practice (Zeng et al., 2024). From a deterrence perspective, environmental courts operate primarily through the enforcement chain: when firms anticipate that penalties are more likely to withstand legal challenge

and be implemented consistently, their perceived probability of effective punishment p increases, strengthening general deterrence.

The first channel is procedural capacity and expertise. Environmental cases typically involve technical evidence, scientific uncertainty, and complex regulatory standards. In general courts, limited subject-matter expertise can lead to prolonged case processing, inconsistent rulings, and greater uncertainty regarding whether administrative sanctions will ultimately be sustained. These frictions reduce enforcement credibility and lower firms' perceived enforcement probability p. By contrast, environmental courts concentrate judicial expertise, standardize evidentiary standards, and streamline adjudication procedures, reducing delays and uncertainty in case resolution. This improvement in procedural efficiency increases firms' expectations that detected violations will translate into timely and predictable legal consequences.

More importantly, environmental courts can strengthen de facto judicial insulation from local economic and political pressures. As discussed earlier, local governments may have incentives to protect economically important polluters, particularly under growth-oriented promotion incentives or in resource-dependent cities. In the absence of judicial insulation, such pressures can weaken the enforcement chain by lowering the likelihood that administrative penalties are upheld in court. Environmental courts, typically established at the municipal or provincial level, are institutionally more insulated from county- or district-level intervention (La Porta et al., 2004). This structural separation can reduce local protectionism and promote more impartial application of environmental law (Zhang et al., 2019). When firms anticipate that courts are less susceptible to local bargaining or interference, the perceived probability that penalties will be enforced and sustained rises, increasing the expected cost of non-compliance.

China's experience with environmental courts provides useful institutional variation to test this mechanism. The first environmental court was established in Qingzhen (Guizhou Province) in 2007, and in 2014 the Supreme People's Court formally institutionalized the environmental and resources court system, triggering a rapid expansion of environmental courts nationwide. We leverage this variation by partitioning the sample according to whether a city has an environmental court and estimating the baseline model separately for each subsample.

The results in Table 7 reveal clear heterogeneity consistent with the credibility channel. In cities with environmental courts, combined penalties exert a larger and more precisely estimated negative effect on SO_2, indicating stronger general deterrence when administrative enforcement is backed by capable adjudication and credible legal follow-through. In cities without environmental courts, the deterrent effect of penalties is markedly weaker, consistent with a lower perceived enforcement probability p. These findings underscore that environmental penalties do not operate in an institutional vacuum: their deterrent power depends on the integrity of the enforcement chain from administrative sanction to judicial validation. By increasing the likelihood that penalties are upheld and meaningfully enforced, specialized environmental courts raise firms' ex ante expected cost of non-compliance and strengthen general deterrence. In this sense, judicial capacity and insulation function as key institutional complements to administrative enforcement, shaping how formal penalties translate into behavioral change.

[Insert Table 7 about here]

6. Temporal and Spatial Limits of General Deterrence

6.1. The short-lived nature of deterrence effects

A natural question is whether the pollution reductions reflect persistent structural change, for example, lower primary energy use, reduced entry of polluting firms, or durable abatement investment, or instead represent short-run compliance responses. To probe this, we first aggregate enforcement measures to the annual level to match the frequency of several candidate channel variables. At the annual frequency, we find no evidence that penalties affect these longer-run mechanisms, and even when using annual average SO_2 as the outcome, the estimated effect remains statistically insignificant.

This annual null is informative: it suggests that the deterrent impact of penalties is transient rather than transformational. We therefore return to the monthly panel, where we can trace the timing of responses after penalties become publicly salient. Using the monthly variation, the deterrence effect is statistically significant but highly short-lived, persisting for roughly three months after disclosure before dissipating (see Figure 2 and 3). This pattern is consistent with a scrutiny cycle interpretation:

firms temporarily curb emissions when enforcement visibility and perceived inspection risk are elevated, but revert once attention fades. The finding also aligns with related evidence that compliance responses to enforcement actions can be temporally bounded (Shimshack and Ward, 2008; Zou, 2021; Campa, 2018).

To sharpen the timing evidence and reduce sensitivity to functional form, we adopt a spike-based design inspired by work on media attention cycles (Nguyen et al., 2022). We define *Combined Penalty Spike* months as those in which a city's combined-penalty frequency is in the top 10% of the city-month distribution, then create indicators for the first four months following a spike (*M1-M4*). Table 8 reveals a clear pattern: the coefficients for *M1-M3* are negative and statistically significant, while *M4* is insignificant. Thus, unusually salient combined-penalty events generate measurable reductions in SO_2 for approximately three months, after which the effect fades.

[Insert Table 8 about here]

Taken together, the annual null and monthly dynamics indicate that penalties primarily operate through short-run deterrence rather than durable technological upgrading. This has a direct policy implication: if enforcement is episodic, the compliance response may be episodic as well. Sustained pollution abatement likely requires either persistent enforcement pressure or complementary policies that change firms' long-run abatement incentives.

6.2. Absence of spatial spillovers across jurisdictions

Air pollution is inherently spatially correlated due to atmospheric transport and shared meteorological conditions. If penalty disclosures generate general deterrence partly through information diffusion beyond jurisdictional boundaries, one might expect enforcement in one city to reduce pollution in neighboring cities. We therefore examine whether deterrence spills across space.

6.2.1. Spatial patterns: Pollution clusters, penalties do not

We begin with descriptive spatial evidence, shown in the Online Appendix. Figure OA2 shows

pronounced geographic clustering in average SO_2 concentrations, indicating that highly polluted cities tend to be geographically proximate. By contrast, Figure OA3 shows that penalty activity is distributed far more idiosyncratically, with no clear spatial agglomeration. This mismatch suggests a basic asymmetry: environmental externalities diffuse across space, while enforcement actions remain largely jurisdiction-bound.

We formalize spatial dependence in pollution using Moran's I, a standard measure of global spatial autocorrelation, ranging from -1 to 1 with positive values indicating spatial clustering and negative values indicating spatial dispersion. Table OA6 reports positive and statistically significant Moran's I for SO_2 throughout 2015-2022, confirming strong and persistent spatial autocorrelation. We further illustrate this with Moran scatterplots for four representative months (spring, summer, autumn, winter). Figure OA4 shows a consistent concentration of observations in the high-high and low-low quadrants, reinforcing the presence of robust spatial dependence.

6.2.2. Spatial econometric test: SDM estimates show no cross-city deterrence

Given the strong spatial autocorrelation in SO_2, we estimate a two-way fixed-effects spatial Durbin panel model (SDM) following LeSage and Pace (2009), which allows spatial dependence in both the dependent variable and the key enforcement variables. Model-selection tests reject spatial autoregressive (SAR) and spatial error (SEM) restrictions, supporting the SDM specification for our setting. We use a standard contiguity matrix W where $W_{ij} = 1$ if cities share a border and 0 otherwise, and adjust inference for spatial dependence.

Table 9 yields three policy-relevant results. First, pollution itself exhibits strong spatial dependence: the spatial lag of pollution $W \times lnSO_2$ is large and highly significant. Second, consistent with previous findings, combined penalties reduce local pollution: *lnCombined* enters negatively and significantly. Third, and central to this section, there is no evidence of cross-city deterrence: the spatial lag of combined penalties $W \times lnCombined$ is negative but statistically insignificant. This implies that while pollution travels across borders, the deterrent effect of enforcement does not. Because local authorities internalize only the environmental damages within their own borders, they

have little incentive to rigorously enforce regulations against pollution that spills over to neighboring regions (Sigman, 2005; Lipscomb and Mobarak, 2016). Consequently, enforcement credibility remains geographically trapped, preventing general deterrence from crossing administrative boundaries.

