American Economic Journal: Applied Economics 2024, 16(1): 213–252
https://doi.org/10.1257/app.20210386

Informed Enforcement: Lessons from Pollution Monitoring in China†

By Sebastian Axbard and Zichen Deng*

Government regulations are often imperfectly enforced by public officials. In this study, we exploit the introduction of air pollution monitors in China to investigate whether real-time monitoring of policy outcomes affects the enforcement of existing regulations. Using assignment criteria established by the central government and new georeferenced data on local enforcement activities, we show that monitoring (i) increases enforcement against local firms, (ii) improves the targeting of enforcement, and (iii) reduces aggregate pollution. These effects are driven by officials facing performance incentives and are stronger when there is limited scope for data manipulation, suggesting that real-time monitoring improves top-down accountability. (*JEL* K32, L51, O13, P25, P28, Q52, Q53)

Across the globe, there is a substantial discrepancy between central government regulations and actual enforcement of those regulations at the local level. This gap exists across a wide range of policy areas and is particularly severe in low- and middle-income countries (Nakagaki, Kompanek, and Tominic 2012; World Bank 2017). A common practice to address the principal–agent problem inherent in the delegation of authority to lower levels of government is to provide high-powered incentives to implementing officials to ensure that their interests are better aligned with those of the policymaker.[1] However, such incentive schemes require reliable information on the actions of agents or local policy outcomes. In many settings, such information either is not widely available, is of poor quality, or could easily be manipulated by local officials who have an interest in misreporting due to the incentives they face (Jacob and Levitt 2003; Figlio and Winicki 2005; Figlio and Getzler 2006; Banerjee, Duflo, and Glennerster 2008; Sandefur and Glassman 2015; Fisman and Wang 2017; Greenstone et al. 2022; Acemoglu et al. 2020).

*Axbard: School of Economics and Finance, Queen Mary, University of London (email: s.axbard@qmul.ac.uk); Deng: School of Economics, University of Amsterdam; Department of Economics, Norwegian School of Economics (email: z.deng@uva.nl). Benjamin Olken was coeditor for this article. We are grateful to seminar participants at University of Bergen, University of Zurich, University of York, BI Norwegian Business School, University of Helsinki, CEIBS, NHH, University of Amsterdam, Renmin University, VU Amsterdam, Queen Mary University of London, Warwick, University of East Anglia, University of Gothenburg, EBRD Conference on Corruption and Anti-Corruption Policies (Kyiv), and NEUDC 2019 (Northwestern University) for many useful comments and suggestions. We also thank Jordan Ashmore, Pengzhan Qian, Yu Xiao, and Dan Xie for excellent research assistance.

† Go to https://doi.org/10.1257/app.20210386 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

[1] The theoretical literature has focused on how incentives could be designed to ensure the motivation of agents while decreasing any distortionary impact on effort (Holmström 1979; Holmström and Milgrom 1991; Baker, Gibbons, and Murphy 1994).

This paper explores how a technology that enables the central government to directly monitor local policy outcomes in real time can overcome the gap in enforcement. More specifically, we study the introduction of air pollution monitors in China—a setting where local officials face strong incentives to reduce pollution under centrally set targets—and investigate how that affects local governments' enforcement of air pollution regulations as well as local pollution levels. Our focus on environmental policy is motivated by recent reporting from the United Nation (2019) arguing that a lack of enforcement of environmental regulations is one of the greatest obstacles that needs to be overcome in order to combat climate change and pollution. Despite international efforts in recent years to improve air quality, more than 90 percent of the world's population in 2016 (WHO 2016) still lived in areas where air pollution exceeded World Health Organization guidelines, with far-reaching consequences for both health and productivity (Neidell and Currie 2005; Greenstone and Hanna 2014; Ebenstein et al. 2017; Jia 2017; Barwick et al. 2018). A large part of this population lived in emerging economies, including China, where pollution levels have exceeded the highest levels recorded in high-income countries.

We begin by investigating how a central-government-led program that introduced 552 pollution monitors in 2015 has shaped the enforcement activities of city governments in China. To conduct this analysis, we collect more than 55,000 environmental enforcement records from local governments. We then classify these records and identify the firm involved, the type of regulation violated, and the punishment imposed. Using this information, we estimate a flexible difference-in-difference (DiD) model, which compares firms located close to a monitor with firms located farther away from the monitor but within the same city. The results show an increase in the probability of enforcement by 72 percent for firms located within 10 kilometers (km) of a monitor, consistent with anecdotal evidence suggesting that cities stepped up enforcement activities close to the monitors after their introduction (see discussion in Section IIIB). The main threat to identification—potential endogenous placement of monitors—is mitigated in this setting because the placement of monitors followed strict guidelines issued by the central government. We investigate the determinants of monitor location and document that the placement is unrelated to prior enforcement activity and that there are no differential pre-trends for firms located at different distances from the monitor. In addition, we show that air pollution monitoring does not affect enforcement related to other types of environmental regulations.

To shed further light on how government actions are affected, we investigate how the type of enforcement carried out changes in the presence of monitoring. We show that local governments impose stricter punishment against high-polluting firms and become more responsive to local pollution shocks once monitors have been introduced. To show the latter, we exploit exogenous variation in pollution induced by variations in rainfall and wind direction. We show that enforcement is higher when rainfall is low (and pollution is high) in the presence of monitoring, but that no such relationship exist when there is no monitoring. Similarly, the enforcement response is stronger against firms whose emissions will impact the monitor recording—i.e., firms that are located upwind from a monitor. This suggests that monitors can ensure a more responsive enforcement of regulations, mitigating concerns that our results are driven by a uniform increase in enforcement around all monitors.

Building on the above evidence that local enforcement efforts against firms increase in the vicinity of monitors, we move on to study the pollution monitoring program's citywide effects. The focus on the city level allows us to capture the aggregate impact of the policy (including any within city spillover or displacement effect).[2] By exploiting plausibly exogenous variation in the number of monitors installed in different cities, and thus the share of polluting activity covered by the program, we can assess the impact on total enforcement and pollution. To capture overall pollution changes at the city level, we follow previous literature and use satellite data on the aerosol optical depth (AOD). The AOD data enable us to measure pollution across the whole city both before and after the introduction of monitors and provide us with a reliable data source that cannot be manipulated by local officials. Our analysis exploits the fact that the central government assigned monitors to cities based on their population and geographical size. Using this information, we employ four different empirical strategies: a DiD specification for cities that installed a different number of monitors, an instrumental variable approach that uses the assigned number of monitors, a regression discontinuity (RD) specification that exploits assignment cutoffs, and a difference-in-discontinuity specification that combines the previous approaches. All four empirical strategies produce consistent estimates and show that one additional monitor, which increases coverage of high-pollution activity by about 30 percent, leads to a step-up in enforcement activities by between 15 and 28 percent and reduces pollution by 3.1 to 4.6 percent. We then investigate spillovers and document that while enforcement activities are concentrated in the high-pollution area close to the monitors, pollution is reduced across the city.

Our preferred interpretation of the above results is that monitors improve the central government's ability to hold local officials accountable for their actions. In this setting, local mayors face promotion incentives and are specifically evaluated on their ability to achieve predefined pollution reduction targets set by the central government. To empirically assess the validity of this interpretation, we follow Xi, Yao, and Zhang (2018) and exploit discontinuities in promotion incentives caused by the age of local mayors at the time of the National People's Congress. Estimating our baseline empirical model for mayors facing different promotion probabilities, we find evidence suggesting that monitoring is the most effective when mayors face performance incentives. Hence, this finding is in line with pollution monitoring strengthening top-down accountability and through that making existing performance incentives more effective. Since air pollution information is made available to the public online, an alternative mechanism explaining our results is that monitors improve bottom-up accountability. However, additional analysis suggests that this is unlikely to be the main driver behind our results.[3]

[2] As depicted in panel B of online Appendix Figure D2, cities are large geographical units. Due to the administrative structure in China and the large distance between the urban centers of different cities, we are not concerned about across city spillovers.

[3] As discussed in Section IV, we do not suggest that the monitors did not improve access to information about local pollution to the public (which in turn may have an impact on government enforcement). However, the response in enforcement and pollution to more comprehensive monitoring that we document at the city level does not seem to be driven by increased dissemination of information to the public.

Finally, as discussed above, two reasons why information about policy outcomes may be lacking or of poor quality in low- and middle-income countries are capacity constraints and misreporting. The policy we study is potentially reducing both of these factors at the same time. To shed some light on the relative importance of the two factors, we take advantage of an additional policy shift—the reassignment of control of the monitors from the local government to external third parties. This reassignment decouples the information provision responsibility from the enforcement of regulation responsibility and was conducted after it was discovered that several local governments tried to manipulate the data from the monitors. By exploiting information from the monitors as well as our satellite-based measure of pollution, we show that the monitor recordings are more strongly correlated with the satellite data when they are under the control of a third party—consistent with a reduction in manipulation. Following this logic, we further document that when monitors are under the control of the independent third party, the effect of an additional monitor on enforcement and pollution is substantially larger. This provides suggestive evidence that not only the capacity to collect information but also the way in which this information is provided is important for top-down accountability.

