

Under the Spotlight: Pollution Dynamics Around Corporate Site Visits

Yuchen Wei, Fujing Xue, Bohui Zhang, Xiaofeng Zhao*

Current version: January 2026

Abstract

This study examines whether firms “clean up” their emissions when they know investors are watching. In China, investors frequently conduct site visits to publicly listed firms to obtain firm-specific information. Using daily satellite data for manufacturing firms, we find that PM2.5 concentrations drop significantly around scheduled site visits. The decline is tightly timed to the visits, suggesting strategic suppression. The effects are stronger for dirtier, high-valuation, and better-governed firms, and are driven by visiting shareholders rather than Principles for Responsible Investment (PRI) signatories. These results document a new form of corporate signal manipulation and highlight limits of observed ESG diligence.

JEL Classifications: G14, G34, Q53, M41

Keywords: Corporate site visits, Strategic pollution, ESG, Real activity management, Satellite data

* Wei is based on Hong Kong University, Pokfulam Road, Hong Kong. Xue is based at Sun Yat-Sen University, No. 66 Gongchang Road, Guangming District, Shenzhen, Shenzhen, China, 518107. Zhang is based at the School of Management and Economics and Shenzhen Institute of Data Economy, The Chinese University of Hong Kong, Shenzhen (CUHK-Shenzhen), Longxiang Boulevard, Longgang District, Shenzhen, China, 518172. Zhao is based at the Faculty of Business, Lingnan University, 8 Castle Peak Road, Tuen Mun, New Territories, Hong Kong. Authors' contact information: Yuchen Wei: ycwei99@connect.hku.hk; Fujing Xue: xuefj@mail.sysu.edu.cn; Bohui Zhang: bohui Zhang@cuhk.edu.cn, (86)755-23518868; Xiaofeng Zhao: x4zhao@ln.edu.hk, (852)26168156. Zixuan Zeng: zengzixuan@cuhk.edu.cn.

1. Introduction

This study examines whether firms temporarily reduce pollution emissions around investor site visits—scheduled, in-person meetings in which investors tour a firm’s facilities and meet with management. Recent research demonstrates that regulated entities engage in strategic evasion, curbing negative externalities primarily when observed. For example, Zou (2021) finds that U.S. polluters emit significantly less on days when government air-quality monitors are scheduled to take measurements, resulting in cleaner air on monitored days and worse air on unmonitored ones—a phenomenon characterized as “unwatched pollution.” In China, Yang *et al.* (2024) show that when environmental targets are tied to monitored pollution levels, local governments concentrate anti-pollution efforts around monitoring stations, effectively gaming the system through spatial reallocation. While these studies reveal strategic compliance behavior in the public sector, the extent to which such behavior permeates the private corporate sector remains unclear.

This study fills the gap in the context of corporate site visits. Unlike regulatory inspections, which are explicitly compliance-oriented, site visits are non-punitive, scheduled interactions designed to gather information about operations and management quality (Cheng *et al.*, 2016; Han *et al.*, 2018). However, because these visits influence investor beliefs, reputation, and market value, managers may have incentives to manage what visitors observe. Specifically, we ask whether firms “clean up” their emissions when they know investors are watching, manipulating the visible environmental quality to make a favorable impression.

Addressing this question contributes to the intersection of environmental economics and finance. By documenting corporate unwatched pollution, we extend the concept of strategic evasion from public monitoring to capital-market scrutiny. This investigation is timely given the surge in ESG (Environmental, Social, and Governance) investing, where reliance on self-reported data is high and verification is costly (Berg *et al.*, 2022; Christensen *et al.*, 2022). If

firms can “greenwash” their physical operations during investor diligence, capital may flow toward firms that excel at staged compliance rather than those with persistently cleaner production. Thus, this study not only uncovers a novel form of signal manipulation but also questions the efficacy of direct observation as a monitoring tool in the presence of strategic agents.

In tackling this question, we also add to the literature on information acquisition. Prior work highlights that site visits can generate meaningful informational advantages: Cheng *et al.* (2016) find that analysts who conduct site visits in China produce more accurate earnings forecasts, and Zhang *et al.* (2025) show that firms frequently visited by investors earn higher subsequent stock returns, consistent with site visits helping investors identify undervaluation. Site visits allow investors to observe operational details, including both “hard” and “soft” signals (e.g., management demeanor, workplace conditions) that are not captured in public disclosures. We posit that managers anticipate this scrutiny and proactively shape what investors observe. By documenting this phenomenon, we extend the investor-site-visit literature beyond information acquisition to consider information distortion, highlighting the endogenous nature of private meetings.

To investigate our research question, we construct a facility-level panel that links disclosed investor site visits to daily, high-resolution measures of ambient air pollution around corporate operating sites for Chinese manufacturing firms listed on the Shenzhen Stock Exchange (SZSE) from 2012 to 2019. Our sample covers 30,948 distinct site visit events involving 2,366 facility sites affiliated with 1,350 publicly listed manufacturing firms. China provides a particularly informative setting for three reasons. First, investor site visits are a prevalent channel of information acquisition in China’s capital markets (Cheng *et al.*, 2016; Han *et al.*, 2018). Second, since 2012 the SZSE has required listed firms to disclose investor site visits within two trading days; these disclosures report the visit date and location as well

as the identities of visiting institutions, allowing us to observe visit timing and venue at scale. Third, industrial pollution is both economically material and highly visible in China, making environmental impressions plausibly consequential for investor beliefs and firm valuation (Yang *et al.*, 2024). We focus on manufacturing because visits are most informative when investors can observe tangible operations and because industrial production generates salient, localized emissions.

To compile the dataset, we identify firms' subsidiaries and production facilities and geocode their addresses into precise coordinates. We then merge each location with daily local air pollution data from a satellite-based source (the China High Air Pollutants (CHAP) dataset), which provides ground-level PM_{2.5} concentrations at 1×1 km resolution. Next, we link these outcomes to the full universe of disclosed investor relations activities, using meeting descriptions to identify in-person visits and to map each visit to the most plausible on-site location. The final panel tracks daily pollution levels for each site, yielding 6.9 million site-day observations.

We test whether a site experiences a reduction in pollution emissions when it receives a visit by controlling for contemporaneous local background pollution and absorbing time-invariant facility characteristics as well as city-by-day shocks (e.g., weather and citywide emission conditions). By exploiting within-facility variation in pollution around scheduled visits, this design allows us to attribute changes in pollution at the visited site to short-run responses to investor presence rather than to broader regional air-quality fluctuations or persistent differences across facilities.

We find that local air pollution declines meaningfully around investor site visits. In baseline regressions, the site-visit indicator loads negatively and is statistically significant, implying that PM_{2.5} concentrations at a facility are lower on visit days than on otherwise comparable non-visit days (after accounting for site and city-date fixed effects). The magnitude

is economically meaningful. On average, daily abnormal PM_{2.5} is about 16% lower on a visit day relative to non-visit days for the same site—roughly a 0.02 $\mu\text{g}/\text{m}^3$ decline against a sample mean of about 0.09 $\mu\text{g}/\text{m}^3$. We observe similar patterns when we use raw pollution readings.

Examining pollution dynamics around site visits supports the interpretation of strategic, short-lived cleanup. We find no discernible trend in the days preceding a visit; however, air quality improves sharply immediately prior to and during the visit, followed by a rapid return to baseline levels. This timing—a transient dip tightly synchronized with the visit—is consistent with deliberate short-term emissions suppression rather than lasting environmental improvement.

A series of robustness exercises reinforces our findings. Placebo tests that sever the link between realized visit timing and pollution outcomes do not reproduce the pattern. The results remain stable when we assess sensitivity to omitted variables (Oster, 2019), reweight observations to improve covariate balance between visit and non-visit days (Hainmueller & Xu, 2013), and vary sample composition across regions with different economic and regulatory environments.

We find substantial heterogeneity that sheds light on the underlying mechanism. The strategic reduction is driven primarily by firms in heavy-polluting industries (e.g., thermal power and chemicals) and by firms with poor environmental records, where regulatory risk and abatement costs are likely highest. Moreover, the effect is concentrated among firms with high market valuations and strong governance. These patterns suggest that managers of high-value firms—who face greater downside from adverse signals and operate under tighter oversight—have stronger incentives to undertake short-run abatement.

We further find that the “unwatched pollution” effect is disciplined by ownership rather than optics: visits by existing shareholders trigger meaningful clean-up, whereas visits by Principles for Responsible Investment (PRI) signatories without ownership stakes do not

generate a differential response. This evidence indicates that strategic pollution reduction is not primarily driven by a reputation channel linked to environmentally conscious visitors; instead, it reflects an ownership-based monitoring channel.

Finally, we conduct additional analyses to rule out alternative explanations. First, virtual interactions that lack physical observation do not coincide with comparable improvements in local air quality, consistent with physical monitoring being the trigger. Second, facilities of the same firm that are not visited do not exhibit parallel pollution changes during a sister facility's visit window, which argues against firm-wide initiatives or contemporaneous shocks as the primary explanation. Third, we do not detect commensurate reductions in satellite-based thermal signatures of operating intensity around visits, suggesting that firms do not achieve cleaner air by suspending production (see Xue *et al.*, 2025). Instead, the evidence is more consistent with short-run abatement effort that decouples emissions from output during periods of scrutiny.

Our study makes several contributions. First, we advance the literature on strategic environmental compliance by documenting a new setting in which strategic abatement arises. Prior work focuses on regulatory monitoring—for example, Chinese local governments manipulating pollution data to meet targets (Ghanem & Zhang, 2014) or U.S. factories shifting emissions to unmonitored times and locations (Grainger *et al.*, 2019; Zou, 2021). We complement these studies by showing that strategic “air cleaning” is not confined to regulatory oversight; it also occurs in anticipation of capital-market scrutiny. To our knowledge, we provide the first firm-level evidence that companies adjust environmental outputs for a non-regulatory audience.

Second, by documenting this form of real activity management, we also contribute to the real earnings management literature in accounting and finance, which shows that firms adjust production, pricing, or discretionary expenditures to shape reported performance

(Roychowdhury, 2006; Cohen & Zarowin, 2010; Gunny, 2010; Zang, 2012). We argue that strategic abatement during short-term investor visits represents a new form of real activity management. This distinction matters because investor scrutiny can generate transient improvements in local air quality while distorting inferences about firms' persistent environmental performance.

Third, we connect environmental economics with finance by clarifying when and how investor attention functions as a monitoring technology. A large literature shows that private interactions, including site visits, can mitigate information asymmetry (Cheng *et al.*, 2016; Kirk & Markov, 2016; Bushee *et al.*, 2018; Han *et al.*, 2018; Cheng *et al.*, 2019; Bradley *et al.*, 2022). We extend this work by showing that investor presence can also induce changes in firms' real behavior—in our setting, pollution emissions. In this sense, investor attention exerts a governance-like effect on environmental actions, though the effect is short-lived. Our evidence is consistent with the view that informed investors influence corporate behavior through monitoring (Brav *et al.*, 2008; Yu, 2008; Levit, 2019), but it also highlights a limitation: firms may respond strategically by managing the observable signal rather than improving underlying performance. This result underscores the endogenous nature of site visits and other forms of private communications—an aspect largely overlooked in the corporate communications literature (see Brown *et al.*, 2015; Kirk & Markov, 2016; Bushee *et al.*, 2018; Bradley *et al.*, 2022).

Finally, our findings have implications for both policy and investment practice. For regulators, the results indicate that predictable oversight can induce short-term adjustments that conceal underlying environmental performance, strengthening the case for continuous monitoring (e.g., real-time emissions sensors). For investors, the evidence suggests that “seeing is not believing” and highlights the need for longitudinal verification in ESG due diligence: a single visit may capture how a firm behaves when observed rather than how it operates on

average. More broadly, effective ESG monitoring should combine episodic on-site assessments with continuous, independently measured environmental data.

The remainder of the paper proceeds as follows. Section 2 describes the research background and develops the hypothesis. Section 3 details the data, variable construction, and empirical design. Section 4 presents the main results. Section 5 examines heterogeneity and mechanisms. Section 6 provides additional analyses to rule out alternative explanations, and Section 7 concludes.

2. Research background and hypothesis

2.1. Research background

2.1.1. Pollution regulation and strategic compliance

Effective environmental regulation is frequently constrained by limited monitoring resources. Traditional oversight typically relies on periodic inspections or stationary monitors that cover only a fraction of a jurisdiction. A substantial body of research in environmental economics demonstrates that regulated entities respond strategically to such intermittent and spatially limited monitoring. A seminal example is the United States' air quality monitoring under the Clean Air Act: Zou (2021) shows that local pollution readings are systematically lower on days when government monitors are scheduled to measure particulate matter, with air quality deteriorating on unmonitored days. Polluters, aware of the once-every-six-day schedule of EPA monitors, temporarily suppress emissions on monitoring days, effectively “shifting” pollution to periods when it is less likely to be recorded.

In China, environmental monitoring has historically faced measurement gaps and manipulation risks. Ghanem and Zhang (2014) provide evidence that, under the previous monitoring regime, Chinese cities misreported or selectively reported air quality data to inflate the number of “Blue Sky Days.” In response, China established an automated, continuous

monitoring network in the mid-2010s. However, local officials adapted strategically: Yang *et al.* (2024) document that following the introduction of real-time monitors, local governments concentrated pollution control efforts in the immediate vicinity of monitoring stations. This spatial reallocation resulted in significantly lower PM_{2.5} levels near stations relative to unmonitored areas, implying that officials deployed short-term measures—such as suppressing dust or restricting traffic—specifically around monitoring equipment. These localized interventions improved monitor readings without necessarily improving citywide air quality. Similarly, Grainger *et al.* (2019) find that U.S. local regulators, when exercising discretion, tend to site monitors in relatively cleaner locations to avoid pollution hotspots, particularly in areas with less federal oversight.

In summary, prior research across contexts reveals a pattern of strategic behavior in pollution monitoring: whether by timing emissions to off-monitor days, selectively placing monitors, or focusing abatement only where monitored, regulators and officials often find ways to game environmental monitoring to meet targets or avoid detection. However, evidence of this strategic behavior beyond formal regulatory oversight is limited. We address this gap by extending the framework to examine informal monitoring by capital market participants.

2.1.2. Remote sensing and pollution data advances

A primary challenge in detecting strategic environmental behavior is measurement, as pollution can be displaced across time or space to evade observation. Recent advances in satellite remote sensing have significantly enhanced the ability to observe pollution continuously and at fine spatial scales. For example, Currie *et al.* (2023) leverage satellite-derived PM_{2.5} concentration data at a 0.01×0.01 -degree resolution (from Di *et al.* (2016)) to measure individual exposure levels across the United States. By linking these granular measures with demographic and regulatory data, they show that the 1990 Clean Air Act amendments disproportionately benefited Black communities, accounting for over 60% of the

convergence in Black–White pollution exposure gaps. Such studies underscore the power of high-resolution data to uncover demographic patterns that would remain undetected with sparse ground monitors.

Satellite data have also been employed to identify monitoring evasion. Grainger *et al.* (2019) combine NASA Aerosol Optical Depth data with ground monitor readings to identify unmonitored areas with high pollution, providing evidence that some U.S. counties avoided placing monitors in the most polluted locations. Likewise, Zou (2021) uses satellite-based readings on non-monitored days to quantify the increase in smog when regulators were not “watching.” In China, nationwide satellite PM_{2.5} data have allowed researchers to validate ground station reports. Yang *et al.* (2024) employ a spatial difference-in-differences approach using satellite data to demonstrate that monitor-focused cleaning led to a 3.2% drop in particulate levels near monitors—an effect not mirrored citywide.

Our study utilizes the high-resolution CHAP dataset to detect short-duration, localized pollution changes at the firm level. Historically, data limitations precluded such analysis, as firm emissions were unobservable without adjacent official monitors. By leveraging granular satellite data, our work aligns with a broader trend in environmental economics that employs advanced measurement technologies (e.g., satellite sensors and dense monitor networks) to reveal previously hidden forms of pollution issues (e.g., Sullivan and Krupnick (2018) and Fowlie *et al.* (2019)). Crucially, the strategic behavior we document persists despite this increased transparency, underscoring the adaptability of firms in signaling compliance.