Interestingly, Table 9 also indicates that the spatial lag of monetary-only penalties is positive and significant ($W \times lnFineOnly$). This pattern is consistent with the protective fines interpretation: weak penalties may be interpreted as tolerable and therefore fail to discourage pollution, potentially even reinforcing permissive norms in nearby jurisdictions, whereas credible combined sanctions deter locally but do not generate measurable spillovers. Overall, the spatial results reveal a fundamental asymmetry: environmental externalities diffuse, but regulatory deterrence remains local. Firms appear to condition compliance primarily on the enforcement environment of their own jurisdiction, where sanctioning authority and follow-through are concentrated.

[Insert Table 9 about here]

7. Conclusions

This paper examines when environmental administrative penalties generate general deterrence, changes in the behavior of non-sanctioned firms that arise when publicly disclosed enforcement actions update beliefs about regulatory severity and follow-through. Using a prefecture-by-month panel (2015-2022) linking administrative penalty disclosures to ambient SO_2, we show that aggregate enforcement intensity is a poor metric for deterrence. Total penalty counts have no robust association with air quality. By contrast, disaggregating by enforcement instrument reveals sharp heterogeneity: only combined penalties, those coupling monetary fines with production restrictions, consistently reduce SO_2. This pattern is consistent with complementarity between financial and operational sanctions and with the interpretation that combined penalties function as credible escalation in the eyes of local peers, raising the expected cost of non-compliance.

We strengthen causal interpretation with quasi-experimental evidence around the 2015 enforcement reform and a battery of robustness and inference checks. The results persist after

controlling for major contemporaneous policy shocks (including ultra-low emission standards and central environmental inspections), absorbing province-by-year shocks, adopting conservative clustering schemes, and implementing permutation-based placebo tests. These exercises support the interpretation that the combined-penalty effect reflects deterrence rather than confounding policy cycles or spurious correlations.

The analysis also provides direct evidence consistent with protective fines. When monetary sanctions are used in isolation, especially in low-fine jurisdictions, the association with air pollution can be weak or even perverse. This result provides macro-level environmental evidence for the classic behavioral insight that weak sanctions can crowd out intrinsic compliance motivations, effectively transforming a legal penalty into a routine market price for rule-breaking (Gneezy and Rustichini, 2000).This finding underscores a central policy lesson: penalties that fail to remove the gains from violating can inadvertently normalize non-compliance, reducing deterrence rather than strengthening it. In this sense, enforcement effectiveness depends not merely on the existence of sanctions, but on whether sanctions are designed and implemented so that regulated firms perceive them as binding constraints rather than routine costs.

Institutional context plays a decisive role. Deterrence is stronger when enforcement is more credible, as proxied by officials' career incentives, lower resource dependence, and stronger environmental adjudication through specialized environmental courts. Interpreted through standard deterrence logic, these institutional features raise firms' perceived probability that violations will lead to enforceable consequences, strengthening general deterrence. Conversely, in environments where discretion, protectionism, or path dependence attenuate enforcement certainty, sanctions convey weaker signals and are less likely to change behavior.

Finally, we document that general deterrence is short-lived and jurisdiction-bound. Pollution responds most strongly within approximately three months of salient combined-penalty disclosures, and spatial econometric evidence indicates that while pollution itself diffuses across borders, the deterrent effect of enforcement does not. This asymmetry highlights a coordination challenge in decentralized regulatory systems: environmental externalities travel, but credible enforcement authority remains local.

Taken together, the results suggest that improving environmental outcomes is not primarily about increasing the number of enforcement actions, but about ensuring enforcement is designed to bite and credible in implementation. Policies that strengthen the complementarity between monetary and operational sanctions, reduce discretion that can render fines protective, and enhance the integrity of the enforcement chain, particularly through credible adjudication, are likely to yield larger deterrence gains than reforms that emphasize case volume alone.

Reference:

Agarwal, S., Han, Y., Qin, Y., & Zhu, H. (2023). Disguised pollution: Industrial activities in the dark. *Journal of Public Economics*, 223, 104904.

Alm, J., Jackson, B. R., & McKee, M. (2009). Getting the word out: Enforcement information dissemination and compliance behavior. *Journal of Public Economics*, 93(3), 392-402.

Atkinson, N. (2023). Profiting from Pollution. *Yale Journal on Regulation*, *41*(1), 1–28.

Auffhammer, M., Bento, A. M., & Lowe, S. E. (2009). Measuring the effects of the Clean Air Act Amendments on ambient PM10 concentrations: The critical importance of a spatially disaggregated analysis. *Journal of Environmental Economics and Management*, 58(1), 15-26.

Bandura, A. (1977). *Social learning theory*. Prentice-Hall.

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

Blundell, W., Gowrisankaran, G., & Langer, A. (2020). Escalation of Scrutiny: The Gains from Dynamic Enforcement of Environmental Regulations. *American Economic Review*, 110(8), 2558-2585.

Buntaine, M. T., Greenstone, M., He, G., Liu, M., Wang, S., & Zhang, B. (2024). Does the Squeaky Wheel Get More Grease? The Direct and Indirect Effects of Citizen Participation on Environmental Governance in China. *American Economic Review*, 114(3).

Campa, P. (2018). Press and leaks: Do newspapers reduce toxic emissions? *Journal of Environmental Economics and Management*, 91, 184-202.

Chen, J., Shi, X., Zhang, M., & Zhang, S. (2024). Centralization of environmental administration and air pollution: Evidence from China. *Journal of Environmental Economics and Management*, 126, 103016.

Chen, Y., Ebenstein, A., Greenstone, M., & Li, H. (2013). Evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy. *Proceedings of the National Academy of Sciences*, 110(32), 12936-12941.

Dang, H.-A. H., & Trinh, T.-A. (2021). Does the COVID-19 lockdown improve global air quality? New cross-national evidence on its unintended consequences. *Journal of Environmental Economics and Management*, 105, 102401.

de Chaisemartin, C., & D'Haultfoeuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, 110(9), 2964-2996.

Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, 109(12), 4178-4219.

DiMaggio, P. J., & Powell, W. W. (1983). The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review*, 48(2), 147.

Duflo, E., Greenstone, M., Pande, R., & Ryan, N. (2013). Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India. *Quarterly Journal of Economics*, 128(4), 1499-1545.

Foulon, J., Lanoie, P., & Laplante, B. (2002). Incentives for pollution control: regulation or information?. *Journal of Environmental Economics and Management*, 44(1), 169-187.

Ghanem, D., & Zhang, J. (2014). 'Effortless Perfection:' Do Chinese cities manipulate air pollution data? *Journal of Environmental Economics and Management*, 68(2), 203-225.

Gray, W. B., & Shimshack, J. P. (2011). The Effectiveness of Environmental Monitoring and Enforcement: A Review of the Empirical Evidence. *Review of Environmental Economics and Policy*, 5(1), 3-24.

Greenstone, M., & Jack, B. K. (2015). Envirodevonomics: A Research Agenda for an Emerging Field. *Journal of Economic Literature*, 53(1), 5-42.

Gneezy, U., & Rustichini, A. (2000). A fine is a price. *Journal of Legal Studies*, 29(1), 1-17.

Hagemann, A. (2019). Placebo inference on treatment effects when the number of clusters is small. *Journal of Econometrics*, 213(1), 190-209.

Harrington, W. (1988). Enforcement Leverage When Penalties Are Restricted. *Journal of Public Economics*, 37(1), 29.

He, G., Wang, S., & Zhang, B. (2020). Watering Down Environmental Regulation in China. *Quarterly Journal of Economics*, 135(4), 2135-2185.