This paper contributes to three strands of literature. First, it relates to a growing empirical literature studying policies aimed at reducing pollution in developing countries. Prior work has documented that regulatory changes can bring about pollution reduction (Greenstone and Hanna 2014; Tanaka 2015; Ebenstein et al. 2017) and that the incentives faced by both local leaders (Kahn, Li, and Zhao 2015) and auditors matter for policy outcomes (Duflo et al. 2013). However, the literature also emphasizes that enforcement of environmental regulations is a major challenge (see, e.g., discussion in Greenstone and Hanna 2014) and that we know little about how to improve it in developing countries (Shimshack 2014). For example, simply increasing the rate of environmental inspections does not seem to have any substantial impact on compliance and environmental outcomes due to the importance of regulatory discretion (Duflo et al. 2018). Our findings suggest that improved monitoring of local pollution—a policy that strengthens top-down accountability without reducing regulatory discretion—could be an effective way of addressing the enforcement gap and reducing pollution. Hence, our work suggests that automatic pollution monitoring could be an effective policy instrument to address high levels of pollution in developing countries. We also relate to two concurrent studies that investigate other dimensions of the same pollution monitoring program (Greenstone et al. 2022; Barwick et al. 2020). Barwick et al. (2020) investigate the impact of sharing air pollution information with the public and show how that leads to avoidance behavior, while Greenstone et al. (2022) study how the updating of monitors in large cities (as opposed to the introduction of new monitors in smaller cities that we study) improved air pollution data quality and reduced the scope for manipulating the data. An additional related concurrent paper is He, Wang, and Zhang (2020), which studies how water pollution monitoring affects firm performance and documents that firms immediately upstream of a water monitor have lower productivity than those immediately downstream. Finally, a literature focusing on the United States documents strategic responses to air quality monitoring (Auffhammer, Bento, and Lowe 2009; Grainger and Schreiber 2019; Zou 2021). Our work complements earlier studies by showing how air pollution monitoring affects the enforcement behavior of local governments and aggregate pollution levels within a fixed regulatory framework. We shed light on the consequences of monitoring for government performance, exploring how the responsiveness and targeting of enforcement is affected. In addition, we investigate the long-term impact on an aggregate and objective measure of pollution (the AOD).[4]

Second, we contribute to an extensive literature showing that monitoring and the provision of information can improve accountability and government performance (Besley and Burgess 2002; Olken 2007; Snyder and Strömberg 2010; Reinikka and Svensson 2005, 2011; Kosack and Fung 2014; Avis, Ferraz, and Finan 2018). While the broader literature has considered the impact of media as well as of audits, we are most closely aligned with recent work showing how information technology affects government performance and efficiency (Duflo, Hanna, and Ryan 2012; Muralidharan, Niehaus, and Sukhtankar 2016; Dhaliwal and Hanna 2017; Banerjee et al. 2020). Proponents of such technological innovations have argued that they could increase efficiency, reduce the scope for manipulation, and be implemented at a relatively low cost. Our study differs from most of the previous work by focusing on monitoring of the final policy outcome (pollution) rather than intermediate inputs in policy production—such as public official attendance (Duflo, Hanna, and Ryan 2012; Dhaliwal and Hanna 2017) or transfer of funds (Muralidharan, Niehaus, and Sukhtankar 2016; Banerjee et al. 2020). While the monitoring of final policy outcomes might not always be feasible, it could mitigate concerns about multitasking (Holmström and Milgrom 1991) associated with intermediate monitoring. We show that policy outcome monitoring can indeed be effective in the context of pollution. In addition, we expand prior work by studying how enforcement of regulations as opposed to public service provision is affected by monitoring. Third, we relate to a literature investigating the potentially distorting effect of high-powered incentives on data reporting (Banerjee, Duflo, and Glennerster 2008; Fisman and Wang 2017; Acemoglu et al. 2020), including manipulating pollution data (Andrews 2008; Chen et al. 2013; Ghanem and Zhang 2014; Oliva 2015). We contribute to this literature by studying how control over the information infrastructure (shifting from local governments to external firms) is correlated with the quality of information as well as government actions and actual policy outcomes. While we are cautious when interpreting the results from this analysis due to the strong assumptions required for causal inference, it has the benefit that we can observe both potentially manipulated data from monitors and satellite data independent of government influence (and therefore also policy impact).

The paper is structured as follows. Section I describes the context as well as the implementation of the pollution monitoring program we investigate. Then the

[4] He, Wang, and Zhang (2020) use a spatial regressions discontinuity design to investigate the impact on *self-reported* water pollution by firms (i.e., the analysis does not capture potential misreporting by firms and aggregate pollution effects including nonmonitored firms). Greenstone et al. (2022) use a temporal RD design to study the short-term impact on air pollution from the upgrading of monitors. Auffhammer, Bento, and Lowe (2009) document larger reductions in pollution close to monitors that violate air quality standards. Grainger and Schreiber (2019) document the strategic placement of air pollution monitors. Zou (2021) shows that there is a temporal reduction in pollution as measured by satellite data during days when monitors under the Clean Air Act are turned on.

data used in this study is described in Section II. The first analysis, which explores firm-level evidence on enforcement, is presented in Section IIIB. The aggregate effect of monitoring on enforcement and pollution at the city level is discussed in Section IIIC. These two sections present both the respective empirical strategies and results. The analysis of the mechanisms is discussed in Section IV. Finally, Section V offers concluding remarks.

I. Background and Policy Description

This section provides background information and describes the context in which the national monitoring program studied in this paper was introduced. In Section IA, we describe the environmental policies in place in China during this period and discuss the local leaders' role in achieving them. After that, the infrastructure put in place to monitor these policies' implementation is described in Section IB. Section IC discusses how the information presented in this section guides our analysis.

A. *Environmental Policies in China*

While the Chinese government's priority during the past decades has largely been to stimulate economic growth, attention has lately shifted toward environmental policies (Zheng and Kahn 2017).[5] Starting in 2013, the National Air Quality Action Plan was set up to improve air quality by the end of 2017. As a part of China's successful "war on pollution" (Greenstone and Schwarz 2018), the plan laid out the general ambition for the whole country and set differentiated goals for each region. In January 2014, the Ministry of Environmental Protection (MEP) entered into "contracts" with all 31 provinces and set up a three-year air quality plan to decrease the concentration of particulate matter (PM) in the whole country. In each "contract," an air quality target for 2017 was set—resulting in different percentage reduction targets of $PM_{2.5}/PM_{10}$ for each province relative to the 2012 level.[6]

These centrally set targets are implemented by local government officials, who are incentivized to fulfill them through performance-based promotions. Promotions are the key instrument used in China to ensure that local officials carry out policies in line with the goals set by the central government (see Zheng and Kahn 2013, 2017, for further discussion of this topic). For a long time, the central government focused on economic performance and emphasized economic growth as the key evaluation criteria for local officials' promotion (Chen, Li, and Lu 2018). However, from the Twelfth Five-Year Plan onward, the central government has used the fulfilment of environmental performance targets as a requirement for the promotion of local mayors (Zheng and Kahn 2013).

[5] The concentration of air pollutants in China is among the world's highest and is a problem with serious health consequences. Average $PM_{2.5}$ (particulate matter with a diameter of 2.5 micrometers or less) concentrations in 2013 were 91 micrograms per cubic meter ($\mu g/m^3$), which is nine times the amount the World Health Organization considers safe. Estimates by Greenstone and Schwarz (2018) suggest that if these levels of pollution are sustained, it will result in a 6.5 year decline in life expectancy for the average resident.

[6] For the list of targets by province, see Table C3 in online Appendix C.

B. *National Monitoring System*

To address issues raised about limited coverage and quality of existing pollution data, the central government introduced a new monitoring system as a part of its 2013 National Air Quality Action Plan. This new system expanded coverage to all of China—introducing monitors in prefecture-level cities that previously had no systematic air pollution monitoring in place. In addition, cities with existing monitors received new updated monitors that could capture the wider range of pollutants included in the revised air pollution standards (notably, $PM_{2.5}$, widely regarded as the key measure of ambient air pollution, was included for the first time). One of the key features of the new system is that all monitoring stations report six pollutants (SO_2, NO_2, CO, PM_{10}, $PM_{2.5}$, and O_3) to the central government in real time (Greenstone et al. 2022). Hourly pollution data are then automatically published online by the central government.

The new monitors were installed in three separate phases. The first phase was conducted in 2013 and focused on 74 major cities that represented the country's key population and economic centers.[7] As part of the National Air Quality Action Plan, these cities were simultaneously targeted by a number of additional policies aimed at reducing air pollution (MEP 2013).[8] The second phase was implemented in 2014 for the 87 designated "environmental role model cities," which face stricter and more frequent evaluation of environmental performance (MEP 2011; Brehm and Svensson 2020). The primary aim of the first two phases was to automate old manual monitors, since 70 percent of these cities already had pollution monitoring in place. The main expansion phase, which is the one we focus on in this paper, was carried out in the following year when all 177 remaining prefecture-level cities (53 percent of all prefecture-level divisions in China) received monitoring for the first time. After this final expansion, all prefecture-level cities had at least one air quality monitor. These monitors all started transmitting information to the central government on January 1, 2015.

The funding for the monitors was provided by the province-level environmental bureaus. Once all equipment had been put in place, the city-level environmental bureau was made responsible for the maintenance and operation of all monitors within the city. The local governments, which have incentives to report low levels of pollution because of the performance targets they face, could potentially do this by manipulating the recordings from the monitors. Such manipulation was facilitated by the direct control of the monitors that the local governments were given. Indeed, several media sources have reported that such manipulation did occur (e.g., by spraying the monitor with water to reduce the recordings).[9]

[7] The Beijing–Tianjin–Hebei metropolitan region, the Yangtze River Delta, the Pearl River Delta, directly administered municipalities, and provincial capitals.

[8] This included, e.g., stricter pollution reduction targets by 2017, prohibiting the construction of new coal-fired power stations from 2015, different motor gasoline standards from 2013, reduced reliance on coal and increased construction of natural gas infrastructure, setting up regional environmental impact assessment and joint law enforcement, etc.

[9] See https://p.dw.com/p/32jqR and http://www.xinhuanet.com/politics/2018-08/09/c_1123244676.htm, for two examples.

Realizing that the data provided by local environmental protection bureaus might not be reliable, the MEP decided to contract the operation of the monitor stations to private companies through a procurement process. According to official documents from the MEP, all of the monitors were operated by private companies on November 1, 2016. Monitors were procured through 12 contracts. Each contract was designed to involve monitors in different provinces spread out over the country, to make it difficult for firms to select a given area. Six companies were selected, and each of them won two contracts. These firms were then paid directly by the MEP to operate the monitors.

In addition to the regular monitors in the built-up area of each city, half of the cities were also assigned one background monitor. There are two main differences between the background monitors and the regular monitors: background monitors are installed outside of the built-up area of the city and are usually placed in a local scenic area; more importantly, the readings from the background monitors are not used in the performance evaluation of local officials. Due to the different nature of the background monitors, we are not including them in the main analysis.[10]

C. *Conceptual Framework and Sample Selection*

As discussed in the previous section, the central government regulates (e.g., sets pollution standards), while the local government is responsible for enforcing these regulations (e.g., by issuing fines to firms' violating existing regulations). Our interest is in understanding to what extent the introduction of monitors helps the central government hold the local government accountable for their actions and how that affects enforcement behavior and pollution at the local level. To capture this effect we focus on the 177 cities that face the same regulations and receive monitoring for the first time.[11] Online Appendix Figure D1 illustrates how the introduction of monitors changes access to information on pollution both within and between cities. Within cities, monitor readings will be mostly affected by firms located close to and upwind from a monitor.[12] Between cities, information on a larger share of high-pollution activity will be available for those cities that installed a greater number of monitors (as shown in Figure 3).

Hence, the monitoring program that we study changes the capacity of the central government to collect information about pollution. This capacity changes both

[10] Including them in the analysis does not alter any of our results. This is due to the fact that there are a limited number of firms located close to the background monitors. We also check robustness of our main results to controlling for whether a city has a background monitor.