2.1.3. Investor information acquisition and corporate site visits.

A distinct strand of literature in finance and accounting examines how market participants gather private information. In-person meetings between investors and firms serve as a critical information channel (Brown *et al.*, 2015), yet data on such interactions have historically been unavailable. In 2012, the SZSE addressed this limitation by requiring listed companies to

disclose site visits on its "Hu Dong Yi" portal (<http://irm.cninfo.com.cn/>) within two trading days. This regulation provides a novel data source for observing such activity. Typically, these visits involve investors touring physical premises to gather operational insights. Although hosting visits is voluntary, firms are encouraged to accommodate these requests to enhance transparency with market participants.

Cheng *et al.* (2025), analyzing the early years of these disclosures, find that analysts who conduct site visits subsequently improve their earnings forecast accuracy relative to non-visiting peers. This suggests that site visits enable analysts to observe operational details and gather both “hard” and “soft” information, yielding an informational advantage.

Subsequent research documents additional implications. Bowen *et al.* (2018) report that corporate insiders conduct a significant portion of their informed stock trades after hosting private meetings, effectively profiting from the information revealed to select outsiders. Cheng *et al.* (2019) find that stock returns begin to drift upward prior to earnings announcements when a site visit has recently occurred, interpreting this as the capitalization of positive private signals. Similarly, Zhang *et al.* (2025) show that an abnormal frequency of site visits predicts superior stock returns, indicating that visits help investors identify mispriced firms. Furthermore, site visits are associated with improved corporate governance: more frequent visits correlate with lower earnings manipulation (Qi *et al.*, 2021) and improved investment efficiency (Cao *et al.*, 2025), suggesting that managers respond to the scrutiny of informed outsiders.

While this literature establishes site visits as valuable events for private information production, it largely treats the firm as a passive subject. Prior research has not fully considered that the firm might deliberately alter its behavior or environment in anticipation of a visit. We address this gap by integrating the site visit literature with the concept of strategic compliance. We posit that site visits are not merely information-revealing events but are endogenously influenced by firm actions. Specifically, we investigate whether the information acquired by

investors is shaped by the firm's strategic management of its environmental attributes.

2.2. Hypothesis

Building on the preceding discussion, we develop our main hypothesis regarding firm behavior around investor site visits.

Firms face a strategic tension regarding site visits. On one hand, hosting investors creates significant value: site visits are associated with enhanced transparency, which can improve firm valuation and lower the cost of capital (Cheng *et al.*, 2016; Zhang *et al.*, 2025). Consequently, managers have strong incentives to accommodate visit requests. On the other hand, the physical nature of these interactions exposes operations to direct observation, potentially revealing negative externalities that firms wish to conceal.

To resolve this conflict between the desire to host investors and the need to hide poor environmental performance, we theorize that firms will engage in temporary remediation. When a firm anticipates an investor visit, it will take short-term actions to improve visible environmental conditions—a form of physical “window dressing.” This behavior stems from an impression management or “signal jamming” motive: managers understand that visitors will scrutinize details otherwise hidden from the market, prompting proactive efforts to shape the firm's presentation.

Environmental performance is particularly visible during a facility tour. Smoke, dust, and odors provide immediate, tangible signals of pollution management quality. Because these environmental cues are salient to anyone walking through a plant, they are prime targets for manipulation. A manager expecting a visit has a strong incentive to temporarily curtail polluting activities and ensure environmental indicators appear favorable while investors are present. Accordingly, we formulate our central hypothesis:

Hypothesis 1: Firms strategically reduce their pollution emissions during periods when investors conduct on-site visits.

3. Methodology

3.1. Sample and data sources

Our study focuses on manufacturing firms listed on the Shenzhen Stock Exchange (SZSE) from January 1, 2012, to December 31, 2019.¹ We restrict our sample to the manufacturing sector for both physical and informational reasons. Physically, manufacturing operations represent significant sources of industrial emissions, providing a setting where strategic abatement is potentially detectable (Shapiro & Walker, 2018). Informationally, the value of site visits is contingent on the observability of operations. Cheng *et al.* (2016) document that the positive effect of site visits on analyst forecast accuracy is significantly stronger for firms in manufacturing industries, particularly those with higher asset tangibility. For these entities, direct observation of production lines, inventory levels, and factory conditions allows visitors to gather valuable "hard" information.

A single publicly listed firm typically operates multiple distinct subsidiaries, factories, or production facilities dispersed across different geographic locations. Our analysis is conducted at the site level, allowing us to exploit variation in pollution dynamics specific to each individual facility.

To construct our site-level dataset, we first identify the universe of subsidiaries and production plants associated with SZSE-listed manufacturing firms. Data on facility addresses and ownership structures are obtained from the *Controlled & Participated Company Database of Listed Companies* (CPCD) provided by the Chinese Research Data Services (CNRDS). We include plants where the listed firm holds at least 50% control. For each facility, we obtain the exact geographic coordinates (latitude and longitude).

¹ The sample period begins in 2012, coinciding with the implementation of the SZSE's disclosure requirement for site visits. We end the sample in 2019 to exclude the confounding effects of the COVID-19 pandemic and associated travel restrictions, consistent with recent literature (e.g., Zhang *et al.*, 2025).

We utilize the CHAP dataset to measure local air pollution. This dataset provides seamless, high-resolution ground-level PM_{2.5} concentrations (Wei *et al.*, 2021). We overlay the coordinates of our identified sites with these daily pollution grids to extract site-specific air quality measures.

We combine this spatial and pollution data with site visit records from the "Records of Investor Relations Activities" in the China Stock Market & Accounting Research (CSMAR) database. We identify site visits based on the "meeting method" field, manually reviewing records to confirm physical tours.² To link these visits to specific geographic locations, we review the meeting records for address details. If the disclosure specifies a particular facility or subsidiary location, we map the visit to that specific site. If the information provided is insufficient to identify a specific site location, we assume the visit occurred at the firm's headquarters. Merging these datasets with firm-level financial characteristics results in a final sample of 2,366 distinct geographic sites associated with 1,350 unique manufacturing firms. We restrict the sample to sites that received at least one visit during the sample period to ensure valid treatment variation. The distribution across manufacturing sub-industries is reported in Table IA1 of the Internet Appendix.³

Figure 1 illustrates the geographic distribution of sites in our sample alongside visit frequency. Consistent with the broader population of listed Chinese firms, sites are concentrated in coastal regions and provincial capitals. Visit intensity varies considerably: while the vast majority (approximately 95%) of sites received fewer than 50 visits, a small subset of highly visible locations hosted over 250 visits. The final panel comprises approximately 6.9 million (*i.e.*, 2,366 sites × 365 days × 8 years = 6,908,720) site-day observations covering 30,948 distinct site visit events, with each site receiving an average of

² We manually review each original record and remove conference calls, roadshows, and communications conducted by phone, video, or email, ensuring our analysis focuses exclusively on face-to-face site visits to firms' plants or headquarters.

³ The total frequency is 1,414—larger than the number of unique firms—because some firms change their industries during our sample period.

13 visits during the sample period.

[Insert Figure 1 about here]

3.2. Variable construction

3.2.1. Measuring local air pollution

The CHAP dataset is generated using machine learning algorithms that integrate ground-based measurements of pollutants, satellite remote sensing products, and atmospheric reanalysis data (Wei *et al.*, 2020; Wei *et al.*, 2021). This dataset provides daily 1 km ground-level PM_{2.5} data for China from 2000 to the present.⁴ The dataset exhibits high quality, with a cross-validation coefficient of determination ($CV - R^2$) of 0.90 on a daily basis (Wei *et al.*, 2021).

[Insert Figure 2 about here]

Using this high-resolution grid, we construct two site-specific pollution measures as visualized in Figure 2. The first measure, raw level of pollution ($Pollution_{i,j,t}$), represents the average ambient pollution exposure at the production facility. It is defined as the average PM_{2.5} concentration ($\mu\text{g}/\text{m}^3$) within a 1-kilometer radius (the red inner circle in Figure 2) centered on site i in city j on day t . The definition is provided in Appendix 1.

The second measure, abnormal pollution ($Pollution_{i,j,t}^{abn}$), is designed to isolate idiosyncratic emission changes attributable to the specific site from broader regional trends. We calculate the difference between the pollution level in the immediate vicinity of the site and the surrounding area. Specifically, this variable is defined as the difference between the average PM_{2.5} concentration within the 1km inner circle ($Pollution_{i,j,t}$) and the average concentration in the surrounding 1km-to-3km annular ring ($Pollution_{i,j,t}^{ring}$) (the blue area in

⁴ Satellite-based pollution measures in prior studies are typically coarser. For example, Currie *et al.* (2023) use PM_{2.5} concentrations at a 0.01-degree by 0.01-degree resolution to study the Black-White gap in pollution exposure, while Zou (2021) uses aerosol concentrations at a 10km×10km resolution to examine the impact of monitoring on pollution emissions.

Figure 2). This spatial difference-in-differences approach effectively nets out background pollution—such as city-wide smog or weather-driven variations—allowing us to isolate idiosyncratic emission changes attributable to the specific site.

$$Pollution_{i,j,t}^{abn} = Pollution_{i,j,t} - Pollution_{i,j,t}^{ring} \quad (1)$$

3.2.2. Identifying site visit events

Site visits are a prevalent channel for information acquisition in which investors, analysts, and fund managers visit a company's physical premises. A typical visit comprises a management briefing and a Q&A session, followed by a tour of operational facilities such as factories, warehouses, or project sites. This component of direct observation distinguishes site visits from remote forms of corporate communication, such as conference calls or investor meetings. Underscoring their importance, recent research indicates that China's top mutual fund firms allocate 40-50% of their working hours to these visits (Zhang *et al.*, 2025).

The "Records of Investor Relations Activities" from the CSMAR database provides detailed information for each site visit, including the date, participants, firm management who participated, and the meeting minutes. We utilize this granular data to construct our primary independent variables. We define a binary variable, $SiteVisit_{i,j,t}$, which equals one if site i receives a visit on day t , and zero otherwise. To capture the potential anticipatory or lagged effects of the visit, we also construct an indicator $SiteVisit_{i,j,t}^{3day}$, which equals one if site i receives a visit within a 3-day window (i.e., day t , day $t-1$, or day $t+1$). This window accounts for the immediate period surrounding the event where operational adjustments might occur.

Appendix 2 details the distribution of visitor types across our 30,948 visits. The participants are diverse, comprising securities firms (e.g., sell-side financial analysts), funds, asset management companies, foreign institutions (e.g., QFII), insurance companies, banks, investment advisors, media, and individual investors. A single visit often involves multiple visitor types simultaneously (e.g., a joint visit by analysts and fund managers). Securities firms

participate in 80% of visits, followed by funds (49%) and asset management companies (42%).

3.3. Empirical model

3.3.1. Baseline model

To estimate the impact of site visits on local pollution, we employ the following specifications:

$$Y_{i,j,t} = \alpha + \beta \times VisitEvent_{i,j,t} + \gamma \times X_{i,j,t} + \eta_i + \mu_{j,t} + \varepsilon_{i,j,t} \quad (2)$$

where the dependent variable $Y_{i,j,t}$ represents either the raw pollution level ($Pollution$) or the abnormal pollution level ($Pollution^{abn}$). $VisitEvent_{i,j,t}$ represents the variable of interest, which can be either the single-day indicator $SiteVisit_{i,j,t}$ or the three-day window indicator $SiteVisit_{i,j,t}^{3d}$, depending on the specification.

We include a vector of controls, $X_{i,j,t}$, to account for site-specific characteristics and localized environmental conditions. Specifically, we control for $Pollution_{i,j,t}^{lag}$, the average pollution level at site i over the window $[t-14, t-7]$. This variable captures the persistence of site-specific emission levels prior to the immediate event window, accounting for operational continuity and previous environmental performance. Additionally, we include for $Pollution_{i,j,t}^{ring}$, the concurrent pollution level in the 1km-3km buffer zone surrounding the site. Controlling for the pollution in this ring area is crucial as it represents the background pollution level in the immediate vicinity, driven by broader regional factors such as weather patterns or city-wide emissions.

We further include site fixed effects (η_i) to absorb time-invariant site characteristics (e.g., industry type, production capacity, geography) and city-by-date fixed effects ($\mu_{j,t}$) to absorb time-varying shocks common to a city on a given day, such as meteorological conditions (wind, precipitation, temperature) and city-wide pollution trends. Consequently, the coefficient β identifies the idiosyncratic pollution variation at a specific site during a visit, relative to the

same site on days without visits and to other sites in the same city on the same day. Standard errors are clustered at the site level to account for serial correlation in errors within sites over time.

3.3.2. Dynamic model

To examine the temporal dynamics of pollution and test for pre-existing trends, we estimate the following specification:

$$Y_{i,j,t} = \alpha + \sum_{n=-7}^7 \beta_n \times VisitWindow_{i,j,t+n} + \gamma \times X_{i,j,t} + \eta_i + \mu_{j,t} + \varepsilon_{i,j,t} \quad (3)$$

where $VisitWindow_{i,j,t+n}$ is an indicator variable equal to one if site i is visited on day $t + n$. We estimate coefficients for a two-week window ranging from 7 days before ($n = -7$) to 7 days after ($n = 7$) the visit. The points outside the 14-day window serve as the reference period. This specification allows us to verify whether pollution reduction is concentrated around the visit dates and ensures that the results are not driven by pre-existing downward trends in emissions.

3.4. Summary statistics

Table IA2 presents the summary statistics for the main variables used in the study. The average PM2.5 concentration ($Pollution$) is $49.57 \mu\text{g}/\text{m}^3$, with the surrounding buffer ring ($Pollution^{ring}$) showing a slightly lower average of $49.48 \mu\text{g}/\text{m}^3$. The average abnormal pollution ($Pollution^{abn}$) is $0.093 \mu\text{g}/\text{m}^3$. The mean of $SiteVisit$ is 0.004, meaning that there are 4 observations of sites for every 1,000 site-daily observation points. It increases to 13 if the site event is defined in a window of 3 days (the mean of $SiteVisit^{3day}$ is 0.013).

4. Main results

4.1. Baseline model results

Table 1 presents the estimates of Equation (2). Columns (1) and (2) report the results using abnormal pollution ($Pollution^{abn}$) as the dependent variable, which effectively nets out common regional trends to isolate site-specific emission dynamics. In Column (1), the coefficient on the

single-day indicator *SiteVisit* is -0.015 and is statistically significant at the 5% level. This magnitude corresponds to a 16.1% reduction relative to the sample mean ($0.093 \mu\text{g}/\text{m}^3$). To capture operational adjustments surrounding the event, Column (2) employs the three-day window indicator, *SiteVisit*^{3day}. The estimated coefficient is -0.016 (statistically significant at the 1% level), representing a 17.2% reduction relative to the mean. These findings suggest that firms actively suppress emissions during periods of external monitoring, consistent with the hypothesis of strategic pollution suppression.

Columns (3) and (4) verify these findings using raw pollution levels (*Pollution*) as the outcome variable. The coefficients remain negative and statistically significant, lending further support to the robustness of the main result. Specifically, the coefficient on *SiteVisit* in Column (3) is -0.015, statistically significant at the 10% level. Similarly, the coefficient on *SiteVisit*^{3day} in Column (4) is -0.017, significant at the 1% level. The estimates are quantitatively similar to those based on abnormal pollution, indicating that our findings are robust to the choice of pollution measures. The control variables behave as expected. Lagged pollution (*Pollution*^{lag}) and pollution in the surrounding buffer ring (*Pollution*^{ring}) are highly predictive of site-level pollution, underscoring the importance of the temporal and spatial difference-in-differences design in absorbing baseline emissions and regional environmental shocks. Collectively, these results provide evidence that manufacturing firms strategically curb pollution emissions during periods of investor oversight.

[Insert Table 1 about here]

4.2. Dynamic model results

To further scrutinize the causal interpretation of our findings and map the temporal dynamics of this behavior, we estimate the dynamic model specified in Equation (3). Figure 3 plots the point estimates and 95% confidence intervals for the coefficients on daily indicators ranging from seven days before to seven days after the site visit. Panel A, which displays the

results for abnormal pollution, reveals two critical patterns. First, the estimated coefficients for the days leading up to the visit (days $t-7$ to $t-3$) are statistically indistinguishable from zero and exhibit no systematic trend. This absence of pre-existing trends supports the parallel trends assumption underlying our identification strategy, suggesting that the observed pollution reduction is not driven by unobserved factors or secular improvements in environmental performance.