Kahn, M. E., Li, P., & Zhao, D. (2015). Water Pollution Progress at Borders: The Role of Changes in China's Political Promotion Incentives. *American Economic Journal: Economic Policy*, 7(4), 223-242.

Karplus, V. J., & Wu, M. (2023). Dynamic responses of SO2 pollution to China's environmental inspections. *Proceedings of the National Academy of Sciences*, 120(17).

Kong, D., & Liu, C. (2024). Centralization and regulatory enforcement: Evidence from personnel authority reform in China. *Journal of Public Economics*, 229.

La Porta, R., Lopez-de-Silanes, F., Pop-Eleches, C., & Shleifer, A. (2004). Judicial checks and balances. *Journal of Political Economy*, 112(2), 445-470.

LeSage, J. P., & Pace, R. K. (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.

Li, H., & Zhou, L.-A. (2005). Political turnover and economic performance: the incentive role of personnel control in China. *Journal of Public Economics*, 89(9-10), 1743-1762.

Lipscomb, M., & Mobarak, A. M. (2017). Decentralization and Pollution Spillovers: Evidence from the Re-drawing of County Borders in Brazil. *Review of Economic Studies*, 84(1 (298)), 464-502.

Nguyen, D. D., Ongena, S., Qi, S., & Sila, V. (2022). Climate Change Risk and the Cost of Mortgage Credit. *Review of Finance*, 26(6), 1509-1549.

Polinsky, A. M., & Shavell, S. (2000). The Economic Theory of Public Enforcement of Law. *Journal of Economic Literature*, 38(1), 45-76.

Qi, Y., Yin, A., Chen, J., Yang, C., & Zhan, P. (2024). Motivating for environmental protection: Evidence from county officials in China. *World Development*, 184.

Rincke, J., & Traxler, C. (2011). Enforcement spillovers. *Review of Economics and Statistics*, 93(4), 1224-1234.

Sachs, J. D., & Warner, A. M. (2001). The curse of natural resources. *European Economic Review*, 45(4-6), 827-838.

Shapiro, J. S., & Walker, R. (2018). Why Is Pollution from US Manufacturing Declining? The Roles of Environmental Regulation, Productivity, and Trade. *American Economic Review*, 108(12), 3814-3854.

Shimshack, J. P., & Ward, M. B. (2005). Regulator reputation, enforcement, and environmental compliance. *Journal of Environmental Economics and Management*, 50(3), 519-540.

Shimshack, J. P., & Ward, M. B. (2008). Enforcement and over-compliance. *Journal of Environmental Economics and Management*, 55(1), 90-105.

Shimshack, J. P., & Ward, M. B. (2022). Costly sanctions and the treatment of frequent violators in regulatory settings. *Journal of Environmental Economics and Management*, 116, 102745.

Sigman, H. (2005). Transboundary spillovers and decentralization of environmental policies. *Journal of Environmental Economics and Management*, 50(1), 82-101.

Stafford, S. L. (2002). The Effect of Punishment on Firm Compliance with Hazardous Waste Regulations. *Journal of Environmental Economics and Management*, 44(2), 290-308.

van Rooij, B., Zhu, Q., Na, L., & Qiliang, W. (2017). Centralizing Trends and Pollution Law Enforcement in China. *The China Quarterly*, 231, 583-606.

Wang, R., Wijen, F., & Heugens, P. P. M. A. R. (2017). Government's green grip: Multifaceted state influence on corporate environmental actions in China. *Strategic Management Journal*, 39(2), 403-428.

Wang, Z., Zhang, Q., & Zhou, L.-A. (2020). Career Incentives of City Leaders and Urban Spatial Expansion in China. *Review of Economics and Statistics*, 102(5), 897-911.

Wu, M., & Cao, X. (2021). Greening the career incentive structure for local officials in China: Does less pollution increase the chances of promotion for Chinese local leaders? *Journal of Environmental Economics and Management*, 107, 102440.

Zeng, H., Ren, L., Chen, X., Zhou, Q., Zhang, T., & Cheng, X. (2024). Punishment or deterrence? Environmental justice construction and corporate equity financing—Evidence from environmental courts. *Journal of Corporate Finance*, 86, 102583.

Zhang, B., Chen, X., & Guo, H. (2018). Does central supervision enhance local environmental enforcement? Quasi-experimental evidence from China. *Journal of Public Economics*, 164, 70–90.

Zhang, Q., Yu, Z., & Kong, D. (2019). The real effect of legal institutions: Environmental courts and firm environmental protection expenditure. *Journal of Environmental Economics and Management*, 98, 102254.

Zou, E. Y. (2021). Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality. *American Economic Review*, 111(7), 2101-2126.

Table 1. Variable definitions and summary statistics

Variable	Definition	N	Mean	SD	Min	Max
lnSO₂	Natural logarithm of the monthly average concentration of sulfur dioxide (SO_2), measured in micrograms per cubic meter ($\mu g/m^3$)	26,821	2.524	0.615	1.320	4.350
lnPenalty	Natural logarithm of (1 + the total number of environmental penalty cases publicly disclosed in the previous month)	26,821	0.508	0.971	0	4.094
lnCombined	Natural logarithm of (1 + the number of penalty cases that combine both monetary fines and production restrictions in the previous month)	26,821	0.176	0.582	0	5.875
lnFineOnly	Natural logarithm of (1 + the number of penalty cases involving only monetary fines in the previous month)	26,821	0.350	0.800	0	3.714
lnRestrictOnly	Natural logarithm of (1 + the number of penalty cases involving only production restrictions in the previous month)	26,821	0.104	0.406	0	2.485
Temperature	Monthly average temperature in degrees Celsius	26,821	14.939	10.279	-14.097	29.529
Precipitation	Monthly average precipitation in millimeters	26,821	3.224	3.100	0.023	13.835
WindSpeed	Monthly average wind speed in meters per second	26,821	9.009	2.050	5.148	15.028
PerGRP	Natural logarithm of gross regional product per capita	26,821	10.891	0.508	9.825	12.091
IndShare	Share of secondary industry output in GDP (%)	26,821	42.069	10.091	15.75	65.27
PopDensity	Natural logarithm of population per square kilometer	26,821	5.735	0.884	2.944	7.216
GreenCoverage	Green coverage rate in built-up areas (%)	26,821	40.973	4.718	21.96	50.4
PublicVehicle	Natural logarithm of (1+the stock of public buses and trolleybuses in operation per 10,000 population at the end of the year)	26,821	1.257	0.628	0.281	2.964
TechExpenditure	Science and technology expenditure as a percentage of total public fiscal expenditure (%)	26,821	1.754	1.743	0.107	9.168

Note: This table presents the definitions and summary statistics for the primary variables used in the empirical analysis. All environmental penalty variables are measured with a one-month lag to account for implementation delays. To reduce the influence of extreme outliers, all continuous variables are winsorized at the 1st and 99th percentiles.