[11] We exclude from our analysis the cities that received a monitor for the first time before 2015 since they were specifically targeted and simultaneously affected by other central government policies aimed at reducing air pollution (as discussed in Section IB). This type of targeting is common practice in China (Wang and Yang 2022). Wang and Yang (2022) emphasize that focusing on areas where policies are initially implemented risks leading to biased estimates of policy impact because of the way these areas are selected and the incentives faced by leaders exposed to early policy experimentation.

[12] There is no exact cutoff for how far away from the monitor pollution could be picked up. For example, anecdotal evidence discussed in online Appendix D suggests that environmental officials are concerned with pollution from firms within 5 km of a monitor. Schlenker and Walker (2015) show that health effects can be picked up 20 km from a polluting source, suggesting that a monitor would be able to pick up differences at such a distance. We take a flexible approach in our analysis and let the data inform us about this cutoff.

at the extensive margin (covering some firms but not others) and at the intensive margin (covering a larger versus smaller share of polluting activity). In addition to the change in monitoring capacity in 2015, the reassignment of monitors from the local government to external third parties in 2016 changes the information provision process and decouples the responsibility of providing information with the responsibility to enforce regulations. The intention of the central government is that this shift should improve data quality and reduce the scope for manipulation. Because third parties are paid directly by the MEP, their incentives are arguably more aligned with those of the central rather than the local government. In our analysis we will mainly focus on the overall effect of the monitors. However, in Section IVB we will shed some light on the potential importance of who is responsible for information provision.

II. Data

In this article, we combine several data sources that provide comprehensive information on the enforcement of environmental regulation and air pollution performance in cities that introduced air pollution monitors in 2015 (MEP 2014). Section IIA describes the new data on local air pollution enforcement that we collect and digitize. After that, Section IIB describes the two sets of data that we use to measure air pollution: a satellite-based measure of the AOD and data from the monitoring stations. Finally, Section IIC discusses the summary statistics for our three main samples. Additional details on data processing and on supplementary datasets used are provided in online Appendix A.

A. *Enforcement Records and Firm Data*

To fully understand the impact of new air quality monitors on enforcement activities and the consequences of those activities, we face some data-related empirical challenges: first, the need to measure the quantity (and the quality) of governments' enforcement activities and, second, the need to link enforcement activities to the location of air quality monitors. We address these challenges by constructing a new dataset on local enforcement of air pollution regulation in China using records collected from local environmental bureaus by the Institute of Public and Environmental Affairs (IPE) (IPE 2017). To the best of our knowledge, this is the first attempt to fully track enforcement activities carried out by local environmental bureaus in China. To identify where these enforcement activities occur, we georeference all major manufacturing firms in China using the Annual Survey of Industrial Firms (ASIF) (National Bureau of Statistics 2013) and link these to the IPE records.

Enforcement Records.—We collected all 55,184 enforcement records carried out from 2010 to 2017 in the 177 prefecture-level cities in our sample. Figure A1 in online Appendix A provides an example of what these records look like and the type of information they contain. Each record includes details about the violating firm, a description of the violation, a reference to the regulation that has been violated, and the local environmental bureau's enforcement action. Using a classification algorithm described in detail in online Appendix A.1, we categorize enforcement records in two dimensions. First, we identify what type of violation has been logged and whether this relates to air pollution, water pollution, waste pollution, or procedural violations. In total, we classify 24,691 records as being related to violations of air pollution regulations. Second, we identify what type of action has been taken by the local environmental bureau and in which quarter and year it was carried out. For 95 percent of the enforcement records related to air pollution, the actions belong to one or several of the following four categories: suspending production (52 percent), ordering replacements/upgrading of the equipment (54 percent), levying fines (48 percent), or issuing a warning (15 percent).

Firm Data and Georeferencing.—To be able to track where and against which firms local environmental bureaus choose to enforce regulations, we use data from the 2013 ASIF. This survey is conducted by the National Bureau of Statistics. It includes all state-owned industrial enterprises and all private industrial enterprises with annual sales exceeding ¥5 million. This corresponds to about 90 percent of all manufacturing firms in China and thus covers all major industrial polluters.[13] Previous versions of the ASIF data have been used in a number of papers (see, e.g., Song, Storesletten, and Zilibotti 2011; Brandt, Van Biesebroeck, and Zhang 2012; Huang et al. 2017). We focus on the 2013 version of the survey, which is the latest available, and restrict our sample to firms that started operating before 2010 (the first year of our analysis). This allows us to gain an understanding of the underlying distribution of manufacturing firms. Before linking the data to the enforcement records, we use detailed firm address information to identify the exact geographical location of all firms in the data. The process used for this georeferencing is outlined in online Appendix A.1. Panel C in online Appendix Figure D2 shows the location of all the ASIF firms in our sample. Finally, we link our collection of enforcement records to the underlying distribution of manufacturing firms in the ASIF. Out of our 55,184 records, 52 percent of them refer to enforcement actions against firms in the ASIF data. Panel D in online Appendix Figure D2 shows the geographical distribution of enforcement activities against these manufacturing firms.

B. *Air Pollution Data*

Monitor Data ($PM_{2.5}$, PM_{10}, and AQI).—Air pollution data for the 552 monitoring stations in the 177 prefecture-level cities in our sample are published online by the MEP from the introduction of the monitors in January 2015 (MEP 2017).[14] The MEP website reports hourly data of SO_2, NO_2, CO, PM_{10}, $PM_{2.5}$, and O_3. An air quality index (AQI) based on these six pollutants is also constructed and reported.[15] The AQI ranges from 0 to 500. It is further divided into six

[13] According to the 2004 economic census, firms in the ASIF represent 89.5 percent of the total revenue of all manufacturing firms in China.

[14] The data can be accessed via this link: http://106.37.208.233:20035/.

[15] The AQI is calculated using the following equation: $AQI = \max\{IAQI_1, IAQI_2, \ldots, IAQI_6\}$, where each Individual Air Quality Index (IAQI) is given by $IAQI_i = \frac{I_h - I_l}{C_h - C_l}(C - C_l) + I_l$. The formula to compute IAQI is

ranges: 0–50, 51–100, 101–150, 151–200, 201–300, and 301–500. In public reports, these are categorized as excellent, good, lightly polluted, moderately polluted, heavily polluted, and severely polluted, respectively. We scrape pollution data from the MEP website and focus on the two main indicators used as targets in the National Air Quality Action Plan (PM_{10} and $PM_{2.5}$) as well as the AQI. To facilitate comparison with our other pollution measure described below, we aggregate the monitor data at the monthly level.

Satellite Data (*AOD*).—Before the expansion of the monitoring system, none of the cities in our sample had any consistent pollution monitoring. To obtain an objective measure of pollution both before and after monitor construction, we use data on AOD captured by the NASA MODIS satellites (NEO 2017). AOD measures the degree to which aerosol particles prevent the transmission of light by absorption or scattering and can therefore be used as a measure of local pollution. Formally, AOD is defined as the negative of the natural logarithm of the fraction of radiation (e.g., light) that is not scattered or absorbed. Hence, estimates of AOD in this paper can be interpreted as percentage changes. Monthly information on AOD is available at 0.1 by 0.1 degrees. We combine measures from the MODIS Aqua and Terra satellites to calculate the mean of AOD in a given month. We aggregate AOD at different geographical levels—ranging from the pixel intersecting with the monitor (approximately 11 km $\times$ 11 km), to the city center around the monitors (10 to 50 km from a monitor) and surrounding areas (beyond 50 km of a monitor). To deal with potential within-city spillovers in pollution, our baseline measure is based on the whole prefecture-level city polygon (ESRI 2017) (as depicted in panel B of online Appendix Figure D2). This figure shows the distribution of average AOD in 2010, the first year of our analysis, across all cities in our sample. As indicated in the figure, there is substantial cross-sectional variation in pollution in our sample.

AOD has been shown to be highly correlated with ground-based measures of pollution (see, e.g., Wang and Christopher 2003; Gupta et al. 2006).[16] While AOD data have been used in various studies to measure air pollution (see, e.g., Chen et al. 2013; Jia 2017), only a few studies have internally verified the correlation between AOD and local ground-based measures. To ensure the validity of AOD data in our setting, we take advantage of the ground-based measures of pollution that are available after the expansion to study the correlation between the AOD data and the two most common measures of air pollution ($PM_{2.5}$ and PM_{10}) as well as the joint AQI. In online Appendix Table C4, we report results from regressions controlling for monitor fixed effects and time fixed effects as well as precipitation, temperature, and mayor's age. Column 1 shows the estimate for $PM_{2.5}$, which is 0.30. This is largely comparable with the correlations reported by Gupta et al. (2006). Estimates for

the same one used in the United States, but with differences in parameters (C_h, C_l, I_h, and I_l). C is the pollutant concentration measured by the air quality monitor. C_h and C_l are the concentration breakpoints, and I_h and I_l the index breakpoints. More details about these parameters can be found here: https://www.mee.gov.cn/ywgz/fgbz/bz/bzwb/jcffbz/201203/t20120302_224166.shtml.

[16] Wang and Christopher (2003) find that the correlation coefficient between the monthly means of AOD and $PM_{2.5}$ is around 0.7 using data in Alabama in 2002. Using much more comprehensive data, Gupta et al. (2006) find that the correlation ranges from 0.14 to 0.6 for a number of cities across the world.

PM_{10} and AQI are smaller but of a broadly similar magnitude. Taken together, this suggests that AOD is a suitable measure for local air pollution and that it most strongly reflects changes in $PM_{2.5}$.

C. *Main Sample and Summary Statistics*

To supplement our analysis, we collect additional data on monthly weather conditions (CMA 2017), quarterly dominant wind direction (CMA 2017), résumés of all mayors during our sample period (Jiang 2017), and city-level aggregates of citizens' online searches for a set of keywords related to pollution (Baidu 2017). Online Appendix A.3 describes this additional data in detail and the procedure used for collecting it. Using the data on pollution and enforcement described above together with these additional sources, we construct three main samples for our analysis—all covering the 177 prefecture-level cities that installed monitors in 2015. Summary statistics for these three samples are presented in online Appendix Table C1.

Panel A reports information for the firm-level data. We rely on the 2013 ASIF and restrict the sample to firms that started operating before 2010 (the first year of our analysis) and that are located within 50 km of an air quality monitor.[17] This leaves us with a total sample of 36,103 firms. The majority of these firms are private (81 percent) and cover a wide range of different industries (online Appendix Table C5 reports the industry composition for our sample). On average, the firms in our sample are located 19 km from a monitor. However, as depicted in online Appendix Figure D9, the spatial distribution of firms is skewed, and 40 percent of firms are located within 10 km of a monitor. For a given firm in our sample, the probability of receiving an air-pollution-related enforcement action in a quarter is 0.5 percent. Such an enforcement action most commonly requests that the firm upgrade their equipment, but suspension of operation and issuing fines are also common. Violations relating to water pollution regulations or conducting a procedural violation are of a comparable magnitude (0.29 percent and 0.52 percent, respectively). Most (more than 75 percent) of the enforcement actions were taken after the introduction of air quality monitors.