Second, there is a sharp and statistically significant decline in pollution beginning two days prior to the visit ($t-2$) and persisting through the day after the visit ($t+1$). The effect magnitude peaks within this window before reverting to baseline levels by day $t+2$. The results based on raw pollution, as shown in Panel B, are qualitatively similar. This "V-shaped" pattern—characterized by a transient dip followed by a quick rebound—is distinct from the permanent pollution reductions typically associated with technological upgrades or persistent regulatory enforcement (Lewis, 2023; Cheng *et al.*, 2025). Instead, it aligns with the "unwatched pollution" phenomenon described in recent literature (Zou, 2021; Karplus & Wu, 2023), where firms strategically ramp down polluting activities when under the direct spotlight of external monitors.

[Insert Figure 3 about here]

4.3. Robustness

We conduct several tests to verify the robustness of our baseline findings, addressing potential concerns regarding chance correlations, visit intensity, selection bias, and sample composition.

4.3.1. Falsification tests

To rule out the possibility that our results are driven by chance, we conduct a falsification test using randomization inference. We generate 1,000 counterfactual samples by randomly permuting the incidence of corporate site visits (*SiteVisit*) across the panel while preserving the

original sample structure. We then re-estimate the baseline model for each placebo sample to obtain the distribution of the spurious coefficients and t-statistics. Figure 4 presents the results, plotting the kernel density estimates of the 1,000 placebo coefficients in Panel A and their corresponding t-values in Panel B. The placebo estimates are centered around zero and follow a normal distribution, indicating that randomly assigned visits do not systematically predict pollution fluctuations. In contrast, the actual estimated coefficient and t-value from our baseline model—marked by the vertical red lines—lie in the extreme left tails of their respective distributions, clearly separated from the placebo estimates. This divergence implies that the likelihood of observing our main results by chance is small.

[Insert Figure 4 about here]

4.3.2. *Visit intensity*

We next examine whether the magnitude of pollution reduction varies with the intensity of the site visit. If the observed pollution suppression is a strategic response to external scrutiny, we expect stronger effects when monitoring intensity is higher. Table 2 reports the results using three alternative measures of visit intensity: the length of the disclosure record ($\text{Ln}(\# \text{ Word})$), the number of visitors ($\text{Ln}(\# \text{ Researcher})$), and the number of visiting institutions ($\text{Ln}(\# \text{ Institution})$). Across all specifications, the coefficients on these intensity measures are negative and statistically significant. For instance, in Columns (3) and (4), the coefficients on $\text{Ln}(\# \text{ Researcher})$ are -0.008 (significant at the 5% level) and -0.007 (significant at the 10% level), respectively, indicating that a larger delegation of visitors is associated with a more pronounced reduction in local air pollution. These findings suggest that firms scale their emission abatement efforts in proportion to the intensity of external monitoring.

[Insert Table 2 about here]

4.3.3. *Selection on unobservables*

To address the concern that our results may be driven by unobserved factors, we employ

the method developed by Oster (2019) to evaluate the potential bias from omitted variables. We calculate the Oster delta statistic, which compares the stability of the coefficient on *SiteVisit* as controls and fixed effects are included. A delta value greater than one (or less than zero) suggests that unobservables do not explain away the estimated treatment effect. Our results in Table IA3 support this conclusion. The calculated delta for the abnormal pollution measure (*Pollution^{abn}*) is 1.11 for the single-day indicator and 1.63 for the three-day window. For the absolute pollution measure (*Pollution*), the delta values are negative (-35.45 and -22.35). These statistics indicate that selection of unobservables seem not to be driver of our main results.

4.3.4. Entropy balancing

To mitigate concerns regarding covariate imbalance between visited and non-visited site-days, we employ entropy balancing, a reweighting technique that matches the covariate distributions of the treatment and control groups on possibly higher moments (Hainmueller & Xu, 2013). This method minimizes model dependence and provides a robust estimate of the treatment effect. Table IA4 presents the results using the reweighted sample. The coefficients on *SiteVisit* and *SiteVisit^{3day}* remain negative and statistically significant across both abnormal and raw pollution measures. For example, the coefficient on *SiteVisit^{3day}* in Column (2) is -0.014, quantitatively similar to the baseline estimate of -0.016. These results confirm that our findings are not driven by distributional differences in observable characteristics between the treatment and control groups.

4.3.5. Geographic variation

Finally, we test the sensitivity of our results to various subsample restrictions to ensure our findings are not driven by specific regions with unique regulatory or economic characteristics. Table IA5 reports these robustness checks. Panel A excludes sites located in provincial capitals and municipalities, where regulatory enforcement might be stricter. Panel B excludes southern provinces, and Panel C excludes coastal provinces, which typically exhibit

higher economic development and environmental awareness. In all subsamples, the coefficients on *SiteVisit* and *SiteVisit*^{3day} remain negative and generally significant. This evidence suggests that the phenomenon we document is widespread across China's manufacturing sector and is not confined to specific geographic or economic zones.

5. Heterogeneity and mechanism

To better understand the drivers of the observed strategic abatement, we examine the underlying mechanisms through the lens of industry characteristics, firm valuation, ESG performance, and the specific attributes of the visiting investors.

5.1. Industry pollution intensity

We first examine whether the propensity to manage pollution varies with the firm's industrial environment. Firms in high-pollution industries typically face higher costs—both regulatory and operational—when their emissions are exposed to the public (Buntaine *et al.*, 2024). Because their operations are inherently emission-intensive, they are the primary targets of public outcry and regulatory enforcement when local air quality deteriorates. For instance, when a smog cloud hangs over a region, a coal-power utility is far more likely to face blame and scrutiny than a catering company. Consequently, these firms face heightened incentives to demonstrate "clean" operations during site visits to signal compliance and avoid the severe penalties associated with perceived negligence.

To test this, we classify industries according to the environmental inspection guidelines released by the Ministry of Ecology and Environment of China.⁵ We define "High-pollution industries" to include sectors such as thermal power, iron and steel, cement, electrolytic aluminum, coal, metallurgy, chemicals, petrochemicals, building materials, paper, brewing, pharmaceuticals, fermentation, textiles, tanning, and mining. All other sectors are classified as

⁵ Ministry of Ecology and Environment, "Classified Management Directory of Environmental Protection Verification for Listed Companies," https://www.mee.gov.cn/gkml/hbb/bgth/200910/t20091022_174891.htm.

"Low-pollution industries." We investigate the heterogeneity across these sectors by re-estimating the baseline Equation (2) separately for the high-pollution and low-pollution subsamples, allowing us to statistically compare the visit coefficients across groups.

Table 3 reports the results of this heterogeneity analysis. Consistent with our prediction, the strategic reduction in pollution is driven primarily by firms in heavy-polluting sectors. The coefficient on $SiteVisit^{3day}$ for high-pollution industries is -0.024 when the dependent variable is abnormal pollution ($Pollution^{abn}$), and is significant at the 5% level. This estimate is nearly double the magnitude of the coefficient for low-pollution industries (-0.013). The difference between the coefficients is statistically significant (p-value = 0.070). We observe a similar pattern when using raw pollution ($Pollution$) as the outcome variable. This evidence suggests that strategic abatement is not a uniform practice but is concentrated in sectors where environmental risks are most material to investors.

[Insert Table 3 about here]

5.2. Firm valuation

Next, we explore how firm valuation influences the motivation for strategic pollution control. Firms with high valuations—often characterized by significant growth opportunities—typically rely more heavily on external financing and are more sensitive to investor sentiment (Baker & Wurgler, 2002). Consequently, the marginal benefit of influencing investor perceptions during a site visit should be higher for these firms. We use *Tobin's Q* as our proxy for firm valuation, calculated as the ratio of the market value of equity plus the book value of liabilities to the book value of assets. We split the sample into "High *Tobin's Q*" and "Low *Tobin's Q*" groups based on the industry median and re-estimate the baseline model for each subsample to test whether market valuation conditions the strategic response to monitoring.

Table 4 tests this prediction using these subsamples. The results reveal a stark dichotomy: the strategic pollution reduction is driven entirely by high-valuation firms. When examining

abnormal pollution ($Pollution^{abn}$), the coefficient for the high *Tobin's Q* subsample is -0.027 and significant at the 1% level, whereas the coefficient for low-valuation firms is virtually zero. The difference between these groups is statistically significant at the 1% level (p-value < 0.0000). A similar pattern is observed for the raw pollution measure ($Pollution$), where the high valuation group shows a significant reduction (-0.030) compared to a null result for the low valuation group. This finding implies that the pressure to maintain market confidence and support high valuations acts as a powerful catalyst for strategic behavior under the spotlight of external monitors.

[Insert Table 4 about here]

5.3. ESG performance

We further investigate how firms' strategic pollution reduction varies with their ESG performance. Incentives to manipulate pollution levels during site visits likely differ across the environmental, social, and governance dimensions. For the Environmental (E) component, we predict that firms with poor records (low E ratings) will be more aggressive in cutting emissions. Analogous to firms in high-pollution industries, these entities face higher legitimacy risks and are under greater pressure to mask their poor baseline performance to mitigate regulatory penalties or investor backlash. The prediction for the Social (S) component is less clear; while low social responsibility may correlate with higher societal costs for negative externalities, the direct link to immediate pollution control is less distinct than for environmental factors.

In contrast, for the Governance (G) component, we predict that firms with strong governance (high G ratings) will exhibit stronger strategic responses. High governance implies strict internal controls and rigorous shareholder monitoring. Managers in these firms are subject to greater discipline and are acutely aware that allowing visitors to witness pollution could severely damage the firm's reputation and valuation. Consequently, these managers are likely to be more "jittery" about upcoming visits and possess the managerial capacity to execute

preemptive cleanups to protect their careers and shareholder value.

To test these predictions, we utilize the ESG ratings provided by the Sino-Securities Index, which scores firms on a scale of 1 to 9 for both composite and sub-scores, with a higher rating indicating better performance.⁶ We classify firms as having a "High Rating" if their score is greater than 4, and a "Low Rating" if their score is equal to or less than 4. By decomposing the composite ESG score into its three specific components (E, S, and G), we can disentangle these competing mechanisms and better understand what drives the strategic behavior.

Table 5 first presents the results for the composite scores. We find that the pollution reduction effect is statistically significant only among firms with low overall ESG ratings (-0.019 for *Pollution^{abn}*), whereas the estimate for high-ESG firms is statistically insignificant.

[Insert Table 5 about here]

Table 6 reveals the nuances behind this aggregate result. In Panel A, consistent with our prediction regarding risk mitigation, the results are driven by firms with Low E ratings. This finding mirrors the results in Section 5.1: entities with the highest "baseline dirtiness" and liability risk are the most aggressive in cutting emissions. In Panel B, we observe that while firms with Low S ratings show a significant reduction in pollution, the difference between the Low and High S groups is not statistically significant. This suggests that the Social dimension is less distinctive in driving this specific form of strategic behavior compared to Environmental and Governance factors.

Crucially, Panel C shows that firms with High G ratings exhibit significantly stronger responsiveness than their Low G counterparts. This aligns with the "jittery manager" hypothesis and the findings in Section 5.2 regarding firm valuation. Just as high-Tobin's Q firms have the market incentive to manage perceptions, high-governance firms possess the managerial capacity and disciplinary incentives to execute these strategic interventions

⁶ See <https://www.unpri.org/>.

effectively.

To verify this governance channel, we conduct an additional test using conventional measures of internal and external governance. We gauge internal governance using managerial compensation (*ExePay*), defined as the logarithm of one plus the total compensation of the top three executives. We gauge external governance using institutional ownership (*IO*), calculated as the number of shares held by institutional investors divided by the total number of shares outstanding. Low executive compensation or high institutional ownership indicate lower levels of agency costs and stronger governance (Core *et al.*, 1999; Appel *et al.*, 2016; Wang *et al.*, 2025).

As reported in Table IA6, the results corroborate our governance findings: the strategic reduction in pollution is driven by firms with stronger governance structures. Regarding internal governance (Panel A), the coefficient on *SiteVisit*^{3day} is negative and statistically significant for the Low *ExePay* subsample (−0.019), whereas it is insignificant for the High *ExePay* subsample (−0.011). Similarly, for external governance (Panel B), the effect is concentrated in the High *IO* subsample, which exhibits a significant coefficient of −0.019, compared to a small and insignificant coefficient (−0.008) for the Low *IO* group. The difference between the subsamples are statistically significant.

Collectively, these findings imply a dual mechanism: the *necessity* to reduce pollution is driven by environmental liability (Low E), while the *ability* and *discipline* to do so are facilitated by strong governance (High G).

[Insert Table 6 about here]

5.4. Visitor ownership stake and PRI signatory

The results thus far indicate that firms' strategic behaviors are driven by environmental risk, valuation, and governance. However, a key question remains: how do these factors transmit to specific firm behaviors? One possibility is a *reputation* channel, whereby firms are

sensitive to scrutiny from environmentally conscious visitors who may expose pollution to more audience, thereby increasing regulatory risk or the cost of capital. A second possibility is an *ownership* channel, where firms respond to visitors holding significant equity stakes. Upon observing pollution, these investors may sell their shares, either due to concerns over regulatory risk or to comply with ESG investment mandates. To distinguish between these mechanisms, we investigate whether the observed pollution reduction is driven by visitors with an ownership stake or by non-shareholder signatories to the Principles for Responsible Investment (PRI).⁷ If the cleanup is primarily motivated by the reputation channel, we would expect PRI signatories to elicit a significant response. Conversely, if the ownership channel is the ultimate driver, strategic pollution reduction should be concentrated among visitors with existing equity stakes.

To systematically identify these visitor types, we construct specific measures for shareholder status and PRI signatory status. We leverage the mandatory disclosure of top ten shareholders under securities regulations and obtain the data from the CSMAR dataset. We match these shareholder names to visitor information using a fuzzy matching algorithm.⁸ Based on these matches, we define *Stake* as a binary indicator equal to one if the visiting group includes at least one shareholder, and *nonStake* as one if it does not. We also construct continuous variables, *Stake_Num* and *nonStake_Num*, to capture the number of shareholders and non-shareholders involved in each visit, respectively. To capture the presence of environmentally conscious investors, we obtain the list of institutional PRI signatories from the United Nations PRI. We employ the same fuzzy matching algorithm to link PRI signatories with the visitor records. We define *PRI* as an indicator equal to one if the visiting group includes a PRI signatory, and *nonPRI* as one if it does not. Similarly, *PRI_Num* and *nonPRI_Num*

⁷ PRI is a United Nations-supported international network of investors working to understand the investment implications of environmental, social, and governance (ESG) factors. Signatories commit to six aspirational principles that encourage incorporating ESG issues into investment analysis and decision-making processes (see <https://www.unpri.org/>).

⁸ We employ the *rapidfuzz* package with a 70% similarity threshold to account for data inconsistencies, such as abbreviations or typographical errors.

measure the respective counts of PRI signatories and non-signatories. To conduct the test, we modify the baseline equation by substituting the single *SiteVisit* variable with the corresponding pair of mutually exclusive visitor indicators (e.g., *Stake* and *nonStake*) or their count equivalents.

[Insert Table 7 about here]

We first test the ownership channel by distinguishing between visits from shareholders and non-shareholders. The results are shown in Panel A of Table 7. The estimates indicate that the strategic reduction in pollution is concentrated during visits by existing shareholders. The coefficient for *Stake* visitors is negative and statistically significant at the 5% level (-0.040) when the dependent variable is abnormal pollution ($Pollution^{abn}$), representing a substantial deviation from the baseline. In contrast, the coefficient for non-stakeholders (*nonStake*) is smaller in magnitude (-0.010) and statistically indistinguishable from zero. The difference between these two groups is statistically significant (p-value = 0.080). When using visitor counts, the pattern persists: the coefficient for *Stake_Num* is significantly negative, whereas the coefficient for *nonStake_Num* shows no such effect. We observe qualitatively similar results when using raw pollution (*Pollution*) as the outcome variable. These findings suggest that managers perceive a higher cost of revealing "dirty" operations to visitors who hold voting power and cash flow rights.

Next, we test the reputation channel by investigating whether the "green" label of visitors—specifically their status as PRI signatories—induces firms to strategically reduce pollution. Panel B shows no evidence that PRI signatories elicit a stronger pollution reduction response than non-signatories. The coefficient for *PRI* visitors is statistically insignificant when examining abnormal pollution ($Pollution^{abn}$), while the coefficient for *nonPRI* visitors is negative and marginally significant. Crucially, the difference between the two coefficients is statistically insignificant (p-value = 0.704), indicating that a visitor's status as a responsible

investor does not, in isolation, trigger a differential strategic response from the firm. Similar patterns are observed for the raw pollution outcome (*Pollution*).