Table 2. The impact of the environmental penalties on SO_2 concentration

Dependent Variable	*lnSO₂*			
	(1)	(2)	(3)	(4)
lnPenalty	-0.002	-0.002	-0.003	
	(0.006)	(0.006)	(0.006)	
lnCombined				-0.023***
				(0.009)
lnFineOnly				0.008
				(0.007)
lnRestrictOnly				-0.003
				(0.013)
Temperature		-0.039***	-0.039***	-0.039***
		(0.002)	(0.002)	(0.002)
Precipitation		-0.024***	-0.024***	-0.024***
		(0.002)	(0.002)	(0.002)
WindSpeed		0.012***	0.011***	0.011***
		(0.004)	(0.004)	(0.004)
PerGRP			-0.080	-0.076
			(0.061)	(0.061)
IndShare			-0.004*	-0.004*
			(0.002)	(0.002)
PopDensity			0.103	0.104
			(0.123)	(0.122)
GreenCoverage			0.003	0.003
			(0.002)	(0.002)
PublicVehicle			-0.074*	-0.074*
			(0.040)	(0.040)
TechExpenditure			-0.026**	-0.026**
			(0.013)	(0.013)
City FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Adjusted R-squared	0.716	0.745	0.747	0.748
Observations	26,821	26,821	26,821	26,821

Note: This table presents the baseline fixed-effects regression results estimating the impact of environmental administrative penalties on ambient air pollution. All specifications include city, year, and month fixed effects. Standard errors, clustered at the city level, are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 3. Difference-in-differences estimates using the 2015 Environmental Protection Law revision

Dependent Variable	$lnSO_2$
	(1)
High Staff per Agency × year2015	-0.062*
	(0.035)
Controls	YES
City FE	YES
Year FE	YES
Month FE	YES
Adjusted R-squared	0.831
Observations	5,216

Note: This table presents difference-in-differences estimates to address potential endogeneity in enforcement activity. The analysis exploits the January 2015 implementation of the revised Environmental Protection Law as a quasi-natural experiment. The sample is restricted to a 24-month window surrounding the policy shock (January 2014 to December 2015). The model includes meteorological and socioeconomic controls, alongside city, year, and month fixed effects. Standard errors are clustered at the city level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 4. The perverse effects of protective fines on air pollution

Dependent Variable	$lnSO_2$			
	High fine		Low fine	
	(1)	(2)	(3)	(4)
lnPenalty	-0.011*		0.008	
	(0.006)		(0.014)	
lnCombined		-0.017**		-0.067
		(0.008)		(0.048)
lnFineOnly		-0.001		0.034***
		(0.007)		(0.012)
lnRestrictOnly		-0.011		-0.000
		(0.015)		(0.019)
Controls	YES	YES	YES	YES
City FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Adjusted R-squared	0.747	0.747	0.756	0.756
Observations	6,700	6,700	20,114	20,114

Note: This table tests the protective-fines mechanism by estimating the heterogeneous effects of enforcement based on penalty severity. The sample is partitioned at the 75th percentile of the average fine magnitude. Columns (1) and (2) report results for the high-fine subsample (the top 25% of observations), while Columns (3) and (4) report results for the low-fine subsample (the bottom 75% of observations). All specifications include the baseline meteorological and socioeconomic controls, alongside city, year, and month fixed effects. Standard errors are clustered at the city level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 5. Heterogeneity in enforcement credibility: officials' career incentive

Dependent Variable	$lnSO_2$			
	High career incentive		Low career incentive	
	(1)	(2)	(3)	(4)
lnPenalty	-0.016*		0.024*	
	(0.009)		(0.012)	
lnCombined		-0.022*		-0.002
		(0.013)		(0.011)
lnFineOnly		-0.002		0.023
		(0.013)		(0.016)
lnRestrictOnly		-0.016		0.049**
		(0.023)		(0.021)
Controls	YES	YES	YES	YES
City FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Adjusted R-squared	0.782	0.782	0.801	0.801
Observations	7,813	7,813	5,588	5,588

Note: This table tests the institutional credibility mechanism by estimating the heterogeneous effects of environmental penalties based on local officials' career advancement incentives. The sample is partitioned at the median of officials' ex ante promotion likelihood. Columns (1) and (2) report estimates for cities where officials face high career incentives (above-median promotion likelihood), while Columns (3) and (4) report estimates for cities with low career incentives (below-median promotion likelihood). All specifications include the baseline meteorological and socioeconomic controls, alongside city, year, and month fixed effects. Standard errors are clustered at the city level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 6. Heterogeneity in enforcement credibility: resource dependence

Dependent Variable	$lnSO_2$			
	Resource-dependent		Non-resource-dependent	
	(1)	(2)	(3)	(4)
lnPenalty	-0.010		-0.001	
	(0.009)		(0.008)	
lnCombined		-0.014		-0.032***
		(0.012)		(0.011)
lnFineOnly		-0.014		0.018*
		(0.011)		(0.009)
lnRestrictOnly		0.011		-0.009
		(0.023)		(0.016)
Controls	YES	YES	YES	YES
City FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Adjusted R-squared	0.752	0.752	0.731	0.732
Observations	10,916	10,916	15,905	15,905

Note: This table tests the institutional credibility mechanism by evaluating how local resource dependence bounds the deterrent effect of environmental penalties. The sample is partitioned based on the official classification in the National Plan for the Sustainable Development of Resource-based Cities (2013-2020) issued by the State Council. Columns (1) and (2) report estimates for the 102 prefecture-level cities classified as resource-dependent, while Columns (3) and (4) report estimates for the remaining non-resource-dependent cities. All specifications include standard meteorological and socioeconomic controls, alongside city, year, and month fixed effects. Standard errors are clustered at the city level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7. Heterogeneity in enforcement credibility: environmental adjudication

Dependent Variable	$lnSO_2$			
	With environmental courts		Without environmental courts	
	(1)	(2)	(3)	(4)
lnPenalty	-0.016		-0.002	
	(0.010)		(0.007)	
lnCombined		-0.029**		-0.012
		(0.013)		(0.011)
lnFineOnly		-0.009		0.003
		(0.014)		(0.007)
lnRestrictOnly		-0.005		-0.008
		(0.018)		(0.018)
Controls	YES	YES	YES	YES
City FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Adjusted R-squared	0.757	0.757	0.758	0.758
Observations	8,638	8,638	17,991	17,991

Note: This table tests the institutional credibility mechanism by estimating how the deterrent effect of environmental penalties varies with local judicial capacity. The sample is partitioned based on whether a municipality has established a specialized environmental court. Columns (1) and (2) present estimates for cities with environmental courts, while Columns (3) and (4) present estimates for cities without such courts. All specifications include standard meteorological and socioeconomic controls, alongside city, year, and month fixed effects. Standard errors are clustered at the city level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 8. The temporal dynamics of deterrence: A spike-based design

Dependent Variable	$lnSO_2$
	(1)
lnCombined × *Combined Penalty Spike (M1)*	-0.013*
	(0.007)
lnCombined × *Combined Penalty Spike (M2)*	-0.014**
	(0.006)
lnCombined × *Combined Penalty Spike (M3)*	-0.018***
	(0.006)
lnCombined × *Combined Penalty Spike (M4)*	-0.005
	(0.008)
lnCombined	-0.003
	(0.008)
lnFineOnly	0.007
	(0.007)
lnRestrictOnly	-0.002
	(0.013)
Controls	YES
City FE	YES
Year FE	YES
Month FE	YES
Adjusted R-squared	0.747
Observations	26,295

Note: This table presents a spike-based temporal analysis evaluating the short-run duration of the general deterrence effect. The variables *Combined Penalty Spike* (*M1*) through (*M4*) are binary indicators equal to one for the first, second, third, and fourth months following an enforcement spike, respectively. A penalty spike is defined as a month in which a municipality's combined-penalty frequency falls within the top 10% of the sample distribution. The specification includes the continuous baseline penalty measures, meteorological and socioeconomic controls, alongside city, year, and month fixed effects. Standard errors are clustered at the city level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 9. Estimation results from the Spatial Durbin Model