Panel B reports the summary statistics for the city-level data. For this sample we consider pollution as well as enforcement at the aggregate city level.[18] The cities we study are small by Chinese standards and have an average population of around 340,000. The average size of our sample (measured by both the urban population and the size of the built-up area) are one-third of the country average. However, the air pollution level in our sample (measured by AOD) is only slightly lower (10 percent) than the country average.[19] On average the cities in our sample have 2.8 monitors installed, and about 4.2 firms face an environmental enforcement action related to air pollution per quarter. Out of these 4.2 firms, 1.5 are ASIF firms. Panel C reports

[17] Note that while we have quarterly information on enforcement actions, our information on firm characteristics is from the 2013 ASIF and therefore cross-sectional.

[18] Hence, this sample is not restricted to firms within 50 km of a monitor and covers the whole city polygon, as depicted in online Appendix Figure D2.

[19] Online Appendix A.2 discusses additional details regarding the representativeness of our sample and compares it to other cities in China.

the summary statistics for the monitor-level data for the three pollution measures we use. These data are aggregated at the monitor-month level and cover data from all 552 monitors installed in the 177 cities that we study. The sample period for these data starts in January 2015, when all the monitors have been installed.

III. Impact of the Monitors

This section describes the impact of monitoring on local government enforcement activities and pollution. We start by discussing what determines where monitors are installed in Section IIIA, since this is key for our empirical analysis. Then, we present our firm-level empirical strategy and the corresponding results in Section IIIB. In Section IIIC, we move from studying the local effects of monitoring to the aggregate city effects—exploiting differences in the number of monitors installed.

A. *Assignment and Location of Monitors*

The MEP provided detailed instructions for how many monitors that should be installed in each city and where they should be located. All the monitors were installed in the so-called "built-up area"—the main urban center of the prefecture-level city. The number of monitors assigned to each city was determined by the city's population size and the geographical size of the built-up area. Cities are assigned to one of four groups (corresponding to the installation of one, two, four, or six monitors), determined by the criteria that assigns it to the highest group. The detailed assignment criterion, which we use for identification in our city-level analysis, is presented in online Appendix Table C2.

The location of these monitors is depicted in the map in panel A of online Appendix Figure D2. While the exact parameters used by the central government for deciding the precise geographic location of each monitor are unknown, official government documents report that the location was chosen by a simulation method that took surrounding buildings, traffic, and the direction of seasonal winds into account to make sure that the monitors captured a fair representation of local pollution. To shed light on what determines the placement decision, we investigate how the location of monitors relates to the spatial distribution and trends in enforcement activities and pollution.

We start by investigating the spatial distribution of enforcement activities and how these change with the introduction of monitors. Figure 1, panel A shows a binned scatterplot of the probability that a firm has any enforcement record related to air pollution in a quarter by the distance to the closest monitor. Black dots indicate the mean probability during the period before air quality monitors were introduced, and red diamonds show the mean probability in the post-period. The graph shows that the average quarterly probability of a firm receiving any air-pollution-related enforcement action is around 0.0021 before 2015, and that this probability does not seem to depend on the distance to the (planned) monitor (i.e., there is no gradient in enforcement activity in the pre-period). This provides some first evidence suggesting that monitors are not endogenously placed in localities with differential levels of enforcement. Figure 1, panel B shows another binned scatterplot that instead plots the trends in enforcement

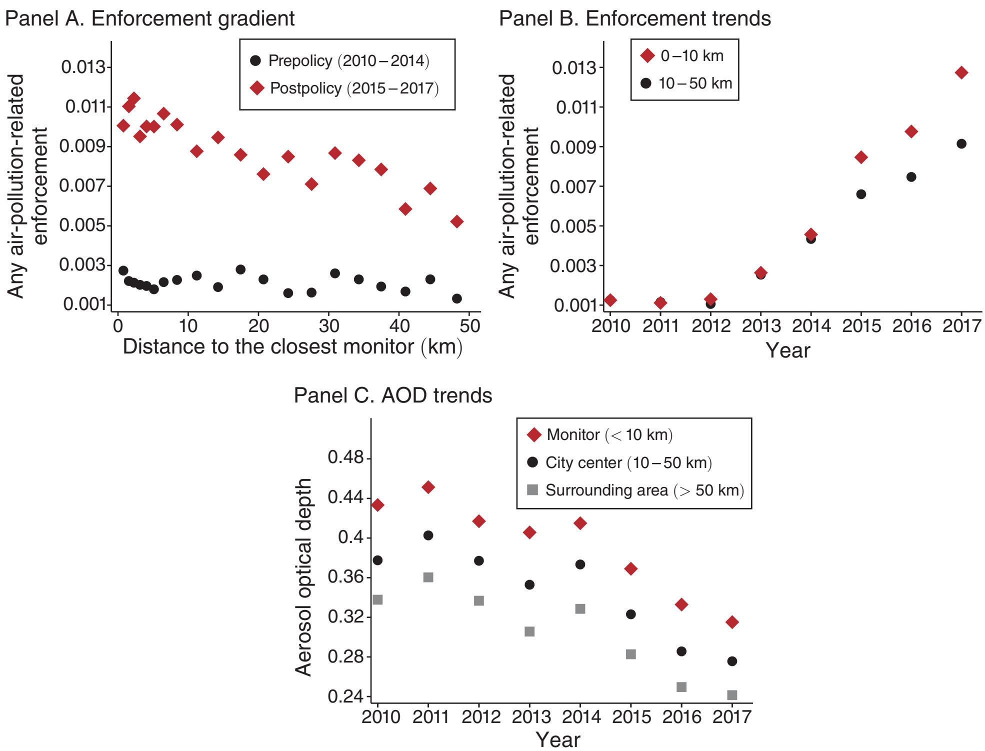

Figure 1. Monitor Location: Air-Pollution-Related Enforcement and Pollution

Notes: Panel A shows a binned scatterplot of the relationship between enforcement activity and distance to the closest monitor. Black dots indicate the average quarterly probability of air-pollution-related enforcement before introducing the air quality monitors, while red diamonds show the probability after the introduction of monitors. Panel B shows a plot of the average share of firms with any air-pollution-related enforcement for our treatment and control firms (within and beyond 10 km of a monitor) over time. Panel C shows a plot of the average AOD over time in three areas: the pixel where the monitor is eventually placed, the city center (10–50 km from a monitor), and the areas surrounding the center (> 50 km from a monitor) but within the boundaries of the prefecture-level city.

for our treatment and control firms (firms located within/beyond 10 km of a monitor). The graph shows that both groups of firms face similar trends as well as levels of enforcement in the years before the monitors are installed.

Next, we investigate the pollution level and trends (measured by AOD) in areas where monitors are eventually installed and compare it with surrounding areas. Due to the nature of the AOD data, we are not able to use as fine-grained spatial information as in the above analysis, but instead compare the AOD in the pixel where the monitor is located (which are approximately 11 km × 11 km) to the AOD in the city center beyond the monitor (10–50 km from a monitor) and the area surrounding the city center (beyond 50 km of a monitor). The first two groups roughly correspond to the treatment and control groups in Figure 1, panel B. Figure 1, panel C shows that the pollution level in the pixels where the monitors are located is above average in both the city center and surrounding area (suggesting that the central government's algorithm is successfully targeting high-pollution areas). However, importantly these areas were on a similar trend to both other areas within the city center as well as the surrounding areas—suggesting that monitors are not placed in areas within the city based on local trends in pollution.[20]

B. *Firm-Level Evidence*

Figure 1, panels A and B suggest that there was a significant increase in enforcement activity against firms located close to a monitor once the monitors were installed in 2015. The results are consistent with extensive media reporting that local environmental bureaus step up environmental inspections close to the monitors. We document some of this evidence in Figure D10 in online Appendix D, which shows a list of news articles generated from a search on the Chinese search engine Baidu using the keywords "monitors," "surrounding area," and "check." The list includes a large number of articles discussing how local governments organize their environmental inspections around the monitors. Some examples[21] include cities that draw special zones around their air quality monitors and send teams of inspectors to those zones, to ensure that firms comply with national environmental regulations. Other sources mention that city governments hire volunteers from the public to inspect air pollution from venues within a certain distance from the monitors. Finally, several sources[22] suggest that mayors take a special interest in these inspections by, for example, directly appointing officials to this task or by visiting surrounding areas. This further underlines the weight that mayors put on the recordings from the monitors because of the performance incentives that they face.

Firm Level: Event Study.—To investigate the relationship between monitors and enforcement formally, we estimate a flexible nonparametric event study specification. If we denote a generic firm by i, with $i \in j,p$, where j denotes a four-digit industry, p a province, and t a generic quarter, our model can be written as

$$y_{ijpt} = \delta_i + \theta_{jt} + \eta_{pt} + \sum_{\substack{d=0\text{–}5\text{ km} \\ d\neq 20\text{–}50\text{ km}}}^{15\text{–}20\text{ km}} \sum_{\substack{k=\text{Q1}/2010 \\ k\neq \text{Q4}/2014}}^{\text{Q4}/2017} \beta_{dk} m_i^{dk} + \epsilon_{ijpt}, \tag{1}$$

where y_{ijpt} is an indicator for enforcement; δ_i is a firm fixed effect; θ_{jt} and η_{pt} represent, respectively, industry-by-time and province-by-time fixed effects; m_i^{dk} is an indicator for any monitor being within d km of a firm in quarter k; and ϵ_{ijpt} is the error term. Because we condition on firm as well as on industry-by-time and province-by-time fixed effects, parameter estimates capture the average (across industries and provinces) effect of monitoring on the differential change in enforcement across firms in the same industry or province. This specification addresses two important concerns. First, we ensure that we estimate the impact of monitoring within the same regulatory environment (pollution reduction targets vary across

[20] The pattern shown in Figure 1, panel C is confirmed by a formal test of the pre-trends, which shows no significant differences between the three areas in the five years prior to monitor installment.
[21] www.163.com/dy/article/H4HC2IIH0534B975.html.
[22] newpaper.dahe.cn/dhb/html/2017-12/28/content_212745.htm.