Finally, we conduct a "horse race" analysis between the two effects in Panel C. This specification includes both the *Stake* and *PRI* indicators (and their count equivalents) simultaneously. When both factors are controlled for, the dominance of the ownership channel becomes even more pronounced. In the specifications of *Pollution^{abn}*, the coefficient on *Stake* increases in magnitude to -0.046 and remains significant at the 5% level, whereas the coefficient on *PRI* turns positive and insignificant. The difference between the shareholder and non-shareholder coefficients remains notable (p-value = 0.074). The pattern is similar in the specifications of *Pollution*. Collectively, these results imply that the "unwatched pollution" phenomenon is primarily a response to the monitoring power derived from ownership rather than the stated environmental preferences of the visitors. Managers appear to strategically clean up for those who own the firm, regardless of whether those owners have explicitly pledged to uphold ESG principles.

In conclusion, these findings elucidate the underlying mechanism: The strategic reduction of pollution is not primarily driven by a reputation channel associated with environmentally conscious visitors. Instead, it appears to be an ownership-driven response. Managers in pollution-intensive or environmentally poor-performing firms (Sections 5.1 and 5.3) who oversee high-valuation assets (Section 5.2) and operate under strict governance regimes (Section 5.3) are acutely sensitive to the scrutiny of major shareholders (Section 5.4). This sensitivity leads them to temporarily sanitize operations, a tactic designed to protect the firm's market value and their own professional standing rather than to achieve genuine environmental sustainability.

6. Additional analysis

We conduct additional analysis to validate our findings and exclude alternative explanations.

6.1. The effect of online visits

To further validate that the observed pollution reduction is driven by the physical presence of external monitors, we employ online site visits as a placebo test. Unlike physical visits, online interactions (such as teleconferences or web-based Q&A sessions) do not involve direct observation of production facilities and thus should not trigger the same immediate "cleanup" incentives. If our argument regarding physical monitoring is correct, we should observe significant pollution abatement only during physical site visits, but not during online interactions.

Figure IA1 presents the dynamic coefficients for online visits, estimated using the same specification as our baseline dynamic model. Consistent with our hypothesis, and in sharp contrast to the distinct "V-shaped" reduction observed during physical visits, the coefficients for online visits are statistically indistinguishable from zero throughout the entire event window. This null result confirms that the "unwatched pollution" phenomenon is uniquely tied to the physical presence of external monitors. It reinforces the mechanism of strategic local abatement, suggesting that managers are motivated to clean up specifically when there is a direct risk of visitors observing "dirty" operations on-site.

6.2. The effect on sister non-visited sites

A potential concern with our identification strategy is that the observed pollution reduction might be driven by unobserved firm-level factors, such as a company-wide initiative to improve ESG performance or other simultaneous shocks affecting the entire organization. To address this, we conduct a placebo test by comparing the pollution dynamics of visited sites to those of non-visited sites within the same firm. If our hypothesis regarding physical monitoring is correct, the strategic abatement should be strictly localized to the specific facility hosting the visitors. Non-visited sites, lacking the immediate pressure of external scrutiny,

should not exhibit similar pollution reductions.

To implement this test, we construct two indicators for non-visited sites belonging to the same firm as the visited site. We define *SameCity* as an indicator equal to one for non-visited sister sites located in the same city as the visited facility. This variable is particularly useful for ruling out coordinated regional responses or city-specific shocks that might affect multiple facilities of the same firm simultaneously. We define *DiffCity* as an indicator equal to one for non-visited sister sites located in different cities. This variable allows us to test for firm-wide policy changes or aggregate shocks that might prompt simultaneous abatement across all production units, regardless of their geographic location.

Table 8 reports the results of this analysis. Panel A presents the estimates for *SameCity* (and its three-day window equivalent, *SameCity*^{3day}), while Panel B presents the estimates for *DiffCity* (and its three-day window equivalent, *DiffCity*^{3day}). Across all specifications, the coefficients on these placebo visit indicators are small in magnitude and statistically insignificant. For instance, the coefficient on *SameCity*^{3day} is 0.005 (p-value > 0.10), indicating no detectable change in pollution at sister sites, even when they are geographically close to the visited location. Similarly, sites in different regions show no significant response. These null results reinforce the conclusion that the pollution reduction is not a general firm-level phenomenon but is uniquely triggered by the physical presence of monitors at the specific site.

[Insert Table 8 about here]

6.3. Changes in operational activities

Finally, we address the question of *how* firms achieve the observed reduction in pollution. A reduction in emissions could result from either a temporary suspension of production activities or the strategic deployment of pollution control technologies (e.g., turning on scrubbers or filters that are otherwise left idle). Distinguishing between these effects is crucial for understanding the channel of "greenwashing." If firms merely halt production, the pollution

drop is a mechanical byproduct of inactivity. However, if production remains constant while pollution falls, it points to the strategic use of abatement technology to mask true emission intensities.

To test this, we utilize satellite-derived thermal infrared radiation (TIR) data as a proxy for real-time industrial activity, following the methodology of Xue *et al.* (2025). Manufacturing processes generate significant heat, making TIR an effective measure of operating intensity that is distinct from pollution output. We define *TIR* as the average daytime thermal infrared radiation value (measured in Kelvin) in the plant area of the specific site on day t . To account for local thermal variations, we also construct an abnormal measure, *AR*, defined as the difference between the site's *TIR* and the average *TIR* of a surrounding 5km buffer zone on day t .

We estimate the impact of site visits on these operational proxies using a specification analogous to our baseline model. Specifically, we regress *TIR* and *AR* on our visit indicators (*SiteVisit* or *SiteVisit*^{3day}), controlling for the site's physical area size (*SiteAreaSize*), lagged operational activity (i.e., *TIR* or *AR* calculated over the [t-14, t-7] window), as well as site and city-date fixed effects. Table 9 presents the results of this analysis. The coefficients on *SiteVisit* and *SiteVisit*^{3day} are statistically insignificant across all columns, indicating that site visits do not trigger a significant reduction in operational heat signatures. Figure IA2 further illustrates that there are no discernible trends in operational activity around the visit window.

These results imply that firms do not achieve cleaner air by shutting down factories. This is reasonable, as a halt or reduction in operations would signal inactivity to visitors, potentially raising concerns about the firm's operational health. Instead, firms appear to decouple emissions from production, likely by temporarily ramping up the usage of end-of-pipe abatement facilities during the visit. This behavior is consistent with a sophisticated form of strategic compliance, where managers ensure that the visible environmental output (air quality)

improves while the underlying economic activity remains unaffected.

[Insert Table 9 about here]

7. Conclusion

This study provides the first firm-level evidence that corporate site visits—a key channel for information acquisition—induce strategic changes in real environmental behavior. We document that manufacturing firms temporarily suppress pollution emissions when investors are present, creating a favorable but transient "green" impression. The effects are most pronounced among firms in heavy-polluting industries, where environmental risks are salient, as well as among firms with high market valuations and strong governance. Furthermore, we find that the effect is driven by visits from existing shareholders rather than by the environmental mandates of visitors (e.g., PRI signatories). This suggests that managers clean up to protect firm value and their own standing before owners, rather than to signal social responsibility.

These findings have significant implications for ESG investing and policy. They demonstrate that the act of observing the firm changes its physical behavior, rendering episodic on-site diligence potentially misleading. Consequently, capital allocators relying on site visits may overestimate a firm's consistent environmental performance. For effective stewardship, investors and regulators must look beyond the curated windows of site visits and integrate continuous, independent monitoring technologies—such as satellite remote sensing and real-time sensors—to capture the true state of corporate externalities.

References

- Appel, I.R., Gormley, T.A., Keim, D.B., 2016. Passive investors, not passive owners. *Journal of Financial Economics* 121, 111-141
- Baker, M., Wurgler, J., 2002. Market Timing and Capital Structure. *The Journal of Finance* 57, 1-32
- Berg, F., Kölbel, J.F., Rigobon, R., 2022. Aggregate confusion: The divergence of ESG ratings. *Review of Finance* 26, 1315-1344
- Bowen, R.M., Dutta, S., Tang, S., Zhu, P., 2018. Inside the “black box” of private in-house meetings. *Review of Accounting Studies* 23, 487-527
- Bradley, D., Jame, R., Williams, J., 2022. Non-deal roadshows, informed trading, and analyst conflicts of interest. *The Journal of Finance* 77, 265-315
- Brav, A., Jiang, W., Partnoy, F., Thomas, R., 2008. Hedge fund activism, corporate governance, and firm performance. *The Journal of Finance* 63, 1729-1775
- Brown, L.D., Call, A.C., Clement, M.B., Sharp, N.Y., 2015. Inside the “black box” of sell-side financial analysts. *Journal of Accounting Research* 53, 1-47
- 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, 815-850
- Bushee, B.J., Gerakos, J.J., Lee, L.F., 2018. Corporate jets and private meetings with investors. *Journal of Accounting and Economics* 65, 358-379
- Cao, S.S., Gong, G., Kim, Y., Shi, H., Wang, A., 2025. Site visits and corporate investment efficiency. *Management Science* 71, 635-658
- Cheng, Q., Du, F., Wang, X., Wang, Y., 2016. Seeing is believing: Analysts’ corporate site visits. *Review of Accounting Studies* 21, 1245-1286
- Cheng, Q., Du, F., Wang, X., Wang, Y., 2019. Do corporate site visits impact stock prices? *Contemporary Accounting Research* 36, 359-388
- Cheng, Q., Lin, A.-P., Yang, M., 2025. Green innovation and firms’ financial and environmental performance: The roles of pollution prevention versus control. *Journal of Accounting and Economics* 79, 101706
- Christensen, D.M., Serafeim, G., Sikochi, A., 2022. Why is corporate virtue in the eye of the beholder? The case of ESG ratings. *The Accounting Review* 97, 147-175
- Cohen, D.A., Zarowin, P., 2010. Accrual-based and real earnings management activities around seasoned equity offerings. *Journal of accounting and Economics* 50, 2-19
- Core, J.E., Holthausen, R.W., Larcker, D.F., 1999. Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics* 51, 371-406
- Currie, J., Voorheis, J., Walker, R., 2023. What caused racial disparities in particulate exposure to fall? New evidence from the Clean Air Act and satellite-based measures of air quality. *American Economic Review* 113, 71-97
- Di, Q., Kloog, I., Koutrakis, P., Lyapustin, A., Wang, Y., Schwartz, J., 2016. Assessing PM_{2.5} exposures with high spatiotemporal resolution across the continental United States. *Environmental Science & Technology* 50, 4712-4721
- Fowlie, M., Rubin, E., Walker, R., 2019. Bringing satellite-based air quality estimates down to earth. In: *AEA Papers and Proceedings*, pp. 283-288. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203
- Ghanem, D., Zhang, J., 2014. ‘Effortless Perfection:’ Do Chinese cities manipulate air pollution data? *Journal of Environmental Economics and Management* 68, 203-225
- Grainger, C., Schreiber, A., Chang, W., 2019. Do regulators strategically avoid pollution hotspots when siting monitors? Evidence from remote sensing of air pollution.

- Department of Economics, University of Wisconsin–Madison Working Paper
- Gunny, K.A., 2010. The relation between earnings management using real activities manipulation and future performance: Evidence from meeting earnings benchmarks. *Contemporary Accounting Research* 27, 855-888
- Hainmueller, J., Xu, Y., 2013. Ebalance: A Stata package for entropy balancing. *Journal of Statistical Software* 54, 1-18
- Han, B., Kong, D., Liu, S., 2018. Do analysts gain an informational advantage by visiting listed companies? *Contemporary Accounting Research* 35, 1843-1867
- Karplus, V.J., Wu, M., 2023. Dynamic responses of SO₂ pollution to China's environmental inspections. *Proceedings of the National Academy of Sciences* 120, e2214262120
- Kirk, M.P., Markov, S., 2016. Come on over: Analyst/investor days as a disclosure medium. *The Accounting Review* 91, 1725-1750
- Levit, D., 2019. Soft shareholder activism. *The Review of Financial Studies* 32, 2775-2808
- Lewis, D., 2023. Air pollution in China is falling—but there is a long way to go. *Nature* 617, 230-231
- Oster, E., 2019. Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics* 37, 187-204
- Qi, Z., Zhou, Y., Chen, J., 2021. Corporate site visits and earnings management. *Journal of Accounting and Public Policy* 40, 106823
- Roychowdhury, S., 2006. Earnings management through real activities manipulation. *Journal of Accounting and Economics* 42, 335-370
- 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, 3814-3854
- Sullivan, D.M., Krupnick, A., 2018. Using satellite data to fill the gaps in the US air pollution monitoring network. *Resources for the Future working paper* 18
- Wang, X., Xie, J., Zhang, B., Zhao, X., 2025. Investor communication and payout policy: A field experiment. *Review of Finance*, forthcoming
- Wei, J., Li, Z., Cribb, M., Huang, W., Xue, W., Sun, L., Guo, J., Peng, Y., Li, J., Lyapustin, A., 2020. Improved 1 km resolution PM 2.5 estimates across China using enhanced space–time extremely randomized trees. *Atmospheric Chemistry and Physics* 20, 3273-3289
- Wei, J., Li, Z., Lyapustin, A., Sun, L., Peng, Y., Xue, W., Su, T., Cribb, M., 2021. Reconstructing 1-km-resolution high-quality PM_{2.5} data records from 2000 to 2018 in China: spatiotemporal variations and policy implications. *Remote Sensing of Environment* 252, 112136
- Xue, Y., Zhang, B., Zhao, X., 2025. Nowcasting firms' operating activities from satellite data on thermal infrared radiation. *Journal of Financial and Quantitative Analysis*, forthcoming
- Yang, L., Lin, Y., Wang, J., Peng, F., 2024. Achieving Air Pollution Control Targets with Technology-Aided Monitoring: Better Enforcement or Localized Efforts? *American Economic Journal: Economic Policy* 16, 280-315
- Yu, F., 2008. Analyst coverage and earnings management. *Journal of Financial Economics* 88, 245-271
- Zang, A.Y., 2012. Evidence on the trade-off between real activities manipulation and accrual-based earnings management. *The Accounting Review* 87, 675-703
- Zhang, R., So, E., Wang, R., 2025. Investor corporate visits and predictable returns. *Journal of Financial and Quantitative Analysis*, 1-34
- Zou, E.Y., 2021. Unwatched pollution: The effect of intermittent monitoring on air quality. *American Economic Review* 111, 2101-2126