Dependent Variable	$lnSO_2$	
	(1)	(2)
$W \times lnSO_2$	0.542***	0.542***
	(0.002)	(0.015)
lnPenalty	-0.004	
	(0.005)	
$W \times lnPenalty$	0.012	
	(0.008)	
lnCombined		-0.017**
		(0.007)
lnFineOnly		0.002
		(0.006)
lnRestrictOnly		-0.006
		(0.011)
$W \times lnCombined$		-0.017
		(0.014)
$W \times lnFineOnly$		0.022**
		(0.011)
$W \times lnRestrictOnly$		0.028
		(0.022)
Controls	YES	YES
W × Controls	YES	YES
City FE	YES	YES
Year FE	YES	YES
Month FE	YES	YES
Observations	25632	25632
R-squared	0.259	0.263
Sigma2	0.065	0.065
Log-likelihood	-2356.631	-2332.387
Wald test spatial lag	195.17***	219.67***
LR test spatial lag	233.07***	257.22***
Wald test spatial error	357.01***	390.80***
LR test spatial error	730.09***	760.39***
Hausman test	308.70***	306.00***

Note: This table presents the estimation results of a two-way fixed-effects Spatial Durbin Model (SDM) to test for cross-jurisdictional spillovers. The spatial weight matrix, denoted by W, is a standard binary contiguity matrix where $W_{ij} = 1$ if cities share a border and 0 otherwise. Model selection diagnostics (Wald, LR, and Hausman tests) strongly reject spatial autoregressive and spatial error restrictions, supporting the SDM specification. All models incorporate baseline meteorological and socioeconomic controls, their spatial lags, and city, year, and month fixed effects. Standard errors, adjusted for spatial dependence, are clustered at the city level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

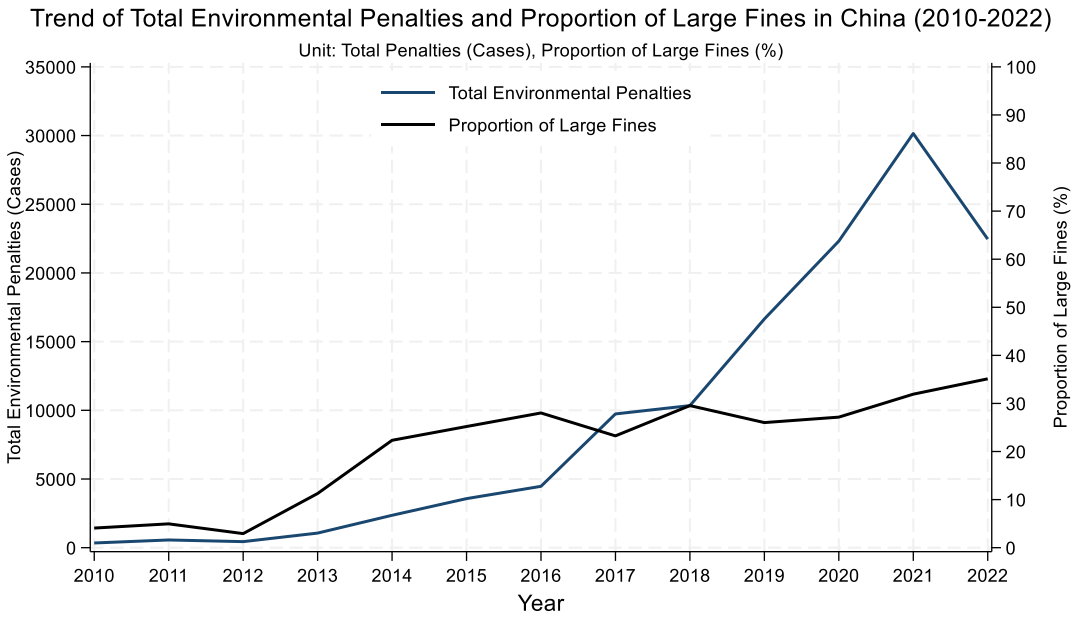


Figure 1. The Evolution of Environmental Enforcement Intensity and Penalty Severity

Note: This figure plots the annual aggregate number of environmental administrative penalty cases (left axis, solid blue line) and the proportion of cases classified as large fines (right axis, solid black line) issued by municipal environmental authorities in China. A structural break in aggregate enforcement volume is evident following the implementation of the revised Environmental Protection Law in 2015. In contrast, penalty severity, proxied by the share of large fines, exhibits a more gradual and secular upward trend over the sample period.

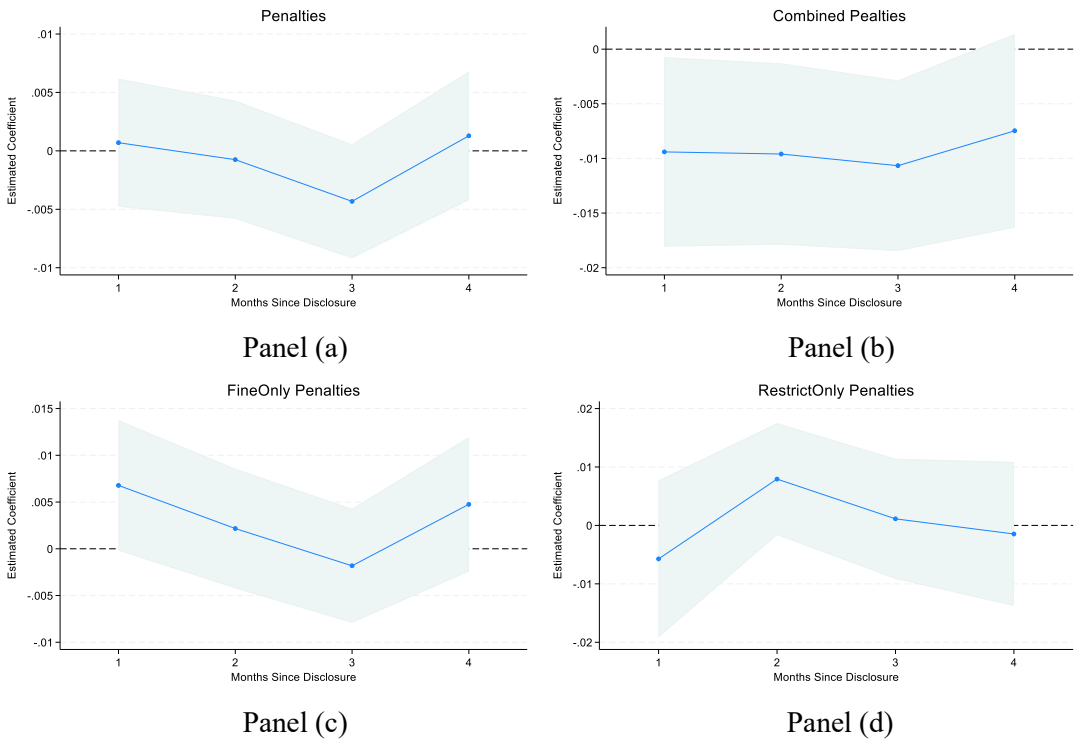


Figure 2. Dynamic Effects of Environmental Penalties on SO_2

Note: This figure plots the dynamic treatment effects of publicly disclosed environmental penalties on ambient SO2 concentrations over the subsequent four months. Panel (a) illustrates the response to aggregate enforcement volume, while Panels (b) through (d) disaggregate these effects by specific penalty instruments. The solid lines represent the point estimates, and the shaded regions indicate the 90% confidence intervals.