provinces as discussed in Section I). Second, we allow for different enforcement trends depending on local industrial composition at baseline. We use the quarter before the introduction of the monitors and firms 20–50 km from the monitor as reference categories and estimate β_{dk} for $d \in \{0\text{–}5\text{ km}, 5\text{–}10\text{ km}, 10\text{–}15\text{ km}, 15\text{–}20\text{ km}\}$. Equation (1) allows us to estimate the temporal and spatial relationship between monitors and enforcement activity. Hence, it is informative about the key identification assumption for our analysis (parallel trends in enforcement for firms located at different distances from the monitors) as well as the spatial reach of monitors. We cluster standard errors at the city level to account for correlation of errors across firms and time within cities.[23]

Figure 2 reports the results from estimating equation (1). We present the estimates in four separate event study graphs, each showing how enforcement activity changes around the introduction of monitors for firms within 0–5 km, 5–10 km, 10–15 km, and 15–20 km of the monitors relative to firms 20–50 km from the monitors (the reference category). In all four graphs, there is no evidence of any differential trends leading up to the intervention—lending credibility to the main identification assumption of parallel trends. After the introduction of the monitors, we see a substantial increase in enforcement activity close to the monitors. This step-up in enforcement is particularly pronounced within 0–5 km of the monitors but is noticeable also for firms 5–10 km from the monitors. For firms 15–20 km from the monitors, there is no differential change in enforcement activity during our sample period.[24]

Firm Level: Main Results.—Guided by the results in the previous section, we use a simplified DiD specification to provide an estimate of the magnitude of the effect. This specification compares firms within and beyond 10 km of a monitor, but our results are robust to using 15 km or 20 km instead. Formally, we estimate

$$y_{ijpt} = \delta_i + \theta_{jt} + \eta_{pt} + \beta m_{it}^{10km} + \epsilon_{ijpt}, \tag{2}$$

where m_{it}^{10km} is an indicator for a firm having a monitor within 10 km. All other variables are the same as in equation (1). The results from estimating this specification are shown in Table 1. The first column of panel A reports estimates on whether any air-pollution-related enforcement took place (i.e., the same outcome as in Figure 2). Results suggest that the probability of a firm within 10 km of a monitor receiving an enforcement action in a quarter is 0.33 percentage points higher compared to firms farther away from the monitor. This suggests that a monitor increases the probability of an air-pollution-related enforcement activity occurring by 72 percent compared to the average quarterly probability of enforcement (0.46). The remaining columns of panel A in Table 1 shed light on what type of action was taken by the local

[23] As a robustness check, we also report standard errors based on the spatial HAC variance estimator proposed by Conley (1999), following the implementation suggested by Hsiang (2010) and Fetzer (2020), which allows for correlation between areas that are geographically close but belong to different administrative units (see panel A of Table 1). These standard errors are smaller, but overall similar, to our baseline standard errors. We focus on the city-level clustered standard errors since these are more conservative.

[24] In online Appendix Figure D3, we estimate the aggregate change in the enforcement gradient using a simplified version of equation (1), which reconfirms the main results reported in Figure 2.

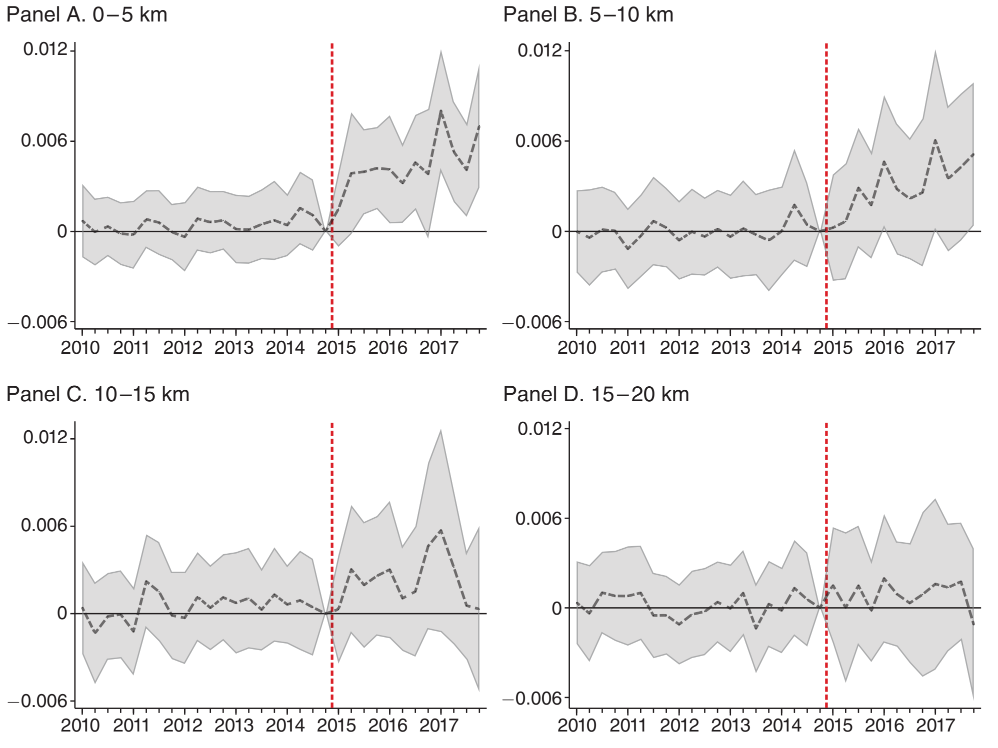

Figure 2. Firm Level: Nonparametric Event Study

Notes: The figure shows the estimates of the nonparametric event study using equation (1). The subfigures report event studies for firms within each distance bin. The reference group is firms located 20–50 km from the closest monitor. The shaded area represents 95 percent confidence intervals, calculated using robust standard errors clustered at the city level.

government, by estimating the same model for the four most common enforcement classifications we identify in the data ("suspension": suspending production for the firm, "upgrading": ordering replacement/upgrading of equipment, levying a "fine," or issuing a "warning"). We find similar estimates for the first three categories and no effect for the last type ("warning"). These results suggest that the local environmental bureau is responding to the monitors by implementing costly punishments on local firms.

Firm Level: Targeting.—In the previous subsection, we showed that enforcement increases following the introduction of monitors. In this subsection we will explore whether monitoring also affects other aspects of enforcement activities. We do this in two ways. First, we study which firms were targeted by the local governments and whether the intensity and strictness of enforcement that they face change with the introduction of monitors. Second, we investigate how monitors shape the responsiveness of enforcement actions to local pollution shocks. To better understand which firms local governments target and whether this targeting changes with the introduction of monitors, we study actions against a set of high polluting firms.

Table 1—Firm Level: Pollution Monitoring and Enforcement Activities

Outcome	Any Air (1)	Suspension (2)	Upgrading (3)	Fine (4)	Warning (5)
Panel A. Any enforcement action related to air pollution					
$Mon_{<10km} \times Post$	0.0033	0.0014	0.0014	0.0014	−0.000058
	(0.00056)	(0.00045)	(0.00041)	(0.00043)	(0.00016)
Mean outcome	0.0046	0.0024	0.0025	0.0022	0.00070
Observations	1,155,296	1,155,296	1,155,296	1,155,296	1,155,296
Conley SE	[0.00040]	[0.00031]	[0.00031]	[0.00030]	[0.00017]
Outcome	# Air	Low intensity	High intensity	Lenient	Strict
Panel B. Intensity and strictness of enforcement action related to air pollution					
$Mon_{<10km} \times Post$	0.0031	0.0027	0.00018	0.00028	0.00070
	(0.00064)	(0.00047)	(0.00013)	(0.00015)	(0.00034)
$Mon_{<10km} \times Post \times H.\ Polluter$	0.040	0.0017	0.014	−0.0037	0.022
	(0.011)	(0.0068)	(0.0028)	(0.0016)	(0.0058)
Mean outcome	0.0052	0.0042	0.00040	0.00065	0.0016
Observations	1,155,296	1,155,296	1,155,296	1,155,296	1,155,296

Notes: This table reports estimates of the impact of air pollution monitoring on the probability of being subject to different air-pollution-related enforcement actions by the local government. All regressions control for fixed effects specific to firm, industry-by-time, and province-by-time interactions. Robust standard errors clustered on the city in parentheses. In panel A, standard errors based on the spatial HAC technique suggested by Conley (1999) are reported in brackets, using a Bartlett kernel and bandwidth of 100 km. Panel B reports heterogeneity for firms identified as high polluters according to ESR during the pre-period. The outcome "low intensity" ("high intensity") corresponds to a dummy variable indicating that a firm received only one (at least two) enforcement actions in a quarter. The outcome "lenient" is a dummy variable that equals one if only one punishment (among "suspension," "upgrading," and "fine") is issued against a firm in a quarter. In contrast, the dummy variable "strict" is defined as one if all three types of punishments are issued against a firm in a quarter.

We rely on the Environmental Survey and Reporting Database (ESR) to identify these firms. The ESR is put together by the central government and includes firms that are considered to be major polluters (in total responsible for 65 percent of local emissions).[25] In panel B of Table 1, we estimate the differential enforcement response against these firms. The estimates in column 1 suggest that there is a larger increase in the number of enforcement actions against these firms (significant at the 1 percent level). The following four columns report what type of enforcement these firms receive. We start by differentiating between low and high enforcement intensity, where we define "low" as receiving one enforcement action in a quarter and "high" as receiving more than one action. The results in columns 2–3 show that low-intensity enforcement is not significantly different between low- and high-polluting firms, but that all high-intensity enforcement focuses on key polluters in the presence of monitoring. Next, we consider the strictness of enforcement action. To capture this, we construct two additional dummy variables that classify enforcement records as either lenient or strict. Since there is no clear ranking of the three main punishment types discussed above ("suspension," "upgrading," and "fine") and enforcement records often include multiple punishments, we consider the two

[25] The ESR database has been used in several recent paper (see, e.g., He, Wang, and Zhang 2020). We use the ESR firms identified between 2010 and 2014, the period before introducing air quality monitors. In total, this corresponds to 1,445 of the firms in our baseline firm sample.

extreme cases where either one (lenient) or all three (strict) punishments are issued against a firm in a year. The last two columns in Table 1 report the results and show that high-polluting firms are less likely to receive lenient treatment and more likely to receive strict enforcement action. Taken together, these results suggest that local governments respond to monitoring by shifting both the intensity and the strictness of enforcement toward high polluters.

Firm Level: Responsiveness.—Next, we investigate whether monitors make local governments' enforcement efforts more responsive to recorded pollution. The main empirical challenge inherent in studying this is the endogeneity of local pollution. To overcome this challenge, we exploit two different sources of plausibly exogenous variation: local rainfall shocks and wind direction.