Appendix 1: Variable definitions and data sources

Variable	Definition	Source
<i>Outcome variables:</i>		
<i>Pollution</i>	The average PM2.5 concentration ($\mu\text{g}/\text{m}^3$) within a 1km radius of the specific site on day t .	CHAP Database
<i>Pollution^{ring}</i>	The average PM2.5 concentration ($\mu\text{g}/\text{m}^3$) within a buffer ring of 1km to 3km radius surrounding the site on day t .	CHAP Database
<i>Pollution^{abn}</i>	The difference between the site-specific pollution and the surrounding environmental pollution ($Pollution - Pollution^{ring}$) on day t .	CHAP Database
<i>Pollution^{lag}</i>	The average level of <i>Pollution</i> calculated over the pre-event window $[t-14, t-7]$.	CHAP Database
<i>TIR</i>	The average daytime thermal infrared radiation value (<i>Kelvin</i>) at the site on day t , used as a proxy for operational activity.	Xue et al. (2025)
<i>AR</i>	The difference between the site's daytime <i>TIR</i> and the average <i>TIR</i> (<i>Kelvin</i>) of a 5km buffer zone surrounding the site on day t .	Xue et al. (2025)
<i>Key independent variables:</i>		
<i>SiteVisit</i>	1 if the site receives a corporate visit on day t , and 0 otherwise.	CSMAR
<i>SiteVisit^{3day}</i>	1 if the site receives a corporate visit on days $t - 1$, t , or $t + 1$, and 0 otherwise.	CSMAR
<i>Visitor type indicators:</i> All visitor types below are defined as dummy variables equal to 1 if the visiting group includes at least one representative from the specified sector, and 0 otherwise.		
<i>Advisor</i>	1 if the visitors include participants from an investment advisory company, and 0 otherwise.	CSMAR
<i>Assman</i>	1 if the visitors include participants from an assets management company, and 0 otherwise.	CSMAR
<i>Bank</i>	1 if the visitors include participants from the banking sector, and 0 otherwise.	CSMAR
<i>Foreign</i>	1 if the visitors include participants from a foreign institution (QFII/RQFII), and 0 otherwise.	CSMAR
<i>Funds</i>	1 if the visitors include participants from a mutual fund or private equity fund, and 0 otherwise.	CSMAR
<i>Futures</i>	1 if the visitors include participants from a futures company, and 0 otherwise.	CSMAR
<i>Individual</i>	1 if the visitors include individual investors, and 0 otherwise.	CSMAR
<i>Insurance</i>	1 if the visitors include participants from an insurance company, and 0 otherwise.	CSMAR
<i>Media</i>	1 if the visitors include participants from a media	CSMAR

	organization, and 0 otherwise.	
<i>Securities</i>	1 if the visitors include participants from a securities firm (investment bank/brokerage), and 0 otherwise.	CSMAR
<i>Others</i>	1 if the visitors include participants who do not fall into the above categories, and 0 otherwise.	CSMAR
<i>Other visit characteristics:</i>		
<i>Ln(# Word)</i>	The natural logarithm of the total word count of the disclosure record filed for the specific site visit event.	CSMAR
<i>Ln(# Institution)</i>	The natural logarithm of the count of distinct visiting institutions participating in the site visit event.	CSMAR
<i>Ln(# Researcher)</i>	The natural logarithm of the count of distinct individuals participating in the site visit event.	CSMAR
<i>Finance</i>	1 if the visit meeting records contain financing-related keywords, and 0 otherwise.	CSMAR
<i>NonFinance</i>	1 if the visit meeting records do not contain financing-related keywords, and 0 otherwise.	CSMAR
<i>Online</i>	1 if the visit was conducted virtually (e.g., teleconference, webinar) rather than physically, and 0 otherwise.	CSMAR
<i>Stake</i>	1 if the visiting group includes a shareholder in the site visit on a day, and 0 otherwise.	CSMAR
<i>nonStake</i>	1 if the visiting group does not include a shareholder in the site visit on a day, and 0 otherwise.	CSMAR
<i>Stake_Num</i>	The number of visitors who are the shareholders of the firm in the site visit on a day, and 0 otherwise.	CSMAR
<i>nonStake_Num</i>	The number of visitors who are not the shareholders of the firm in the site visit on a day, and 0 otherwise.	CSMAR
<i>PRI</i>	1 if the visiting group includes a Principles for Responsible Investment (PRI) signatory in the site visit on a day, and 0 otherwise.	https://www.unpri.org/ & CSMAR
<i>nonPRI</i>	1 if the visiting group does not include a Principles for Responsible Investment (PRI) signatory in the site visit on a day, and 1 otherwise.	https://www.unpri.org/ & CSMAR
<i>PRI_Num</i>	The number of visitors who are signatories to the Principles for Responsible Investment (PRI) in the site visit on a day, and 0 otherwise.	https://www.unpri.org/ & CSMAR
<i>nonPRI_Num</i>	The number of visitors who are signatories to the Principles for Responsible Investment (PRI) in the site visit on a day, and 0 otherwise.	https://www.unpri.org/ & CSMAR
<i>SameCity</i>	1 for non-visited sites belonging to the same firm and located in the same city as the visited site on	CSMAR

	the event date, and 0 otherwise.	
<i>DiffCity</i>	1 for non-visited sites belonging to the same firm but located in a different city than the visited site on the event date, and 0 otherwise.	CSMAR
<i>SiteAreaSize</i>	The area size (m^2) of the plant site.	Xue et al. (2025)
<i>Other variables:</i>		
<i>ESG Rating</i>	Composite ESG rating provided by Sino-Securities Index.	WIND Database
<i>E/S/G Rating</i>	Individual Environmental (E), Social (S), and Governance (G) sub-ratings provided by the Sino-Securities Index.	WIND Database
<i>HighPollution</i>	1 if the firm operates in a heavily polluting industry (e.g., thermal power, steel, cement, chemicals), and 0 otherwise.	CSMAR
<i>Tobin's Q</i>	The ratio of market value of equity plus book value of liabilities to the book value of assets.	CSMAR
<i>ExePay</i>	Logarithm value of 1 plus total compensation of the top three executives.	CSMAR
<i>IO</i>	The number of shares held by institutional investors over the number of total shares outstanding.	CSMAR

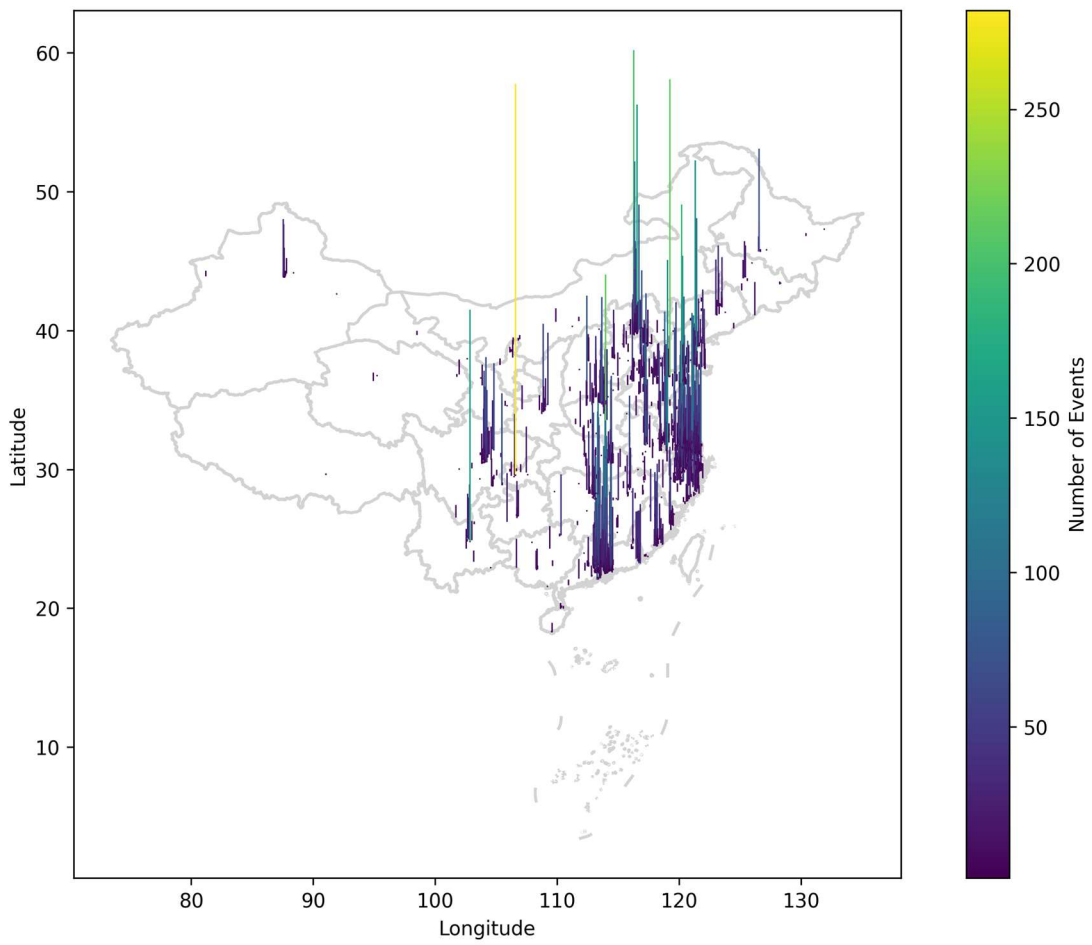
Note: For all variables denoting a specific visitor type (e.g., *Bank*, *Funds*, *Media*) or location characteristic (e.g., *SameCity* and *DiffCity*), we also construct a corresponding 3-day window variable (superscripted with *3day*). These variables are equal to 1 if the respective indicator is active on day $t-1$, t , or $t+1$, and 0 otherwise.

Appendix 2: The distribution of visits by visitor type

Type	Number	Percentage
Total number of site visits (<i>SiteVisit</i>)	30,948	100.00
Visits by type:		
- Securities firm (<i>Securities</i>)	24,769	80.03
- <i>Funds</i>	15,095	48.78
- Assets management companies (<i>Assman</i>)	12,909	41.71
- Foreign institutions (<i>Foreign</i>)	4,427	14.30
- Insurance companies (<i>Insurance</i>)	2,815	9.10
- Individuals (<i>Individual</i>)	972	3.14
- Banks (<i>Bank</i>)	815	2.63
- Investment advisor companies (<i>Advisor</i>)	617	1.99
- <i>Media</i>	571	1.85
- <i>Others</i>	401	1.30
- Futures companies (<i>Futures</i>)	162	0.52

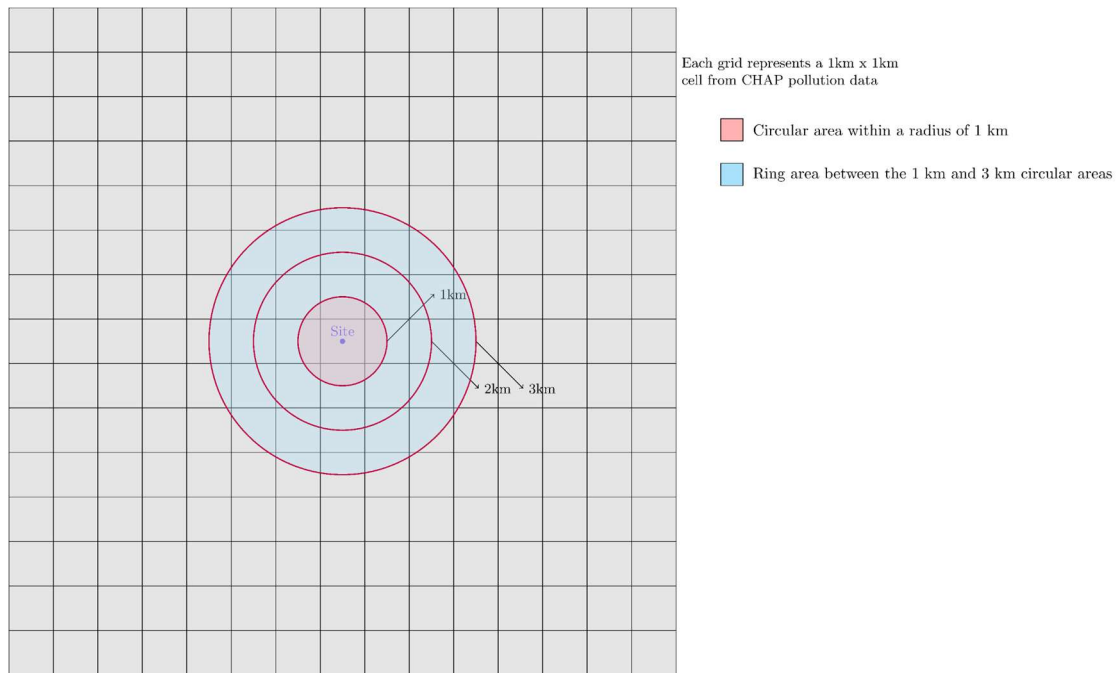
Notes: This table summarizes the frequency and percentage distribution of corporate site visits across different visitor categories for the full sample of 30,948 visit events.

Figure 1: Geographic distribution of sites and frequency of visits across China



Note: This figure displays the geographic distribution of sampled sites across China. Vertical bars mark specific site locations, with both the height and color of the bars representing the number of visits or events that occurred at that specific site.

Figure 2: The constructure of site pollution

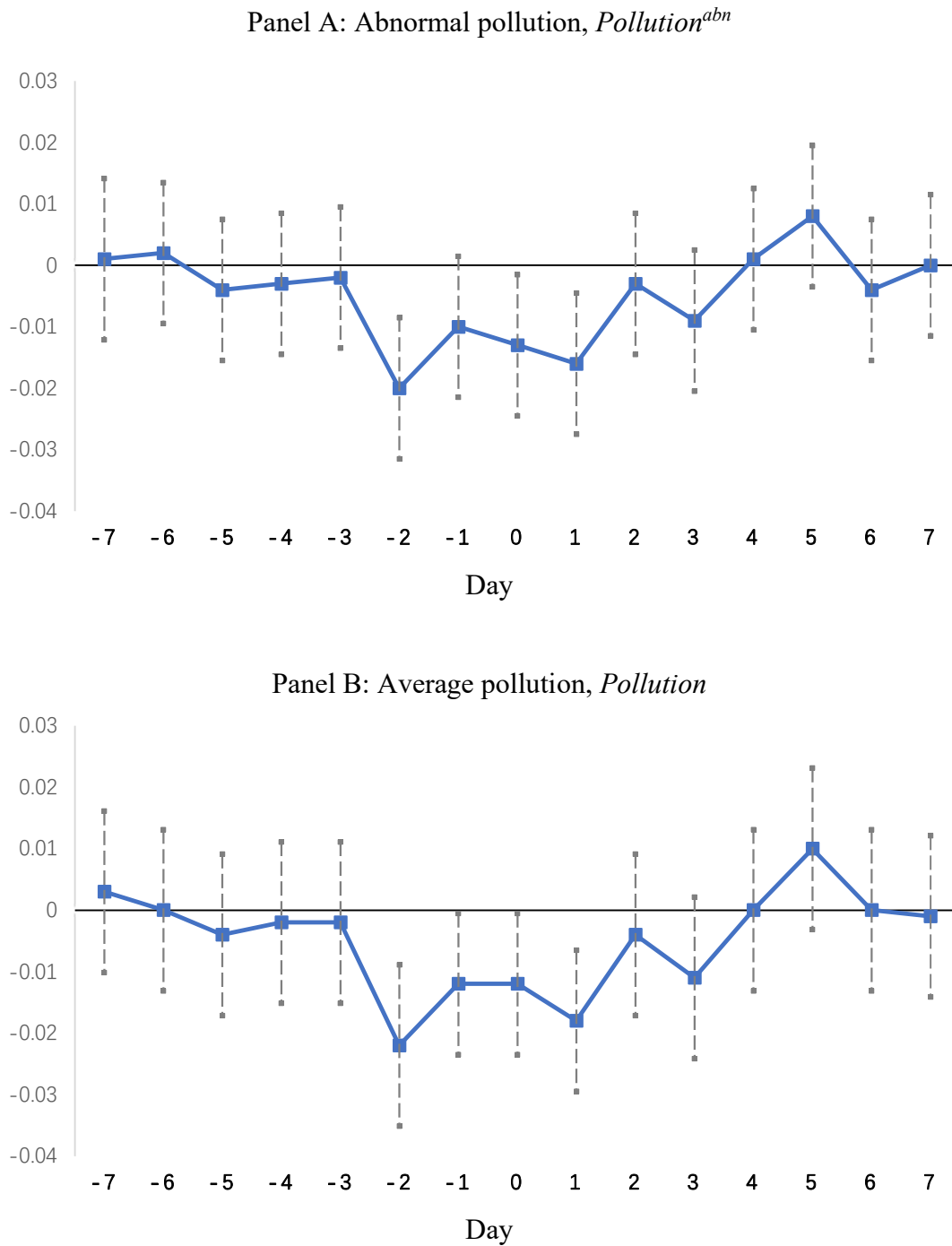


Pollution: The average PM2.5 concentration ($\mu\text{g}/\text{m}^3$) within a 1km radius of the specific site on a day.

Pollution^{ring}: The average PM2.5 concentration ($\mu\text{g}/\text{m}^3$) within a buffer ring of 1km to 3km radius surrounding the site on a day.