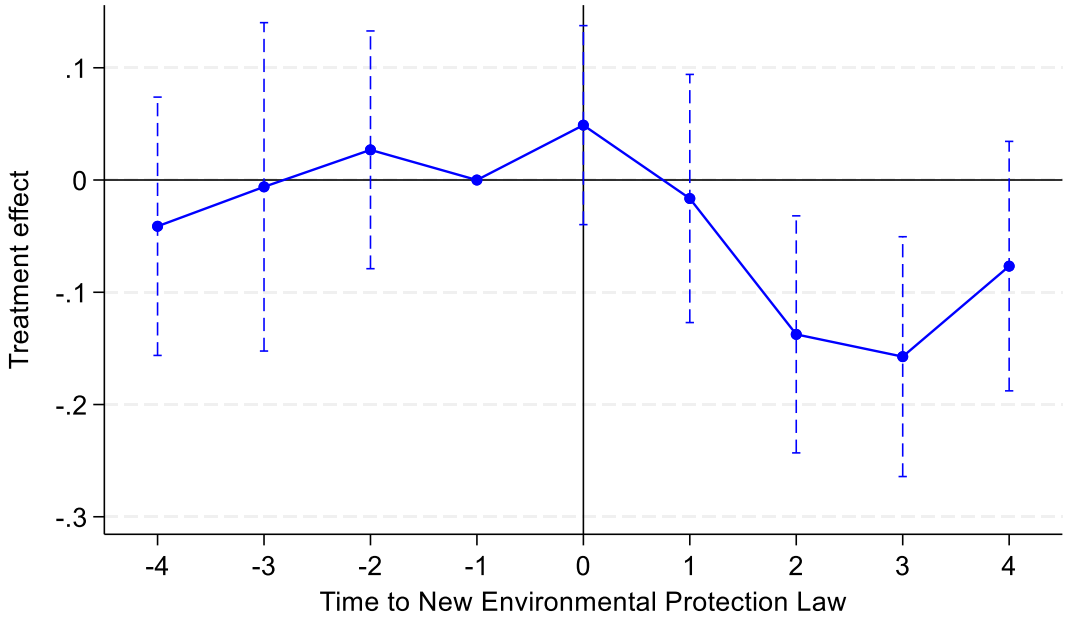


Figure 3. Event Study Estimates of the 2015 Environmental Protection Law Revision

Note: This figure displays the event-study coefficients and 95% confidence intervals evaluating the dynamic effects of the revised Environmental Protection Law. The reference period is the month prior to implementation (t=-1). Time periods exceeding four months from the policy date are binned into the respective endpoints (-4 and 4). Point estimates for the interaction between the post-treatment indicators and the high-capacity city indicator are shown alongside their 95% confidence intervals.

Online Appendix

Table OA1. Excluding special observations

Dependent Variable	$lnSO_2$			
	Excluding observations from the COVID-19 pandemic period		Excluding observations of provincial capital cities	
	(1)	(2)	(3)	(4)
lnPenalty	-0.002		-0.002	
	(0.007)		(0.006)	
lnCombined		-0.028***		-0.023**
		(0.010)		(0.009)
lnFineOnly		0.010		0.011
		(0.009)		(0.008)
lnRestrictOnly		0.015		-0.009
		(0.018)		(0.014)
Controls	YES	YES	YES	YES
City FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Adjusted R-squared	0.770	0.770	0.743	0.743
Observations	16,813	16,813	24,337	24,337

Note: This table reports robustness checks testing the sensitivity of the baseline estimates to specific time periods and city subsets. Columns 1 and 2 exclude observations from the COVID-19 pandemic period (restricting the sample to 2015-2019) to mitigate the influence of macroeconomic and mobility shocks. Columns 3 and 4 exclude provincial capital cities to ensure the findings are not driven exclusively by major urban centers with higher administrative capacity and policy salience. All specifications include meteorological and socioeconomic controls, alongside city, year, and month fixed effects. Standard errors are clustered at the city level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table OA2. Controlling for concurrent policies

Dependent Variable	*lnSO₂*			
	(1)	(2)	(3)	(4)
lnPenalty	-0.002		-0.002	
	(0.006)		(0.006)	
lnCombined		-0.023***		-0.023***
		(0.009)		(0.009)
lnFineOnly		0.007		0.008
		(0.007)		(0.007)
lnRestrictOnly		-0.001		-0.003
		(0.013)		(0.013)
ULES	-0.062***	-0.062***		
	(0.019)	(0.019)		
Central Inspection			-0.022**	-0.023**
			(0.009)	(0.009)
Controls	YES	YES	YES	YES
City FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Adjusted R-squared	0.748	0.748	0.748	0.748
Observations	26,821	26,821	26,821	26,821

Note: The dependent variable is the logarithm of SO_2 concentration. Estimates include binary controls for two concurrent environmental policies (ULES and Central Inspection) to address potential confounding, respectively. Meteorological and socioeconomic controls as specified in the main empirical specification are included. Standard errors are reported in parentheses and clustered at the city level. *p < 0.1, **p < 0.05, ***p < 0.01.

Table OA3. Controlling for province-by-year shocks

Dependent Variable	$lnSO_2$	
	(1)	(2)
lnPenalty	-0.006	
	(0.005)	
lnCombined		-0.015**
		(0.006)
lnFineOnly		0.002
		(0.005)
lnRestrictOnly		-0.021*
		(0.011)
Controls	YES	YES
City FE	YES	YES
Province × Year FE	YES	YES
Month FE	YES	YES
Adjusted R-squared	0.779	0.779
Observations	26,821	26,821

Note: This table reports robustness checks addressing potential confounding from unobserved, time-varying shocks at the provincial level. By replacing standard year fixed effects with province-by-year fixed effects, the models absorb macroeconomic fluctuations, province-wide enforcement campaigns, and regionally coordinated industrial policies. Consequently, the deterrence effect is identified strictly using within-province, within-year variation across cities. The specifications include standard meteorological and socioeconomic controls, as well as city and month fixed effects. Standard errors are clustered at the city level and reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table OA4. Two-way clustering of standard errors

Dependent Variable	$lnSO_2$	
	(1)	(2)
lnPenalty	-0.003	
	(0.006)	
lnCombined		-0.023**
		(0.009)
lnFineOnly		0.008
		(0.007)
lnRestrictOnly		-0.003
		(0.013)
Controls	YES	YES
City FE	YES	YES
Year FE	YES	YES
Month FE	YES	YES
Adjusted R-squared	0.747	0.748
Observations	26,821	26,821

Note: This table reports robustness checks addressing potential serial correlation and common shocks across cities in the same month. To evaluate whether the statistical significance of the baseline results depends on the clustering choice, standard errors are computed using two-way clustering at the city and year-month levels. The specifications include standard meteorological and socioeconomic controls, alongside city, year, and month fixed effects. Standard errors are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table OA5. Alternative specifications for semi-continuous penalty variables

Dependent Variable	$lnSO_2$			
	(1)	(2)	(3)	(4)
AnyPenalty	0.013			
	(0.012)			
AnyCombined		-0.033**		
		(0.014)		
AnyFineOnly		0.026**		
		(0.012)		
AnyRestrictOnly		0.004		
		(0.016)		
lnPenalty			-0.011*	
			(0.006)	
lnCombined				-0.017**
				(0.008)
lnFineOnly				0.001
				(0.006)
lnRestrictOnly				-0.014
				(0.015)
Controls	YES	YES	YES	YES
City FE	YES	YES	YES	YES
Year FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
Adjusted R-squared	0.747	0.748	0.740	0.740
Observations	26,821	26,821	7,411	7,411

Note: This table reports robustness checks addressing the semi-continuous nature of the enforcement data. In Columns (1) and (2), the treatment variables (*AnyPenalty*, *AnyCombined*, *AnyFineOnly*, and *AnyRestrictOnly*) are binary indicators equal to one if the respective penalty type was imposed in the previous month, and zero otherwise, testing the extensive margin. Columns (3) and (4) retain the continuous log-transformed treatment measures but restrict the sample strictly to city-months that recorded at least one penalty of any type (N = 7,411) to focus on variation in penalty composition. All specifications include the baseline meteorological and socioeconomic controls, alongside city, year, and month fixed effects. Standard errors are clustered at the city level and are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table OA6. Global spatial autocorrelation of ambient SO_2 concentrations (2015 to 2022)