We start by investigating rainfall shocks. First, we establish that local rainfall shocks are important determinants of local pollution. For each city, we construct an indicator ($Rain_{>\tilde{x}}$) for whether the quarterly rainfall is above the median for that city or not. Online Appendix Table C6 shows estimates of the relationship between the average pollution recordings across the monitors in the city and this high rainfall indicator, controlling for city and time fixed effects. Results show that average pollution recordings are consistently about 7–9 percent lower in quarters with above-median rainfall. These effects are substantially stronger at higher levels of pollution, which are arguably more important for local policy response, where, e.g., the share of days that have any reading classified as heavily polluted ($AQI > 200$) is reduced by 22 percent ($0.024/0.11$). We then explore how enforcement activities respond to pollution monitoring in the presence of rainfall shocks by estimating an augmented version of our baseline model (equation (2)) that introduces interactions with the quarterly rainfall shock.[26]

The results from this analysis are reported in panel A of Table 2. Column 1 adds the rainfall shock and its interaction with our main treatment variable (m_{it}^{10km}). Two main takeaways emerge. First, enforcement activities do not respond to rainfall shocks in the absence of monitors—suggesting that there is no direct effect of rainfall on enforcement. Second, the effect of monitors on enforcement is substantially smaller ($0.0043 - 0.0018 = 0.0025$) when rainfall levels are above the median (i.e., when pollution levels are lower) as opposed to when rainfall levels are below the median (0.0043) (i.e., when pollution levels are higher). These results suggest that the information captured by the monitors is important for the enforcement actions taken by the local governments and that the monitors make the local government responsive to changes in local pollution. To investigate whether this result reflect a general increase in the enforcement responsiveness in times of high pollution (as opposed to simply a stronger willingness to reduce the monitors' pollution recordings), we estimate a model with the full set of interactions. Column 2 reports the results from estimating this model and shows suggestive evidence (significant at the 10 percent level) that, in the presence of monitoring, there is a general increase in enforcement during periods of high pollution also in areas farther away from the

[26] We prefer taking this reduced-form approach instead of instrumenting local pollution levels since we only have pollution data from the monitors from 2015 onward.

TABLE 2—FIRM LEVEL: ENFORCEMENT RESPONSE BY RAINFALL AND WIND DIRECTION

	Any air			Any air	
	(1)	(2)		(3)	(4)
Panel A. Rainfall			*Panel B. Wind Direction*		
$Rain_{>\tilde{x}}$	−0.00027	0.00031	*Upwind*	0.00033	0.000043
	(0.00033)	(0.00027)		(0.00023)	(0.00031)
$Mon_{<10km} \times Post$	0.0043	0.0041	$Mon_{<10km} \times Post$	0.0024	0.0027
	(0.00069)	(0.00074)		(0.00057)	(0.00067)
$Mon_{<10km} \times Rain_{>\tilde{x}}$		0.00014	$Mon_{<10km} \times Upwind$		−0.00035
		(0.00024)			(0.00045)
$Mon_{<10km} \times Post \times Rain_{>\tilde{x}}$	−0.0018	−0.0015	$Mon_{<10km} \times Post \times Upwind$	0.0035	0.0024
	(0.00068)	(0.00076)		(0.00082)	(0.0011)
$Post \times Rain_{>\tilde{x}}$		−0.0018	$Post \times Upwind$		0.0015
		(0.0011)			(0.00075)
Mean outcome	0.0046	0.0046	Mean outcome	0.0046	0.0046
Observations	1,155,296	1,155,296	Observations	1,155,296	1,155,296

Notes: This table reports results from augmented versions of our baseline model (equation (2)). Panel A (columns 1–2) adds interactions with an indicator for rainfall being above the median ($Rain_{>\tilde{x}}$); i.e., the interaction captures the differential effect on enforcement when pollution is relatively low (see online Appendix Table C6). Panel B (columns 3–4) adds interactions with an indicator for a firm being upwind from a monitor (online Appendix Figure D4 illustrates this classification); i.e., the interaction captures the differential effect for firms whose emissions are moved by the wind toward the monitor. All regressions control for fixed effects specific to firm, industry-by-time, and province-by-time interactions. Robust standard errors clustered on the city in parentheses.

monitor. However, as indicated by the triple interaction, the effect is even stronger for firms close to a monitor. These results also provide additional support for the validity of this exercise by showing that there is no differential response to rainfall shocks in the pre-period in areas close to the monitors compared to areas farther away.

We then move on to explore how the direction of winds affects the enforcement response. Emissions from firms that are upwind from a monitor will be moved by the wind toward the monitor, while emissions from all other firms are moved away from the monitor. Since upwind firms arguably have a larger impact on the pollution recorded by the monitor, local government officials have a greater incentive to enforce regulations for these firms. To test this we follow previous work (Freeman et al. 2019) and identify a firm as "upwind" if it is within 45 degrees of the dominant quarterly wind vector that passes through the monitor. Figure D4 in the online Appendix illustrates how we classify upwind firms. Following the same approach as above, we interact whether a firm is upwind in a quarter with our main treatment variables. The results are reported in panel B of Table 2. Column 3 shows that the coefficient for the upwind indicator is small and not significantly different from zero—suggesting that firms do not face differential enforcement by quarterly winds before the introduction of monitors. However, following the introduction of monitors, the enforcement response is stronger against upwind firms (as shown by the coefficient for $Mon_{<10km} \times Post \times Upwind$). Upwind firms face an increase in enforcement when monitored that is more than twice as large $(0.0024 + 0.0035 = 0.0059)$ as that faced by other firms (0.0024). Column 4 further documents that while this step-up in enforcement is particularly pronounced for firms close to the monitor, it is also detectable for firms farther than 10 km from a monitor (as shown by the estimate for $Post \times Upwind$). This suggests that more distant upwind firms also affect the monitor's recording of pollution and therefore face increased enforcement.

Firm-Level Robustness: Time-Specific Shocks and Spillovers.—Our interpretation of the above results is that monitors lead to an increase in enforcement of environmental regulations. In this section, we will discuss and investigate two threats to this interpretation of our results.

The first concern we will address is the risk that our results are affected by time-specific policies or shocks that coincide with the introduction of monitors. This could be the case if firms close to a monitor respond differently to such a shock compared to firms farther away. For this to bias our results, firms within the same industry/province would need to be differently affected depending on their distance to a monitor (since our baseline specification controls for both four-digit industry-by-time and province-by-time fixed effects). While we are not aware of any such policies or shocks, we investigate this potential concern further by conducting three additional tests. First, we use equation (2) to look at environmental enforcement that is not related to air pollution. The results are reported in panel A of Table C7 in the online Appendix. For enforcement related to water pollution, solid waste pollution, and procedure violation, estimates are small and statistically insignificant. Second, panel B of online Appendix Table C7 gradually includes controls that interact baseline characteristics with time fixed effects to our main specification. First we address the fact that our sample period overlaps with the relocation of basic manufacturing away from coastal regions stipulated in the Twelfth Five-Year Plan. Column 2 addresses this issue by including the distance from the city to the coast interacted with time fixed effects. Then we address the fact that policies or shock exposure might be related to firm characteristics in column 3 and interact the number of employees and ownership status with time fixed effects. Finally, we control for city-by-time fixed effects in column 4. Estimates remain of a comparable magnitude and are highly statistically significant across all specifications. Third, we conduct placebo tests in which we estimate the specification used to produce Figure 2, but instead of the distance to the closest monitor, we use the distance to the local environmental bureau or the distance to the city's firm centroid. We do this to investigate whether there is a differential response in enforcement after 2015 depending on plausible spatial determinants related to both the costs and returns from enforcement.[27] Figure D5 of online Appendix D shows that there are no detectable changes in the gradients of enforcement activity pre- and post-2015. To further validate our main results, we include the distances to both the environmental bureau and the firm centroid in our main specification and interact the distance bins with the time fixed effects. Due to the high correlation between monitor location and these measures, our results are slightly less precise, but largely unaffected when including this full

[27] We argue that the costs of enforcement are likely to be lower close to the environmental bureau (where monitors are often placed) and that the returns from enforcement are likely greater in central areas (due to a greater concentration of population and economic activity). To calculate firm centroids for each city, we use the geographical distribution of all ASIF firms. The firm centroid is a single point representing the barycenter of all firms.

set of controls.[28] Taken together, these results suggest that the step-up in enforcement behavior that we observe is indeed driven by the monitors.

The second concern we will investigate is whether the results we observe above are affected by spillovers. This could be the case if our control group (firms farther away from the monitor) is also affected by the introduction of monitors, e.g., if firms observe and respond to enforcement actions against neighbors, or if the local government reduces enforcement efforts against firms far from a monitor. Depending on the nature of the spillovers, these could lead us to either over- or underestimate the impact of monitoring. Our first attempt to address this issue is to reexamine the patterns in the raw data presented in Figure 1, panels A and B. The figures show that before 2015, enforcement levels and trends are the same in both treated and control firms. After 2015, there is a clear step-up of enforcement against firms within 10 km of a monitor, while firms farther away continue on a similar upward-sloping enforcement trend. If these firms would have been affected by spillovers, we would have expected a change in their enforcement trend post-2015. While these patterns in the raw data are reassuring, they do not necessarily rule out spillovers, since we do not know the counterfactual trends. Therefore, we move on to conduct an analysis at the city level in the following section where we analyze aggregate outcomes that take any within city spillovers into account. In addition, we specifically investigate the impact of the intensity of monitoring on nonmonitored areas following an approach similar in spirit to the work by Crépon et al. (2013).

C. *City-Level Evidence*

To study the impact of monitoring at the city level, we exploit variation in the intensity of the monitoring program. This approach allows us to infer the overall impact of more extensive monitoring on both enforcement and pollution. To conduct this analysis, we exploit the criteria set up by the central government when implementing the monitoring program (listed in Table C2 in online Appendix C) and compare outcomes in cities that installed different numbers of monitors. The argument behind this approach is that a larger number of monitors will ensure that officials are held accountable for a greater share of the overall pollution in a city. Figure 3 illustrates this point for our data by showing how the number of monitors assigned to a city is related to the share of high-pollution activity that is covered by a monitor. We again rely on the ESR database to identify high-polluting firms. To adjust for firm size, our baseline measure of high-pollution activity covered calculates the share of a city's high-polluting firms' revenue that fall within 10 km of a monitor.[29] The figure shows that the share of high-pollution activity covered within 10 km of a monitor increases monotonically with the number of monitors, from 20 percent for cities with 1 monitor to more than 80 percent for cities with 6

[28] The main estimate in column 1 of panel A in Table 1 changes from 0.0033 to 0.0017 and is significant at the 5 percent level.