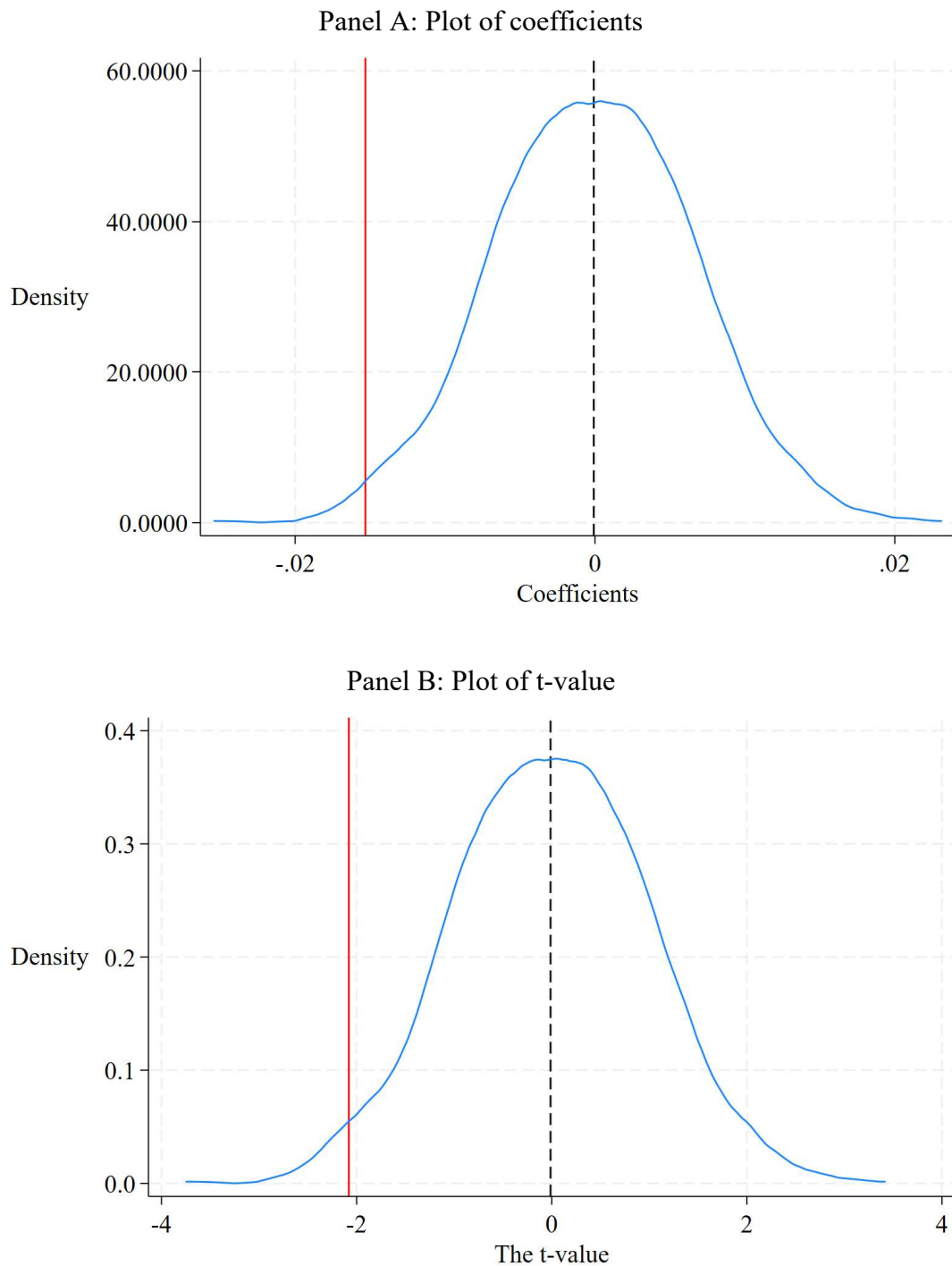
Pollution^{abn}: The difference between the site-specific pollution and the surrounding environmental pollution ($\text{Pollution} - \text{Pollution}^{\text{ring}}$) on a day.

Figure 3: Dynamic impacts of site visits on local air pollution over a 14-day window



Notes: This figure illustrates the dynamic effects of corporate site visits on local air pollution across a 14-day window surrounding the visit event. The analysis plots the estimated coefficients and corresponding confidence intervals (95% level) for a series of dummy variables representing specific days relative to the site visit, ranging from 7 days before ($t-7$) to 7 days after ($t+7$) the event. Panel A displays the results for abnormal pollution ($Pollution^{abn}$), while Panel B shows the results for average pollution ($Pollution$). The horizontal axis represents the timeline relative to the visit day (Day 0), and the vertical axis indicates the magnitude of the estimated impact on pollution levels.

Figure 4: Falsification tests



Notes: This figure presents the results of a placebo test using a randomization procedure to validate the baseline findings. The analysis involves 1,000 simulations where the *SiteVisit* variable is randomly assigned to observations, after which the regression model is re-estimated. Panel A displays the kernel density plot of the estimated coefficients from these 1,000 placebo regressions, with the vertical red line indicating the actual coefficient estimate from the baseline model and the dashed line marking zero. Panel B shows the kernel density plot of the corresponding *t*-values, with the vertical red line representing the actual *t*-value from the baseline model and the dashed line marking zero.

Table 1: The impact of site visits on local air pollution

Dependent variable	(1)	(2)	(3)	(4)
	<i>Pollution^{abn}</i>		<i>Pollution</i>	
<i>SiteVisit</i>	-0.015** (0.007)		-0.015* (0.008)	
<i>SiteVisit^{3day}</i>		-0.016*** (0.006)		-0.017*** (0.006)
<i>Pollution^{ring}</i>	0.026*** (0.001)	0.026*** (0.001)	1.030*** (0.002)	1.030*** (0.002)
<i>Pollution^{lag}</i>	0.011*** (0.001)	0.011*** (0.001)	0.013*** (0.002)	0.013*** (0.002)
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	6,908,720	6,908,720	6,908,720	6,908,720
Adj. R-sq	0.156	0.156	0.957	0.957

Notes: This table presents the regression estimates of the effect of corporate site visits on air pollution surrounding firm sites, using a sample of daily observations at the site level from January 1, 2012, to December 31, 2019. The dependent variable in columns (1) and (2) is *Pollution^{abn}*, defined as the difference between the average PM2.5 concentration within a 1km radius of the site (*Pollution*) and the average concentration within a surrounding 1km to 3km buffer ring (*Pollution^{ring}*). In columns (3) and (4), the dependent variable is the raw pollution level (*Pollution*). The primary independent variables are *SiteVisit*, an indicator equal to one if the site was visited on day t , and *SiteVisit^{3day}*, an indicator equal to one if a visit occurred on day $t-1$, t , or $t+1$. All models include controls for concurrent surrounding pollution (*Pollution^{ring}*) and lagged pollution (*Pollution^{lag}*), calculated as the average pollution level over the pre-event window $[t-14, t-7]$, alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 2: The impact of site visit intensity on local air pollution

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	<i>Pollution^{abn}</i>	<i>Pollution</i>	<i>Pollution^{abn}</i>	<i>Pollution</i>	<i>Pollution^{abn}</i>	<i>Pollution</i>
<i>Ln(# Word)</i>	-0.002** (0.001)	-0.002* (0.001)				
<i>Ln(# Researcher)</i>			-0.008** (0.004)	-0.007* (0.004)		
<i>Ln(# Institution)</i>					-0.007* (0.004)	-0.007 (0.004)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes	Yes	Yes
N	6,908,720	6,908,720	6,908,720	6,908,720	6,908,720	6,908,720
Adj. R-sq	0.156	0.957	0.156	0.957	0.156	0.957

Notes: This table presents regression estimates of the effect of corporate site visit intensity on air pollution surrounding firm sites, using a sample of daily observations at the site level from January 1, 2012, to December 31, 2019. The dependent variable in columns (1), (3), and (5) is *Pollution^{abn}*, defined as the difference between the average PM2.5 concentration within a 1km radius of the site (*Pollution*) and the average concentration within a surrounding 1km to 3km buffer ring (*Pollution^{ring}*). In columns (2), (4), and (6), the dependent variable is the raw pollution level (*Pollution*). The primary independent variables measure visit intensity: *Ln(# Word)* (columns 1 and 2) is the natural logarithm of the total word count of the visit disclosure record; *Ln(# Researcher)* (columns 3 and 4) is the natural logarithm of the count of visitors; and *Ln(# Institution)* (columns 5 and 6) is the natural logarithm of the count of visiting organizations. These variables take a value of zero if no visit occurred. All models include controls for concurrent surrounding pollution (*Pollution^{ring}*) and lagged pollution (*Pollution^{lag}*), calculated as the average pollution level over the pre-event window $[t-14, t-7]$, alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3: The heterogeneous effects of site visits by industry pollution levels

Dependent variable	<i>Pollution^{abn}</i>		<i>Pollution</i>	
	(1)	(2)	(3)	(4)
Subsample	Low-pollution industries	High-pollution industries	Low-pollution industries	High-pollution industries
<i>SiteVisit^{3day}</i>	-0.013* (0.007)	-0.024** (0.011)	-0.013* (0.007)	-0.028** (0.013)
Difference between the coefficients of <i>SiteVisit^{3day}</i>	(2)-(1): -0.011 P-value: 0.070		(4)-(3): -0.015 P-value: 0.020	
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	4,715,800	2,192,920	4,715,800	2,192,920
Adj. R-sq	0.150	0.195	0.957	0.956

Notes: This table presents regression estimates of the heterogeneous effects of corporate site visits on local air pollution, distinguishing between firms based on industry pollution characteristics. The sample consists of daily observations at the site level from January 1, 2012, to December 31, 2019. The dependent variable in columns (1) and (2) is *Pollution^{abn}*, defined as the difference between the average PM2.5 concentration within a 1km radius of the site (*Pollution*) and the average concentration within a surrounding 1km to 3km buffer ring (*Pollution^{ring}*). In columns (3) and (4), the dependent variable is the raw pollution level (*Pollution*). The sample is split into two groups based on industry type: "High-pollution industries" (thermal power, iron and steel, cement, electrolytic aluminum, coal, metallurgy, chemicals, petrochemicals, building materials, paper, brewing, pharmaceuticals, fermentation, textiles, tanning, and mining) and "Low-pollution industries" (the remaining industries). The primary independent variable is *SiteVisit^{3day}*. The difference between the coefficients of the subsamples and the p-value from a test of their equality are reported. All models include controls for concurrent surrounding pollution (*Pollution^{ring}*) and lagged pollution (*Pollution^{lag}*), calculated as the average pollution level over the pre-event window [$t-14$, $t-7$], alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 4: The heterogeneous effects of site visits by firm valuation

Dependent variable	<i>Pollution^{abn}</i>		<i>Pollution</i>	
	High <i>Tobin's Q</i>	Low <i>Tobin's Q</i>	High <i>Tobin's Q</i>	Low <i>Tobin's Q</i>
<i>SiteVisit^{3day}</i>	-0.027*** (0.007)	-0.000 (0.010)	-0.030*** (0.008)	-0.000 (0.010)
Difference between the coefficients of <i>SiteVisit^{3day}</i>	(2)-(1): 0.027 p-value: 0.000		(4)-(3): 0.030 p-value: 0.000	
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	3,906,963	3,001,757	3,906,963	3,001,757
Adj. R-sq	0.171	0.185	0.956	0.958

Notes: This table presents regression estimates of the heterogeneous effects of corporate site visits on local air pollution, distinguishing between firms based on market valuation. The sample consists of daily observations at the site level from January 1, 2012, to December 31, 2019. The dependent variable in columns (1) and (2) is *Pollution^{abn}*, defined as the difference between the average PM2.5 concentration within a 1km radius of the site (*Pollution*) and the average concentration within a surrounding 1km to 3km buffer ring (*Pollution^{ring}*). In columns (3) and (4), the dependent variable is the raw pollution level (*Pollution*). The sample is split into two groups based on firm valuation: "High *Tobin's Q*" (*Tobin's Q* above the industry median) and "Low *Tobin's Q*" (*Tobin's Q* below the industry median). The primary independent variable is *SiteVisit^{3day}*. The difference between the coefficients of the subsamples and the p-value from a test of their equality are reported. All models include controls for concurrent surrounding pollution (*Pollution^{ring}*) and lagged pollution (*Pollution^{lag}*), calculated as the average pollution level over the pre-event window [$t-14$, $t-7$], alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5: The heterogeneous effects of site visits by ESG rating

Dependent variable	<i>Pollution^{abn}</i>		<i>Pollution</i>	
	(1)	(2)	(3)	(4)
Subsample	High ESG Rating	Low ESG Rating	High ESG Rating	Low ESG Rating
<i>SiteVisit^{3day}</i>	-0.009 (0.009)	-0.019*** (0.007)	-0.009 (0.010)	-0.021*** (0.008)
Difference between the coefficients of <i>SiteVisit^{3day}</i>	(2)-(1): -0.010 P-value: 0.100		(4)-(3): -0.012 P-value: 0.070	
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	2,949,944	3,958,776	2,949,944	3,958,776
Adj. R-sq	0.174	0.178	0.959	0.955

Notes: This table presents regression estimates of the heterogeneous effects of corporate site visits on local air pollution, distinguishing between firms based on ESG performance. The sample consists of daily observations at the site level from January 1, 2012, to December 31, 2019. The dependent variable in columns (1) and (2) is *Pollution^{abn}*, defined as the difference between the average PM2.5 concentration within a 1km radius of the site (*Pollution*) and the average concentration within a surrounding 1km to 3km buffer ring (*Pollution^{ring}*). In columns (3) and (4), the dependent variable is the raw pollution level (*Pollution*). The sample is split into two groups based on Sino-Securities Index Bond ESG rating: "High ESG Rating" denotes a score higher than 4, and "Low ESG Rating" denotes a score equal to or lower than 4. The primary independent variable is *SiteVisit^{3day}*. The difference between the coefficients of the subsamples and the p-value from a test of their equality are reported. All models include controls for concurrent surrounding pollution (*Pollution^{ring}*) and lagged pollution (*Pollution^{lag}*), calculated as the average pollution level over the pre-event window [$t-14$, $t-7$], alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 6: Heterogeneous effects of site visits by E, S, and G ratings

Panel A: Subgroup by E rating

Dependent variable	(1)		(2)	
	<i>Pollution^{abn}</i>		<i>Pollution</i>	
Subsample	High E Rating	Low E Rating	High E Rating	Low E Rating
<i>SiteVisit^{3days}</i>	0.010 (0.024)	-0.016*** (0.006)	0.011 (0.024)	-0.017*** (0.006)
Difference between the coefficients of <i>SiteVisit^{3day}</i>	(2)-(1): -0.026 p-value: 0.010		(4)-(3): -0.028 p-value: 0.010	
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	1,093,913	5,814,807	1,093,913	5,814,807
Adj. R-sq	0.188	0.163	0.956	0.957

Panel B: Subgroup by S rating

Dependent variable	(1)		(2)	
	<i>Pollution^{abn}</i>		<i>Pollution</i>	
Subsample	High S Rating	Low S Rating	High S Rating	Low S Rating
<i>SiteVisit^{3days}</i>	-0.011 (0.008)	-0.013* (0.008)	-0.011 (0.009)	-0.015* (0.008)
Difference between the coefficients of <i>SiteVisit^{3day}</i>	(2)-(1): -0.002 p-value: 0.400		(4)-(3): -0.004 p-value: 0.290	
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	3,469,696	3,439,024	3,469,696	3,439,024
Adj. R-sq	0.168	0.183	0.960	0.954

Panel C: Subgroup by G rating

Dependent variable	(1)		(2)	
	<i>Pollution^{abn}</i>		<i>Pollution</i>	
Subsample	High G Rating	Low G Rating	High G Rating	Low G Rating
<i>SiteVisit^{3days}</i>	-0.020*** (0.006)	-0.009 (0.013)	-0.022*** (0.007)	-0.004 (0.013)
Difference between the coefficients of <i>SiteVisit^{3day}</i>	(2)-(1): 0.011 p-value: 0.220		(4)-(3): 0.018 p-value: 0.060	
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	5,567,354	1,341,366	5,567,354	1,341,366
Adj. R-sq	0.161	0.235	0.957	0.956

Notes: This table presents additional results on the heterogeneous effects of site visits, distinguishing between firms based on the individual Environmental (E), Social (S), and Governance (G) components of their ESG ratings. The dependent variable in columns (1) and

(2) is $Pollution^{abn}$, defined as the difference between the average PM2.5 concentration within a 1km radius of the site ($Pollution$) and the average concentration within a surrounding 1km to 3km buffer ring ($Pollution^{ring}$). In columns (3) and (4), the dependent variable is the raw pollution level ($Pollution$). The primary independent variable is $SiteVisit^{3day}$. The sample is split into two groups based on Sino-Securities Index Bond ESG rating: Panel A compares firms with High E Rating versus Low E Rating; Panel B compares High S Rating versus Low S Rating; and Panel C compares High G Rating versus Low G Rating. For each category, a "High" rating denotes a score higher than 4, while a "Low" rating denotes a score equal to or lower than 4. The difference between the coefficients of the subsamples and the p-value from a test of their equality are reported. All models include controls for concurrent surrounding pollution ($Pollution^{ring}$) and lagged pollution ($Pollution^{lag}$), calculated as the average pollution level over the pre-event window $[t-14, t-7]$, alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 7: Visitor ownership stake and PRI signatory