Year Month	Moran's I	Year Month	Moran's I	Year Month	Moran's I
Jan.2015	0.547***	Sept.2017	0.420***	May2020	0.297***
Feb.2015	0.642***	Oct.2017	0.338***	Jun.2020	0.391***
Mar.2015	0.548***	Nov.2017	0.499***	Jul.2020	0.324***
Apr.2015	0.358***	Dec.2017	0.525***	Aug.2020	0.281***
May2015	0.503***	Jan.2018	0.590***	Sept.2020	0.334***
Jun.2015	0.433***	Feb.2018	0.633***	Oct.2020	0.465***
Jul.2015	0.337***	Mar.2018	0.453***	Nov.2020	0.431***
Aug.2015	0.328***	Apr.2018	0.305***	Dec.2020	0.495***
Sept.2015	0.425***	May2018	0.354***	Jan.2021	0.466***
Oct.2015	0.463***	Jun.2018	0.344***	Feb.2021	0.571***
Nov.2015	0.571***	Jul.2018	0.174***	Mar.2021	0.356***
Dec.2015	0.692***	Aug.2018	0.187***	Apr.2021	0.272***
Jan.2016	0.689***	Sept.2018	0.313***	May2021	0.265***
Feb.2016	0.670***	Oct.2018	0.356***	Jun.2021	0.279***
Mar.2016	0.552***	Nov.2018	0.488***	Jul.2021	0.314***
Apr.2016	0.431***	Dec.2018	0.599***	Aug.2021	0.257***
May2016	0.412***	Jan.2019	0.681***	Sept.2021	0.275***
Jun.2016	0.332***	Feb.2019	0.697***	Oct.2021	0.314***
Jul.2016	0.235***	Mar.2019	0.457***	Nov.2021	0.318***
Aug.2016	0.279***	Apr.2019	0.348***	Dec.2021	0.350***
Sept.2016	0.440***	May2019	0.376***	Jan.2022	0.560***
Oct.2016	0.408***	Jun.2019	0.310***	Feb.2022	0.493***
Nov.2016	0.613***	Jul.2019	0.256***	Mar.2022	0.204***
Dec.2016	0.613***	Aug.2019	0.315***	Apr.2022	0.185***
Jan.2017	0.678***	Sept.2019	0.322***	May2022	0.281***
Feb.2017	0.658***	Oct.2019	0.302***	Jun.2022	0.248***
Mar.2017	0.559***	Nov.2019	0426***	Jul.2022	0.200***
Apr.2017	0.391***	Dec.2019	0.487***	Aug.2022	0.167***
May2017	0.360***	Jan.2020	0.720***	Sept.2022	0.223***
Jun.2017	0.381***	Feb.2020	0.633***	Oct.2022	0.192***
Jul.2017	0.231***	Mar.2020	0.407***	Nov.2022	0.272***
Aug.2017	0.322***	Apr.2020	0.266***	Dec.2022	0.373***

Note: This table reports the monthly global Moran's *I* index for ambient SO_2 concentrations across the sample cities from January 2015 to December 2022. Moran's *I* serves as a standard measure of spatial dependence, with positive values indicating spatial clustering and negative values indicating dispersion. The consistently positive and statistically significant indices confirm a strong and persistent spatial autocorrelation in air pollution throughout the sample period, validating the use of spatial econometric specifications. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

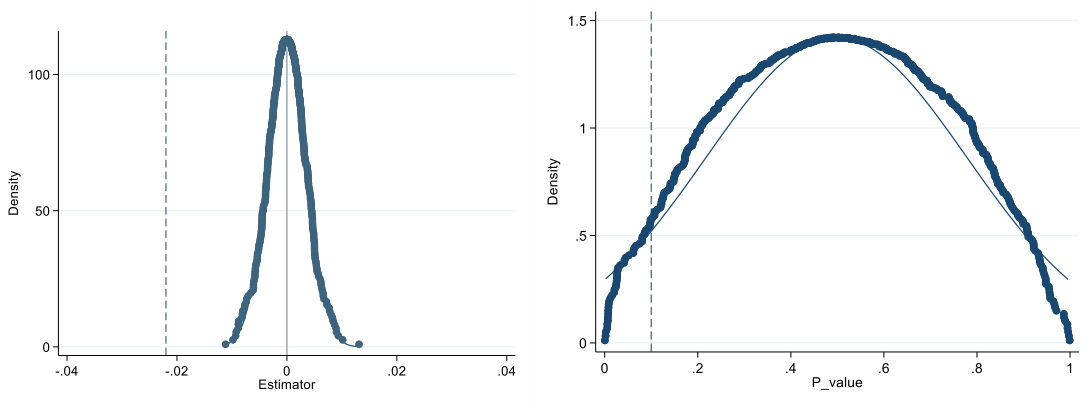


Figure OA1. Placebo test

Note: This figure presents the results of a randomization inference placebo test based on 500 random permutations of the treatment timing. The left panel plots the empirical density of the estimated coefficients from the placebo regressions, which is centered tightly around zero. The vertical dashed line indicates the actual point estimate from the true baseline model. The right panel plots the density of the p-values corresponding to these 500 placebo estimates, with the vertical dashed line marking the conventional 10% significance level ($p = 0.10$). The distribution demonstrates that the vast majority of the randomly generated placebo treatments yielded statistically insignificant results, reinforcing that the true baseline deterrence effect is highly unlikely to be driven by random chance.

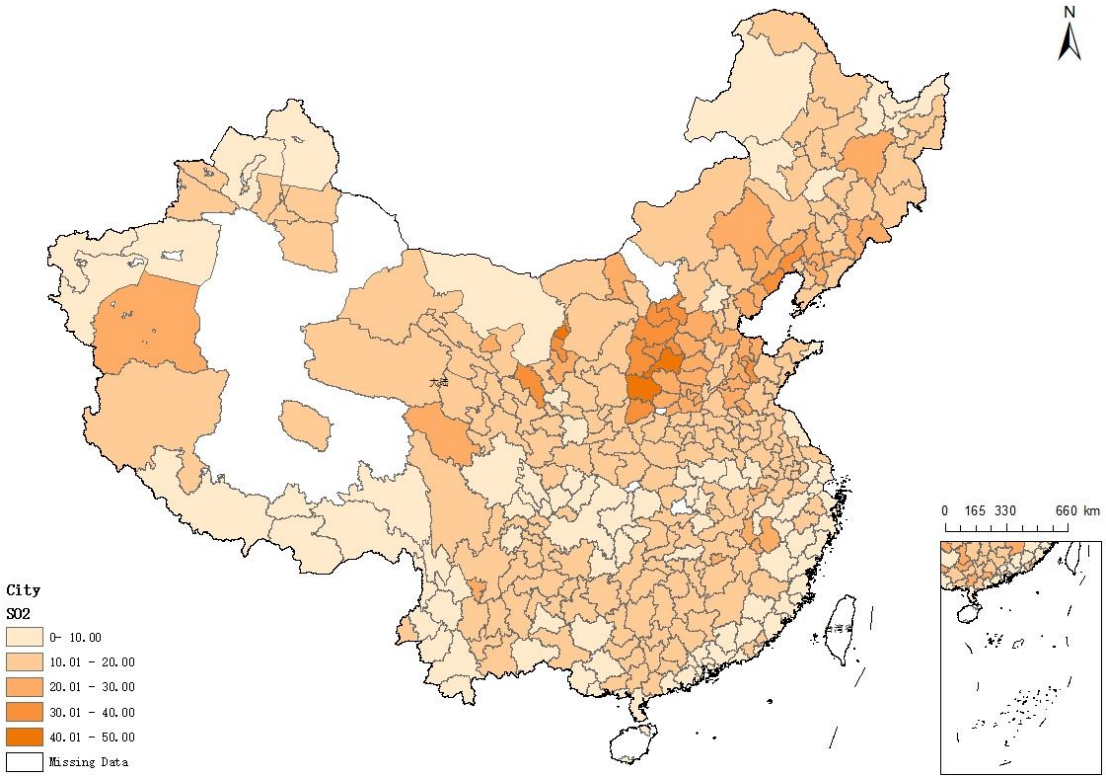


Figure OA2. Spatial distribution and geographic clustering of ambient SO_2 concentrations