[29] The ESR database allows us to identify high/low-polluting firms (i.e., whether the firm is on the list or not). We use the revenue share to scale our measure by firm size and thereby get closer to the share of overall pollution that is covered by the monitors. Our results are robust to using alternative measures of firm size as reported in Table C8 in the online Appendix.

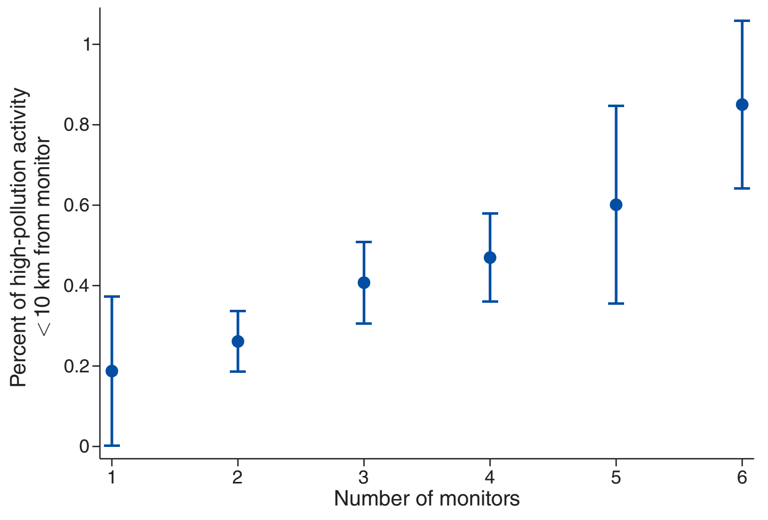

Figure 3. Number of Monitors and Coverage of High-Pollution Activity

Notes: This figure documents the relationship between the number of monitors in a city and the share of the city's high-pollution activity that is covered by monitors. This measure is constructed by calculating the share of the city's ESR firms' revenue that falls within 10 km of a monitor. Marginal effects and alternative measures are reported in online Appendix Table C8.

monitors. We report marginal effects in online Appendix Table C8. Our baseline estimate suggests that an additional monitor increases coverage by 11 percentage point (a 30 percent increase compared to the mean). These results are robust to using alternative distances to the monitor and other indicators for firm size as well as exploiting variation in both the actual and assigned number of monitors (reported in online Appendix Table C8).

City Level: Event Study.—To study the effects of monitoring at the city level, we first estimate a standard event study specification. If we denote a generic city by c, with $c \in r$, where r denotes a pollution reduction target group in online Appendix Table C3, and t is a generic time period (month for the pollution analysis and quarter for the enforcement analysis), our model can be written as

$$y_{crt} = \delta_c + \gamma_{rt} + \sum_{\substack{k=\mathrm{Q1}/2010 \\ k \neq \mathrm{Q4}/2014}}^{\mathrm{Q4}/2017} \beta_k m_c^k + \lambda \mathbf{X}_{ct} + \epsilon_{crt}, \tag{3}$$

where y_{crt} is either an aggregate measure of a city's monthly AOD or the log of the total number of firms that receive any enforcement related to air pollution in a quarter, m_c^k is either the actual number of monitors in the city or the predicted number of monitors according to online Appendix Table C2 in a given quarter k, δ_c are city fixed effects, and γ_{rt} are pollution target group by time fixed effect (month–year for the pollution specification and quarter–year for the enforcement specification).[30] The variable $\mathbf{X}_{ct}$ represents time-varying city controls including total precipitation, average temperature, the age of the mayor in office, and the geographical and population size of the city at baseline interacted with the post variable. The error term is denoted by ϵ_{crt}, which we cluster at the city level to account for potential serial correlation of the errors over time. Because we condition on city as well as on pollution-target-by-time fixed effects, parameter estimates capture the average effect of monitoring on the differential change in pollution/enforcement across cities with the same pollution reduction target.

To causally identify the impact of an additional monitor, we rely on common trends across cities with different numbers of monitors. To assess the validity of this assumption, we start by investigating the AOD trend for each group. In Figure 4, panel A, we plot de-meaned city-level AOD trends in four groups, which are determined according to the minimum number of monitors assigned by the central government. Two important patterns can be noted. First, there is a relatively flat AOD trend in cities assigned one monitor, suggesting that there was no major change in pollution in these cities and that it is therefore a suitable control group. Second, raw AOD data in all four groups share a common trend before 2015, after which AOD diverges—with a more substantial reduction for cities assigned a larger number of monitors.

To formally test this, we estimate equation (3), setting the average pollution in the quarter before monitors were installed as the baseline. Estimates from this specification are shown in Figure 4, panel B. We first estimate a standard event study specification using the actual number of monitors installed in a city as our independent variable of interest. These estimates are reported by the black dashed line in the figure. Point estimates are imprecise in the early time period, but results corroborate the findings above that there are no differential trends in AOD leading up to the intervention. We also see a substantial drop in pollution in the post-period for cities that installed additional monitors—effects that are even stronger in the second and third year.[31] Results on enforcement follow the corresponding pattern: there is no evidence of differential pre-trends and a large increase in enforcement in cities with additional monitoring after 2015 (reported in online Appendix Figure D6).

One potential concern with the above specification is that it might lead to biased results if cities were able to influence the number of monitors installed. The estimates would be biased if, for example, cities that expected lower pollution in the future installed a larger number of monitors. To address this concern, we use the minimal number of monitors set by the MEP as an instrument for the actual number of monitors (m_c). The instrumental variable estimates are marked by the dashed gray line in Figure 4, panel B. The IV coefficients are less precise but follow the OLS estimates closely in the pre-period and are even larger in the post-period. There is no evidence of differential trends leading up to the intervention, supporting the

[30] We add $+1$ to the number of firms with any enforcement related to air pollution before taking the log to avoid generating missing observations. Our results are robust to using the inverse hyperbolic sine transformation instead of the log.

[31] Note that monitors are operational from January 1, so all periods in the quarter of adoption are treated. We report estimates by quarter rather than by month to facilitate comparison with the enforcement results.

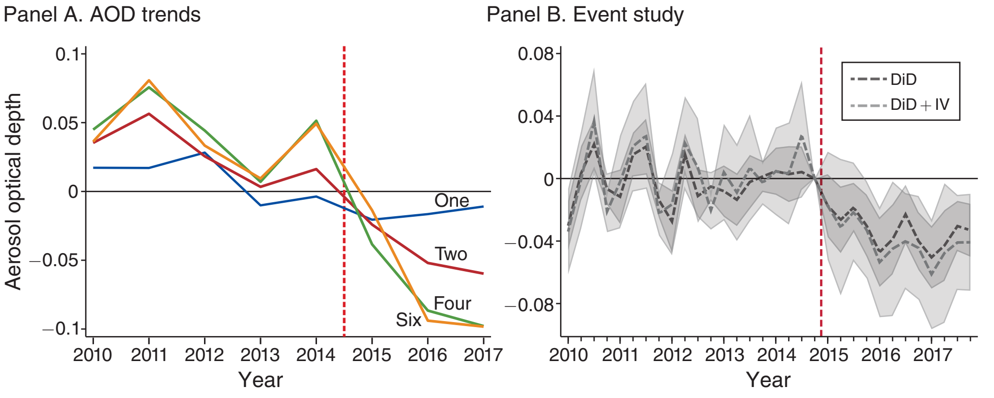

Figure 4. City Level: Impact of Number of Monitors on AOD

Notes: Panel A presents de-meaned city-level AOD trends in four groups. Groups are determined according to the minimum number of monitors assigned according to the regulation. The red line marks the introduction of air quality monitors. Panel B presents the estimates from equation (3) using variation in either the actual (DiD) or the assigned number of monitors (DiD + IV). Dashed black lines represent the coefficients from DiD, whereas gray dashed lines represent DiD + IV estimates. Shaded areas represent 95 percent confidence intervals. AOD is formally defined as the negative of the natural logarithm of the fraction of light that is not scattered or absorbed. Hence, these estimates can be interpreted as percentage changes in pollution.

common trend assumption between cities assigned different numbers of monitors. Again these results are mirrored by estimates for enforcement (reported in online Appendix Figure D6).

City Level: Regression Discontinuity Plots.—While the analysis above suggests that we capture the causal effect of the number of monitors on pollution and enforcement, a potential remaining concern is that the above approach is not adequately capturing potential confounding effects of city size (i.e., our controls for population and the size of the built-up area are not sufficient). This could be an issue if the incentives to reduce pollution change differently across cities of varying sizes after 2015. We do not have any reason to suspect that this is the case. However, to formally address this potential concern, we conduct an additional analysis in which we explore discontinuities in the number of monitoring stations assigned by the central government. Cities are assigned to one of four groups (corresponding to the installation of one, two, four, or six monitors) based on either the city's population or the geographical size of the built-up area in 2014 (as shown in Table C2 in online Appendix C), whichever assigns it to the highest group.[32]

We exploit this variation to implement a fuzzy RD design. Compared to the standard RD design that uses one running variable and cutoff, this setting provides us with two potential running variables (population and geographical size) and three potential cutoffs per running variable (1–2, 2–4, 4–6). However, in practice, we document that there are no discontinuous changes in the number of monitors at the

[32] For example, a city with a population of 200,000 individuals and a built-up area of 21 square kilometers is assigned two monitors.

population thresholds (estimate 0.38 and p-value 0.23). We therefore only exploit variation in the size of the built-up area. To reduce the impact of outliers, we focus on the first two cutoffs since the final group only contains eight cities (five within the optimal bandwidth).[33] We pool all observations, and use the distance to the closest geographical threshold as the running variable. Following the suggestion by Cattaneo, Keele, and Titiunik (2021), we control for the baseline value of the outcome as well as threshold fixed effects to improve precision.[34]

We start with a visual inspection of the data following the approach suggested by Calonico, Cattaneo, and Titiunik (2014). First, we document the difference in the number of monitors installed for cities on opposing sides of the assignment cutoffs. Figure 5, panel A illustrates the results by showing a binned scatterplot of the number of monitors in each city on the geographical size of the city's built-up area, with negative values for cities below the cutoffs and positive values for cities above the respective cutoffs. Cutoff fixed effects, as well as the baseline value of the outcome variable for time-varying data, are absorbed before plotting the data, and the graph also reports a fitted second-degree polynomial. The number of monitors exhibits a sharp jump when moving from the left to the right of the threshold. The first-stage estimates show that cities just above the threshold have installed approximately 1.3 additional monitors. Figure 5, panels B and C use the same approach as above and show the reduced-form estimates on average monthly AOD and the log of the quarterly number of firms that received any enforcement related to air pollution in the post-period. We see clear jumps in both variables when moving from the left to the right of the threshold.