Panel A: Stake visitors				
Dependent variable	(1)	(2)	(3)	(4)
	<i>Pollution^{abn}</i>		<i>Pollution</i>	
<i>Stake</i>	-0.040**		-0.040**	
	(0.016)		(0.018)	
<i>nonStake</i>	-0.010		-0.009	
	(0.008)		(0.009)	
<i>Stake_Num</i>		-0.018*		-0.019*
		(0.010)		(0.011)
<i>nonStake_Num</i>		-0.000		-0.000
		(0.000)		(0.000)
Difference between <i>Stake</i> and <i>nonStake</i>	-0.031		-0.031	
	p-value: 0.080		p-value: 0.124	
Difference between <i>Stake_Num</i> and <i>nonStake_Num</i>		-0.018		-0.018
		p-value: 0.080		p-value: 0.103
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	6,908,720	6,908,720	6,908,720	6,908,720
Adj. R-sq	0.156	0.156	0.957	0.957
Panel B: PRI visitors				
Dependent variable	(1)	(2)	(3)	(4)
	<i>Pollution^{abn}</i>		<i>Pollution</i>	
<i>PRI</i>	-0.013		-0.013	
	(0.010)		(0.011)	
<i>nonPRI</i>	-0.018*		-0.017	
	(0.011)		(0.012)	
<i>PRI_Num</i>		-0.001		-0.001
		(0.002)		(0.002)
<i>nonPRI_Num</i>		-0.001		-0.000
		(0.000)		(0.000)
Difference between <i>PRI</i> and <i>nonPRI</i>	0.005		0.004	
	p-value: 0.704		p-value: 0.814	
Difference between <i>PRI_Num</i> and <i>nonPRI_Num</i>		-0.000		-0.000
		p-value: 0.833		p-value: 0.833
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	6,908,720	6,908,720	6,908,720	6,908,720
Adj. R-sq	0.156	0.156	0.957	0.957
Panel C: Horse race between stake and PRI visitors				
Dependent variable	(1)	(2)	(3)	(4)
	<i>Pollution^{abn}</i>		<i>Pollution</i>	
<i>Stake</i>	-0.046**		-0.044**	
	(0.019)		(0.022)	

<i>nonStake</i>	-0.014 (0.011)		-0.012 (0.013)	
<i>Stake_Num</i>		-0.020* (0.011)		-0.020 (0.012)
<i>nonStake_Num</i>		-0.000 (0.000)		-0.000 (0.000)
<i>PRI</i>	0.008 (0.014)		0.006 (0.016)	
<i>PRI_num</i>		0.001 (0.002)		0.001 (0.003)
Difference between <i>Stake</i> and <i>nonStake</i>	-0.032 p-value: 0.074		-0.032 p-value: 0.117	
Difference between <i>Stake_Num</i> and <i>nonStake_Num</i>		-0.019 p-value: 0.086		-0.019 p-value: 0.108
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	6,908,720	6,908,720	6,908,720	6,908,720
Adj. R-sq	0.156	0.156	0.957	0.957

Notes: This table presents regression estimates of the effect of visitors' ownership stake and Principles for Responsible Investment (PRI) signatory on local air pollution. The sample consists of daily observations at the site level from January 1, 2012, to December 31, 2019. The dependent variable in columns (1) and (2) is $Pollution^{abn}$, defined as the difference between the average PM2.5 concentration within a 1km radius of the site ($Pollution$) and the average concentration within a surrounding 1km to 3km buffer ring ($Pollution^{ring}$). In columns (3) and (4), the dependent variable is the raw pollution level ($Pollution$). Panel A distinguishes visits based on stake status, where *Stake* (*nonStake*) equals one if the visiting group includes (does not include) a shareholder. Panel B distinguishes visits based on PRI status, where *PRI* (*nonPRI*) equals one if the group includes (does not include) a PRI signatory. Panel C presents a horse race including both categories. The variables suffixed with *_Num* represent the count of visitors in each respective category. The differences between coefficients and the corresponding p-values for tests of equality are reported. All models include controls for concurrent surrounding pollution ($Pollution^{ring}$) and lagged pollution ($Pollution^{lag}$), calculated as the average pollution level over the pre-event window $[t-14, t-7]$, alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 8: The effects on sister non-visited sites

Panel A: Sites of the same company in the same city				
Dependent variable	(1)	(2)	(3)	(4)
	<i>Pollution^{abn}</i>		<i>Pollution</i>	
<i>SameCity</i>	0.006 (0.008)		0.008 (0.008)	
<i>SameCity^{3day}</i>		0.005 (0.006)		0.006 (0.006)
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
Adj. R-sq	0.156	0.156	0.957	0.957
N	6,908,720	6,908,720	6,908,720	6,908,720

Panel B: Sites of the same company in different cities				
Dependent variable	(1)	(2)	(3)	(4)
	<i>Pollution^{abn}</i>		<i>Pollution</i>	
<i>DiffCity</i>	-0.007 (0.013)		-0.010 (0.014)	
<i>DiffCity^{3day}</i>		-0.015 (0.009)		-0.017 (0.011)
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
Adj. R-sq	0.156	0.156	0.957	0.957
N	6,908,720	6,908,720	6,908,720	6,908,720

Notes: This table presents the regression estimates examining the effects of site visits on other sites within the same company. The sample consists of daily observations at the site level from January 1, 2012, to December 31, 2019. The dependent variable in columns (1) and (2) is *Pollution^{abn}*, defined as the difference between the average PM2.5 concentration within a 1km radius of the site (*Pollution*) and the average concentration within a surrounding 1km to 3km buffer ring (*Pollution^{ring}*). In columns (3) and (4), the dependent variable is the raw pollution level (*Pollution*). Panel A reports estimates for non-visited sites belonging to the same firm and located in the same city as the visited site (*SameCity* and *SameCity^{3day}*). Panel B reports estimates for non-visited sites belonging to the same firm but located in a different city (*DiffCity* and *DiffCity^{3day}*). *SameCity* and *DiffCity* indicate the observations of non-visited sites for day t , while *SameCity^{3day}* and *DiffCity^{3day}* indicate the observations of non-visited sites in the 3-day window ($t-1$ to $t+1$). All models include controls for concurrent surrounding pollution (*Pollution^{ring}*) and lagged pollution (*Pollution^{lag}*), calculated as the average pollution level over the pre-event window $[t-14, t-7]$, alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 9: Change in operational activities

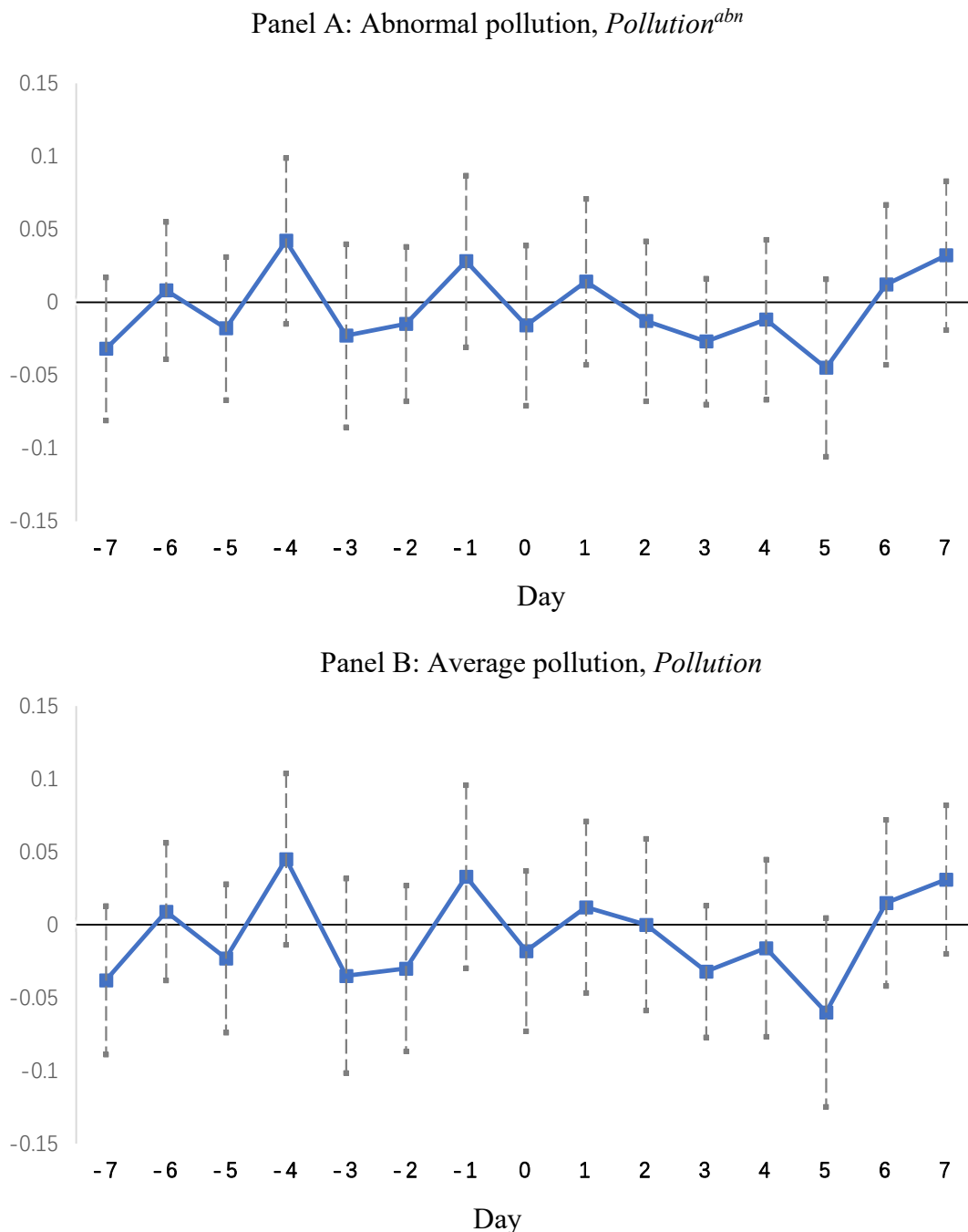
Dependent variable	(1)	(2)	(3)	(4)
	<i>TIR</i>		<i>AR</i>	
<i>SiteVisit</i>	0.012 (0.018)		-0.005 (0.016)	
<i>SiteVisit</i> ^{3day}		0.000 (0.012)		-0.004 (0.010)
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	712,027	712,027	712,027	712,027
Adj. R-sq	0.987	0.987	0.668	0.668

Notes: This table presents the regression estimates of the effect of corporate site visits on firms' operating activities, using a sample of daily observations at the site level from January 1, 2012, to December 31, 2019. The dependent variable in columns (1) and (2) is *TIR*, defined as the average daytime thermal infrared radiation value of the site on day t . The dependent variable in columns (3) and (4) is *AR*, calculated as the difference between the average *TIR* of the site and that of a buffer zone within 5km on day t . The primary independent variables are *SiteVisit*, an indicator equal to one if the site was visited on day t , and *SiteVisit*^{3day}, an indicator equal to one if a visit occurred on day $t-1$, t , or $t+1$. All models include controls for the area size of the site (*SiteAreaSize*) and the lagged *TIR/AR*, calculated as the average *TIR/AR* over the pre-event window $[t-14, t-7]$, alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.

Internet Appendix

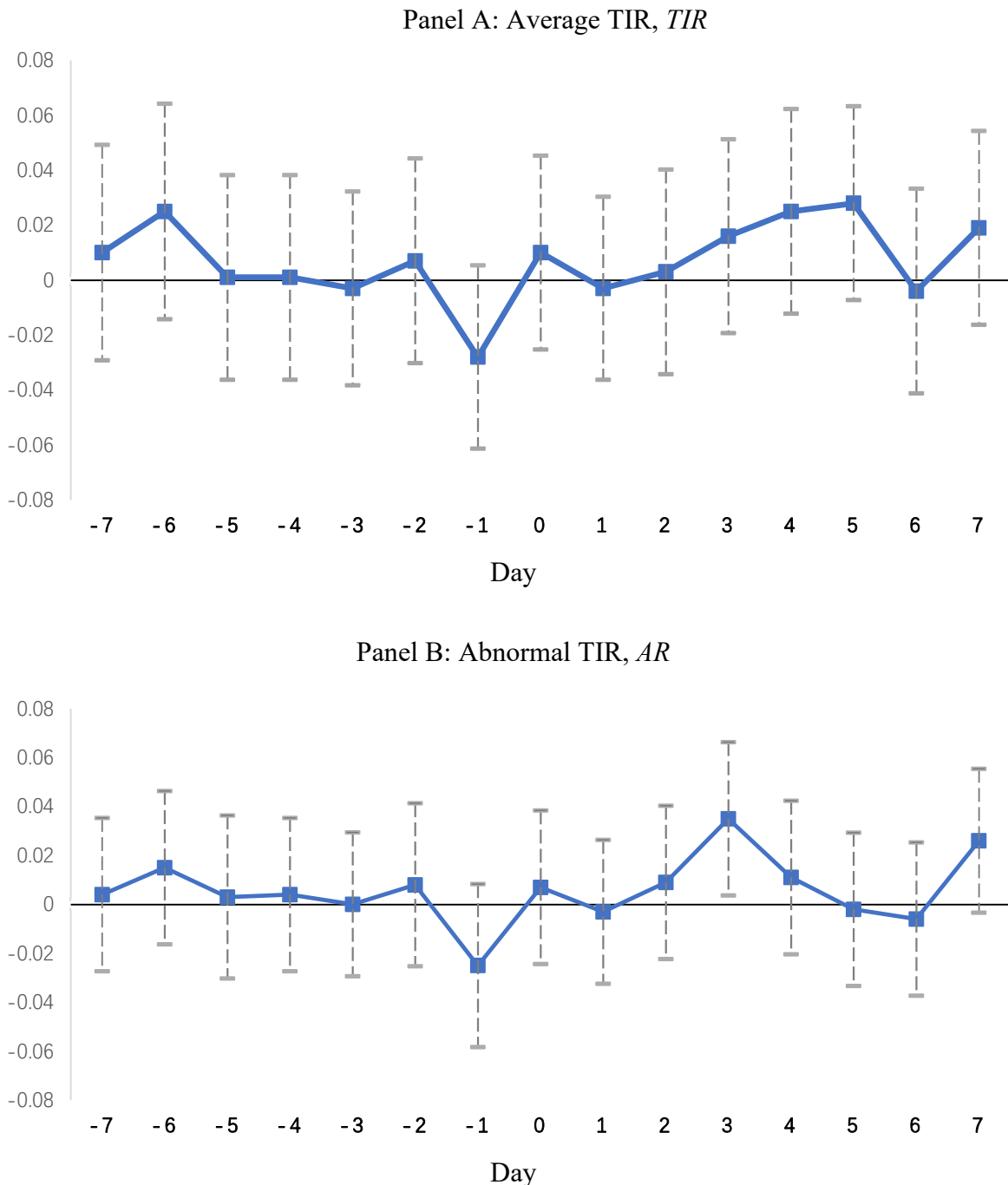
Under the Spotlight: Pollution Dynamics Around Corporate Site Visits

Figure IA1: Dynamic impacts of online visits on local air pollution over a 14-day window



Notes: This figure illustrates the dynamic effects of online site visits on local air pollution across a 14-day window surrounding the event. The analysis plots the estimated coefficients and corresponding confidence intervals (95% level) for a series of dummy variables representing specific days relative to the site visit, ranging from 7 days before ($t-7$) to 7 days after ($t+7$) the event. Panel A displays the results for abnormal pollution ($Pollution^{abn}$), while Panel B shows the results for average pollution ($Pollution$). The horizontal axis represents the timeline relative to the visit day (Day 0), and the vertical axis indicates the magnitude of the estimated impact on pollution levels.

Figure IA2: Dynamic impacts of site visits on operational activities over a 14-day window



Note: This figure illustrates the dynamic effects of site visits on operational activities, measured by thermal infrared radiation (TIR), across a 14-day window surrounding the visit event. The analysis plots the estimated coefficients and corresponding confidence intervals for a series of dummy variables representing specific days relative to the site visit, ranging from 7 days before ($t-7$) to 7 days after ($t+7$) the event. Panel A displays the results for average TIR (TIR), while Panel B shows the results for abnormal TIR (AR), defined as the difference between the site's TIR and that of a buffer zone within 5km. The horizontal axis represents the timeline relative to the visit day (Day 0), and the vertical axis indicates the magnitude of the estimated impact on TIR levels.