Note: This figure displays the spatial distribution of average ambient SO_2 concentrations across Chinese prefecture-level cities from 2015 to 2022. The map reveals pronounced geographic clustering, indicating that jurisdictions with high pollution levels tend to be geographically proximate, which is consistent with the spatial diffusion of environmental externalities. Map lines delineate study areas and do not necessarily depict accepted national boundaries.

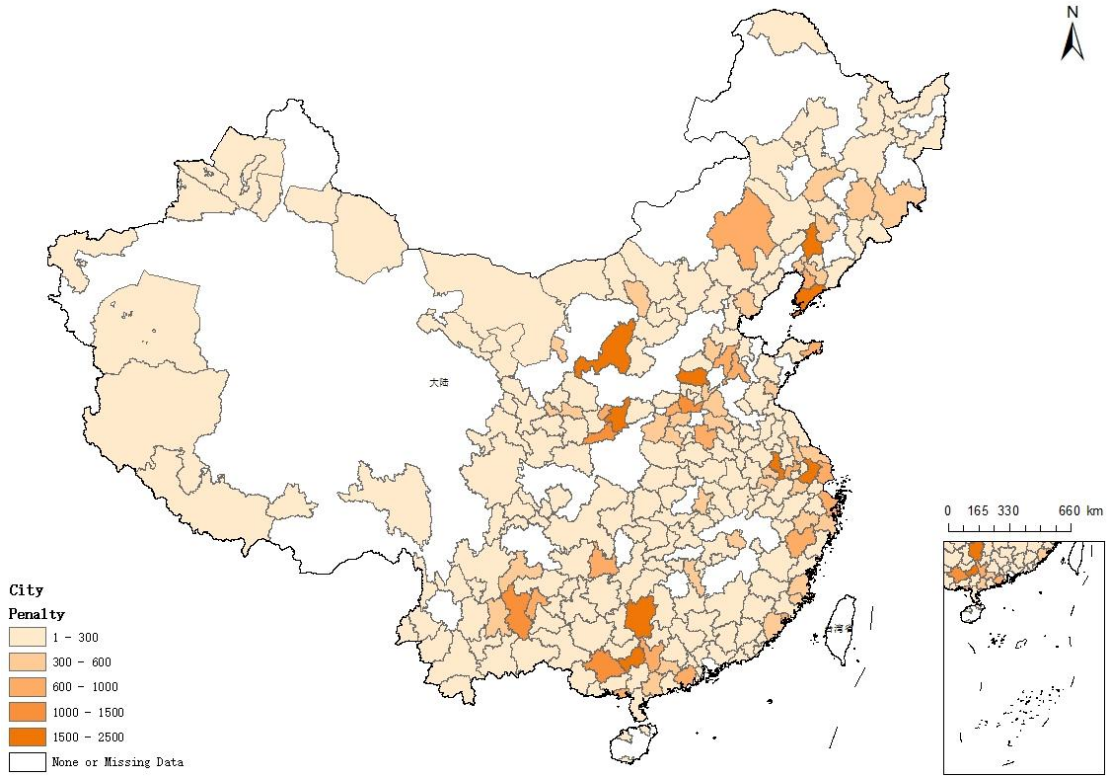


Figure OA3. Spatial distribution of environmental administrative penalties

Note: This figure plots the aggregate frequency of environmental administrative penalties across Chinese prefecture-level cities from 2015 to 2022. In contrast to the clustered pollution patterns observed in Figure 5, enforcement activity is distributed highly idiosyncratically with no clear spatial agglomeration. Map lines delineate study areas and do not necessarily depict accepted national boundaries.

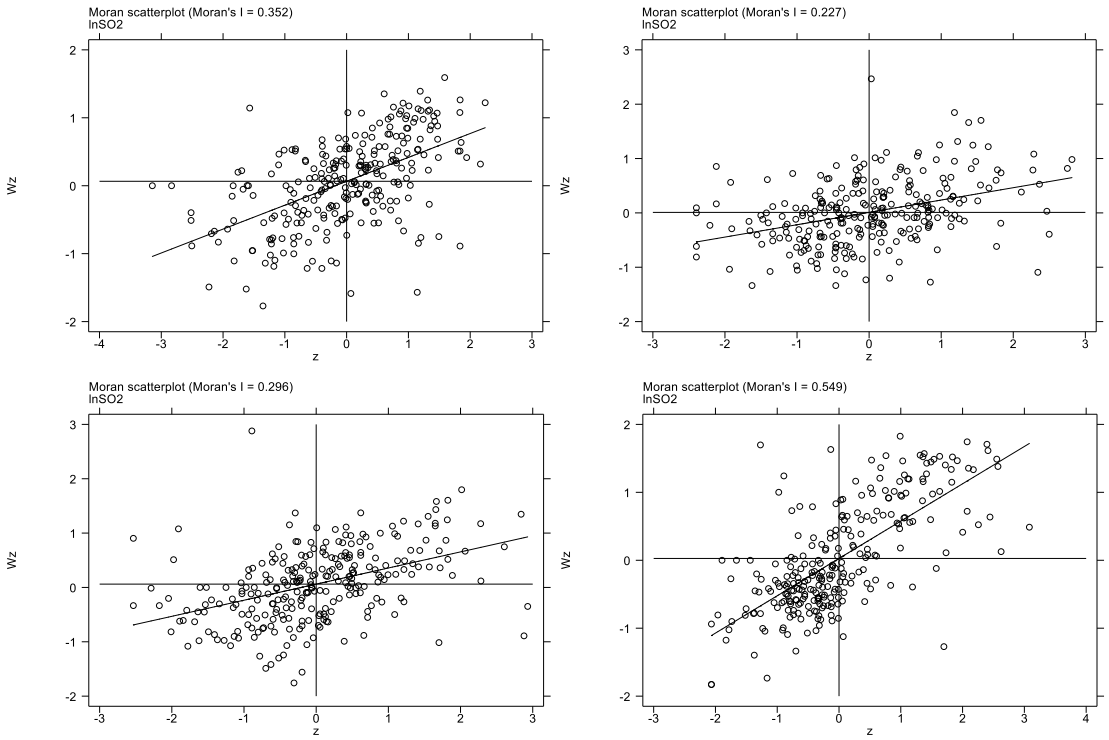


Figure OA4. Moran Scatterplots of Ambient SO_2 Concentrations for representative months

Note: This figure illustrates the global spatial autocorrelation of air pollution using Moran scatterplots for four representative seasonal months: April 2015 (left in Line 1), July 2017 (right in Line 1), October 2019 (left in Line 2), and January 2022 (right in Line 2). The plotted data reveal a consistent concentration of city-level observations within the high-high and low-low quadrants, regardless of the season.