Table IA1: Sample distribution by CSRC 2012 industry classification guidelines

Code	Industry name	Freq.	Percent
	Food, Beverage, and Tobacco		
C13	Agricultural Food Processing Industry	34	2.4
C14	Food Manufacturing	17	1.2
C15	Wine, Beverage and Refined Tea Manufacturing Industry	13	0.92
	Textile, Apparel, and Footwear		
C17	Textile Industry	20	1.41
C18	Textile, Clothing and Apparel Industry	26	1.84
C19	Leather, fur, feather and their products and footwear industry	3	0.21
	Wood and Furniture		
C20	Wood processing and wood, bamboo, rattan, palm and grass products industry	5	0.35
C21	Furniture Manufacturing	9	0.64
	Paper, Printing, and Cultural Products		
C22	Paper-making and paper products industry	15	1.06
C23	Printing and recording media reproduction industry	6	0.42
C24	Cultural, educational, industrial and aesthetic, sports and recreational goods manufacturing industry	12	0.85
	Petroleum, Chemical, and Biopharmaceutical		
C25	Petroleum Processing, Coking and Nuclear Fuel Processing Industry	7	0.5
C26	Chemical raw materials and chemical products manufacturing	144	10.18
C27	Pharmaceutical Manufacturing	125	8.84
C28	Chemical Fiber Manufacturing	16	1.13
C29	Rubber and plastic products industry	52	3.68
	Metal and Non-metal		
C30	Non-metallic Mineral Products Industry	51	3.61
C31	Ferrous Metals Smelting and Calendering Industry	11	0.78
C32	Nonferrous Metals Smelting and Calendering Industry	38	2.69
C33	Metal Products Industry	41	2.9
	Specialized, General, and Transportation Equipment		
C34	General Equipment Manufacturing	92	6.51
C35	Special Purpose Equipment Manufacturing	142	10.04
C36	Automobile Manufacturing	55	3.89
C37	Railway, Shipbuilding, Aerospace and Other Transportation Equipment Manufacturing Industry	20	1.41
	Electrical, Electronic, and Communication Equipment		
C38	Electrical Machinery and Equipment Manufacturing	153	10.82
C39	Computer, communications and other electronic equipment manufacturing	246	17.4
	Instruments and Meters		
C40	Instrumentation Manufacturing	41	2.9
	Other Manufacturing Industries		
C41	Other manufacturing industries	18	1.27
C42	Waste Resources Comprehensive Utilization Industry	2	0.14
Total		1,414	100

Table IA2: Summary statistics

Variable	N	Mean	SD	P25	Median	P75
<i>Outcome variables:</i>						
<i>Pollution</i>	6,908,720	49.573	34.539	26.50	40.43	61.55
<i>Pollution^{abn}</i>	6,908,720	0.093	1.209	-0.42	0.03	0.55
<i>Pollution^{lag}</i>	6,908,720	49.857	28.378	30.01	43.05	62.24
<i>Pollution^{ring}</i>	6,908,720	49.479	34.383	26.49	40.37	61.40
<i>TIR</i>	708,311	296.448	11.037	288.70	298.25	305.20
<i>AR</i>	708,311	1.350	1.844	0.16	1.14	2.39
<i>Key independent variables:</i>						
<i>SiteVisit</i>	6,908,720	0.004	0.067	0.00	0.00	0.00
<i>Sitevisit^{3day}</i>	6,908,720	0.013	0.111	0.00	0.00	0.00
<i>Visitor type indicators:</i>						
<i>Advisor</i>	6,908,720	0.000	0.009	0.00	0.00	0.00
<i>Advisor^{3day}</i>	6,908,720	0.000	0.016	0.00	0.00	0.00
<i>Assman</i>	6,908,720	0.002	0.043	0.00	0.00	0.00
<i>Assman^{3day}</i>	6,908,720	0.005	0.074	0.00	0.00	0.00
<i>Bank</i>	6,908,720	0.000	0.011	0.00	0.00	0.00
<i>Bank^{3day}</i>	6,908,720	0.000	0.019	0.00	0.00	0.00
<i>Foreign</i>	6,908,720	0.001	0.025	0.00	0.00	0.00
<i>Foreign^{3day}</i>	6,908,720	0.002	0.043	0.00	0.00	0.00
<i>Funds</i>	6,908,720	0.002	0.047	0.00	0.00	0.00
<i>Funds^{3day}</i>	6,908,720	0.006	0.079	0.00	0.00	0.00
<i>Futures</i>	6,908,720	0.000	0.005	0.00	0.00	0.00
<i>Futures^{3day}</i>	6,908,720	0.000	0.008	0.00	0.00	0.00
<i>Individual</i>	6,908,720	0.000	0.012	0.00	0.00	0.00
<i>Individual^{3day}</i>	6,908,720	0.000	0.020	0.00	0.00	0.00
<i>Insurance</i>	6,908,720	0.000	0.020	0.00	0.00	0.00
<i>Insurance^{3day}</i>	6,908,720	0.001	0.035	0.00	0.00	0.00
<i>Media</i>	6,908,720	0.000	0.009	0.00	0.00	0.00
<i>Media^{3day}</i>	6,908,720	0.000	0.016	0.00	0.00	0.00
<i>Securities</i>	6,908,720	0.004	0.060	0.00	0.00	0.00
<i>Securities^{3day}</i>	6,908,720	0.010	0.101	0.00	0.00	0.00
<i>Others</i>	6,908,720	0.000	0.008	0.00	0.00	0.00
<i>Others^{3day}</i>	6,908,720	0.000	0.013	0.00	0.00	0.00
<i>Other visit characteristics:</i>						
<i>Ln(# Word)</i>	6,908,720	0.031	0.468	0.00	0.00	0.00
<i>Ln(# Institution)</i>	6,908,720	0.007	0.124	0.00	0.00	0.00
<i>Ln(# Researcher)</i>	6,908,720	0.007	0.125	0.00	0.00	0.00
<i>Finance^{3day}</i>	6,908,720	0.001	0.035	0.00	0.00	0.00
<i>NonFinance^{3day}</i>	6,908,720	0.012	0.109	0.00	0.00	0.00
<i>Online</i>	6,908,720	0.000	0.014	0.00	0.00	0.00
<i>Stake</i>	6,908,720	0.001	0.028	0.00	0.00	0.00
<i>nonStake</i>	6,908,720	0.004	0.060	0.00	0.00	0.00
<i>Stake_Num</i>	6,908,720	0.001	0.045	0.00	0.00	0.00
<i>nonStake_Num</i>	6,908,720	0.034	1.449	0.00	0.00	0.00
<i>PRI</i>	6,908,720	0.002	0.049	0.00	0.00	0.00
<i>nonPRI</i>	6,908,720	0.002	0.045	0.00	0.00	0.00
<i>PRI_Num</i>	6,908,720	0.008	0.235	0.00	0.00	0.00
<i>nonPRI_Num</i>	6,908,720	0.027	1.362	0.00	0.00	0.00

<i>SameCity</i>	6,908,720	0.004	0.059	0.00	0.00	0.00
<i>SameCity</i> ^{3day}	6,908,720	0.010	0.099	0.00	0.00	0.00
<i>DiffCity</i>	6,908,720	0.002	0.039	0.00	0.00	0.00
<i>DiffCity</i> ^{3day}	6,908,720	0.004	0.066	0.00	0.00	0.00
<i>Other variables:</i>						
<i>ESG Rating</i>	6,908,720	4.174	1.032	4.00	4.00	5.00
<i>E Rating</i>	6,908,720	2.340	1.501	1.00	2.00	3.00
<i>S Rating</i>	6,908,720	4.289	1.581	3.00	5.00	5.00
<i>G Rating</i>	6,908,720	5.230	1.310	5.00	5.00	6.00
<i>HighPollution</i>	6,908,720	0.317	0.465	0.00	0.00	1.00
<i>Tobin's Q</i>	6,908,720	2.141	1.320	1.40	1.90	2.31
<i>ExePay</i>	6,908,720	14.354	0.669	13.96	14.35	14.69
<i>IO</i>	6,908,720	34.826	23.321	13.39	34.83	52.38

Table IA3: Oster delta statistic analysis

Dependent variable	Independent variable	Delta	R-Squared (uncontrolled):	R-Squared (controlled):
<i>Pollution^{abn}</i>	<i>SiteVisit</i>	1.11198	0.000	0.225
<i>Pollution^{abn}</i>	<i>SiteVisit^{3day}</i>	1.63392	0.000	0.225
<i>Pollution</i>	<i>SiteVisit</i>	-35.45095	0.000	0.957
<i>Pollution</i>	<i>SiteVisit^{3day}</i>	-22.34914	0.000	0.957

Notes: The table reports the calculated Delta statistic of Oster (2019), along with the uncontrolled and controlled R-squared values, for our baseline regression models where the dependent variables are *Pollution^{abn}* and *Pollution*, and the independent variables are *SiteVisit* and *SiteVisit^{3day}*. Controls include concurrent surrounding pollution (*Pollution^{ring}*) and lagged pollution (*Pollution^{lag}*), calculated as the average pollution level over the pre-event window [$t-14, t-7$], alongside site and city-date fixed effects.

Table IA4: Entropy balancing approach

Dependent variable	(1)	(2)	(3)	(4)
	<i>Pollution^{abn}</i>		<i>Pollution</i>	
<i>SiteVisit</i>	-0.013*		-0.015**	
	(0.007)		(0.008)	
<i>SiteVisit^{3day}</i>		-0.014*		-0.016**
		(0.007)		(0.008)
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	6,908,720	6,908,720	6,908,720	6,908,720
Adj. R-sq	0.419	0.419	0.955	0.955

Notes: This table presents the baseline regression estimates using an entropy balancing approach (Hainmueller, 2012; Hainmueller and Xu, 2013). The dependent variable in columns (1) and (2) is *Pollution^{abn}*, defined as the difference between the average PM2.5 concentration within a 1km radius of the site (*Pollution*) and the average concentration within a surrounding 1km to 3km buffer ring (*Pollution^{ring}*). In columns (3) and (4), the dependent variable is the raw pollution level (*Pollution*). The primary independent variables are *SiteVisit*, an indicator equal to one if the site was visited on day t , and *SiteVisit^{3day}*, an indicator equal to one if a visit occurred on day $t-1$, t , or $t+1$. The analysis constructs a counterfactual control group in which we assign a weight (between 0 and 1) to each observation so that the variables we incorporate as matching dimensions (mainly place and time) are balanced with or without site visit. All models include controls for concurrent surrounding pollution (*Pollution^{ring}*) and lagged pollution (*Pollution^{lag}*), calculated as the average pollution level over the pre-event window [$t-14$, $t-7$], alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.

Table IA5: Geographic variation

Panel A: Excluding provincial capitals and municipalities				
	(1)	(2)	(3)	(4)
Dependent variable	<i>Pollution^{abn}</i>		<i>Pollution</i>	
<i>SiteVisit</i>	-0.016*		-0.016	
	(0.009)		(0.010)	
<i>SiteVisit^{3day}</i>		-0.012*		-0.013*
		(0.007)		(0.008)
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	4,146,400	4,146,400	4,146,400	4,146,400
Adj. R-sq	0.186	0.186	0.952	0.952
Panel B: Exclude southern provinces				
	(1)	(2)	(3)	(4)
Dependent variable	<i>Pollution^{abn}</i>		<i>Pollution</i>	
<i>SiteVisit</i>	-0.032*		-0.027	
	(0.017)		(0.020)	
<i>SiteVisit^{3day}</i>		-0.036***		-0.040***
		(0.013)		(0.015)
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	1,810,400	1,810,400	1,810,400	1,810,400
Adj. R-sq	0.158	0.158	0.966	0.966
Panel C: Exclude coastal provinces				
	(1)	(2)	(3)	(4)
Dependent variable	<i>Pollution^{abn}</i>		<i>Pollution</i>	
<i>SiteVisit</i>	-0.036**		-0.031*	
	(0.016)		(0.019)	
<i>SiteVisit^{3day}</i>		-0.033***		-0.034**
		(0.012)		(0.013)
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	1,988,520	1,988,520	1,988,520	1,988,520
Adj. R-sq	0.144	0.144	0.962	0.962

Notes: This table presents geographic variation of the baseline results. The dependent variable in columns (1) and (2) is *Pollution^{abn}*, defined as the difference between the average PM2.5 concentration within a 1km radius of the site (*Pollution*) and the average concentration within a surrounding 1km to 3km buffer ring (*Pollution^{ring}*). In columns (3) and (4), the dependent variable is the raw pollution level (*Pollution*). The primary independent variables are *SiteVisit*, an indicator equal to one if the site was visited on day t , and *SiteVisit^{3day}*, an indicator equal to one if a visit occurred on day $t-1$, t , or $t+1$. Panel A excludes provincial capitals and municipalities; Panel B excludes southern provinces (Shanghai, Anhui, Guangdong, Jiangxi, Guangxi Zhuang Autonomous Region, Jiangsu, Hubei, Zhejiang, Hainan, Fujian, Yunnan, Guizhou, Sichuan, and Chongqing); and Panel C excludes coastal provinces (Shanghai, Liaoning, Guangdong, Guangxi Zhuang Autonomous Region, Jiangsu, Zhejiang, Hainan, Fujian, Tianjin, Hebei, and Shandong). All models include controls for concurrent surrounding

pollution ($Pollution^{ring}$) and lagged pollution ($Pollution^{lag}$), calculated as the average pollution level over the pre-event window $[t-14, t-7]$, alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.

Table IA6: The heterogeneous effects of site visits by governance

Panel A: Internal governance: Managerial compensation

Dependent variable	<i>Pollution^{abn}</i>		<i>Pollution</i>	
	(1)	(2)	(3)	(4)
Subsample	Low <i>ExePay</i>	High <i>ExePay</i>	Low <i>ExePay</i>	High <i>ExePay</i>
<i>SiteVisit^{3day}</i>	-0.019** (0.009)	-0.011 (0.008)	-0.022** (0.010)	-0.010 (0.008)
Difference between the coefficients of <i>SiteVisit^{3day}</i>	(2)-(1): 0.008 p-value: 0.110		(4)-(3): 0.0103 p-value: 0.070	
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	3,111,248	3,797,472	3,111,248	3,797,472
Adj. R-sq	0.178	0.166	0.957	0.956

Panel B: External governance: Institutional ownership

Dependent variable	<i>Pollution^{abn}</i>		<i>Pollution</i>	
	(1)	(2)	(3)	(4)
Subsample	Low <i>IO</i>	High <i>IO</i>	Low <i>IO</i>	High <i>IO</i>
<i>SiteVisit^{3day}</i>	-0.008 (0.008)	-0.019** (0.008)	-0.009 (0.009)	-0.019** (0.009)
Difference between the coefficients of <i>SiteVisit^{3day}</i>	(2)-(1): 0.011 p-value: 0.080		(4)-(3): 0.010 p-value: 0.100	
Controls	Yes	Yes	Yes	Yes
Site FE	Yes	Yes	Yes	Yes
City-date FE	Yes	Yes	Yes	Yes
N	3,038,865	3,869,855	3,038,865	3,869,855
Adj. R-sq	0.190	0.166	0.958	0.956

Notes: This table presents regression estimates of the heterogeneous effects of corporate site visits on local air pollution, distinguishing between internal and external governance mechanisms. The sample consists of daily observations at the site level from January 1, 2012, to December 31, 2019. Panel A focuses on internal governance proxied by managerial compensation (*ExePay*), defined as the logarithm of one plus the total compensation of the top three executives. Panel B focuses on external governance proxied by institutional ownership (*IO*), measured as the number of shares held by institutional investors divided by the total number of shares outstanding. The dependent variable in columns (1) and (2) is *Pollution^{abn}*, defined as the difference between the average PM2.5 concentration within a 1km radius of the site (*Pollution*) and the average concentration within a surrounding 1km to 3km buffer ring (*Pollution^{ring}*). In columns (3) and (4), the dependent variable is the raw pollution level (*Pollution*). The sample is split into two subsamples based on whether the governance variable is above or below the industry median. The primary independent variable is *SiteVisit^{3day}*. The difference between the coefficients of the subsamples and the p-value from a test of their equality are reported. All models include controls for concurrent surrounding pollution (*Pollution^{ring}*) and lagged pollution (*Pollution^{lag}*), calculated as the average pollution level over the pre-event window [$t-14$, $t-7$], alongside site and city-date fixed effects. All variables are defined in Appendix 1. Standard errors are clustered at the site level and are shown in parentheses, with *, **, and *** denoting statistical significance at the 10%, 5%, and 1% levels, respectively.