The above estimation results rest on the standard assumption that there is no manipulation of the running variable and that other characteristics of cities are smooth at the thresholds. If mayors were able to manipulate the size of the built-up area and sort below the threshold to avoid an additional monitor, our estimates could suffer from selection bias. Figure D8 in online Appendix D is reassuring about the absence of manipulation, as there is no jump in the distribution at any threshold. To test whether municipalities could have manipulated the running variable, we take advantage of the McCrary (2008) observation that in the absence of manipulation, the density of the running variable should be continuous around the threshold. To formally test whether the density of the running variable is continuous at the threshold, we use the local polynomial density estimator and test statistic as described in Cattaneo, Jansson, and Ma (2018). Online Appendix Figure D8c plots the estimated empirical density. The graphical representation clearly suggests that the running variable is continuous at the threshold. The p-value for the null hypothesis that the density of the running variable is continuous at the threshold is 0.642.

To test the second assumption, we study the main threat to this identification strategy, i.e., that cities with a different number of monitors face different pollution reduction targets. We look at targets for cities close to the thresholds using the

[33] All results are robust to include the third threshold in the analysis as well.

[34] The reasoning for introducing these controls follows the same logic as in experiments where baseline outcomes are commonly used to improve precision. We control for the AOD or the log of the number of firms facing any enforcement action in 2010, respectively. We show in online Appendix Table C12 that estimates without controls are very similar, but less precise.

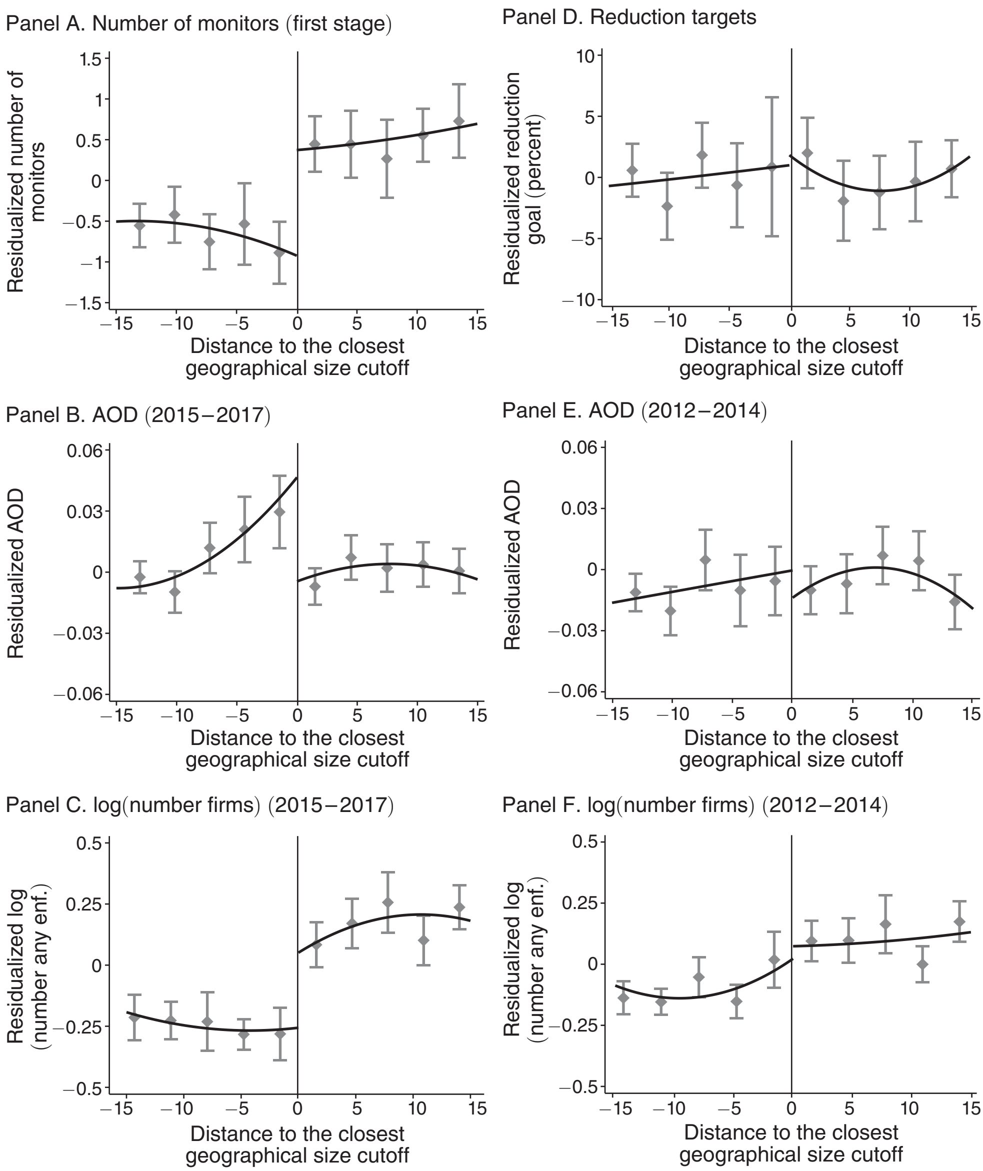

Figure 5. City Level: Regression Discontinuity Plots

Notes: This figure reports RD plots using the procedure suggested by Calonico, Cattaneo, and Titiunik (2014). The running variable is the distance to the closest geographical size threshold (as defined in online Appendix Table C2). Panel A shows how the number of monitors differ for cities around the geographical size thresholds, while panel D shows the variation in pollution reduction targets. The difference in AOD and log of the number of firms facing any enforcement after the monitors are installed are shown in panels B and C, while the corresponding outcomes before the monitors are installed are shown in panels E and F. To reduce sampling variability, we follow existing literature (Lee and Lemieux 2010; Cattaneo, Keele, and Titiunik 2021) and residualize our dependent variables before plotting the RD graphs. We use cutoff fixed effects and the baseline (2010–2011) value of the outcome variables (AOD and enforcement, respectively) for the time-varying data to perform this residualization. Robust standard errors are clustered at the city level. Error spikes represent 95 percent confidence intervals.

same cross-sectional specification we used above to estimate the first-stage impact on the number of monitors. Figure 5, panel D reports the results from this exercise and shows that pollution reduction targets are smooth around the thresholds. This suggests that differential pollution reduction targets do not drive our results. As additional checks, we present RD plots (Figure 5, panels E and F) for our main outcomes during the prepolicy period (2012–2014). Contrary to the postpolicy periods (2015–2017), we see no jumps at the threshold—suggesting that other relevant characteristics are smooth at the threshold.

City Level: Main Results and Robustness.—We now turn to quantify the magnitude of the effect from the different strategies described above. First, we estimate aggregate DiD and IV effects using the following simplified version of equation (3):

$$y_{crt} = \delta_c + \gamma_{rt} + \beta m_{ct} + \lambda \mathbf{X}_{ct} + \epsilon_{crt}, \tag{4}$$

where m_{ct} is the number of monitors installed in city c at time t and all other variables are the same as in equation (3). The first two columns of Table 3 report estimates of β using the DiD and IV strategy, respectively. Panel A shows the effect of monitoring on air pollution measured by aerosol optical depth, and panel B shows the results on the log of the number of firms that faced any air-pollution-related enforcement activity. The estimates show that one additional monitor leads to a 15–19 percent increase in enforcement and a 3.1–4.6 percent decrease in AOD. To put these estimates in perspective, we can calculate the enforcement response elasticity with respect to increased coverage of high-pollution activity using our estimates from Figure 3. Since one additional monitor is associated with a 30 percent increase in coverage of high-pollution activity, our results imply an enforcement response elasticity between 0.5 and 0.7. However, we express caution when interpreting these estimates since our measure of pollution coverage is imprecise and does not necessarily capture all important pollution sources within a city.

Tables C9, C10, and C11 in the online Appendix explore the robustness of these results. We start by interacting the time fixed effects with additional city characteristics in online Appendix Table C9. First, we show that estimating a slightly more demanding specification where we interact baseline city population and the geographical size of the built-up area with time fixed effects instead of the post variable does not alter our main estimates. We then interact the time fixed effects with additional city characteristics: GDP in 2010 and whether a city is assigned a background monitor. Due to the large number of fixed effects, estimates on enforcement are somewhat smaller and less precise, but overall results are of a comparable to our main results. Online Appendix Table C10 reports estimates when we drop data from the provinces Xinjiang and Tibet because the areas covered by cities in these two provinces are much larger than for the rest of the country. The estimates for both pollution and enforcement using the restricted sample are again of a comparable magnitude to our baseline estimates. As a final robustness check, we further investigate the impact of monitoring on the total number of firms that face any enforcement in a city in a quarter (this includes firms that are not covered by the ASIF survey).

Table 3—City Level: Impact of Monitoring

Empirical strategy:	DiD (1)	DiD + IV (2)	RD (3)	Diff-in-disc (4)
Panel A. Aerosol optical depth				
# Monitors	−0.031	−0.046	−0.039	−0.029
	(0.0069)	(0.013)	(0.015)	(0.018)
Observations	16,335	16,335	3,209	8,508
Panel B. log(*# firms receiving any air pollution enforcement*)				
# Monitors	0.15	0.19	0.26	0.28
	(0.046)	(0.098)	(0.10)	(0.16)
Observations	5,664	5,664	1,116	2,976
Panel C. Number of monitors (*first stage*)				
Estimate		0.72	1.28	1.28
		(0.11)	(0.23)	(0.23)
Kernel			Uniform	Uniform
Bandwidth			11.3	11.3

Notes: This table reports the main results for the two main outcomes for each of the four different empirical strategies used in the city-level analysis. Panel A reports results for aerosol optical depth, and panel B for the log of the number of firms receiving any enforcement action related to air pollution. Columns 1 and 2 show the estimates from equation (4), controlling for city fixed effects, time-by-pollution-reduction-target fixed effect, population and the geographical size of the built-up area at baseline interacted with the post variable, and time-varying controls for total precipitation, average temperature, and the age of the mayor. Column 1 exploits variation in the actual number of monitors installed, while column 2 instruments the actual number of monitors with the assigned number. The corresponding first-stage estimate is reported in panel C. Column 3 reports regression discontinuity estimates using local linear regression, a uniform kernel, and the MSE-optimal bandwidth proposed by Calonico, Cattaneo, and Titiunik (2014). We control for cutoff fixed effects and the average of the outcome in 2010–2011, details in online Appendix B.1. The first-stage effect of being assigned to the group above the cutoff on the number of monitors installed is reported in panel C. Column 4 reports estimates from the difference-in-discontinuity approach suggested by Grembi, Nannicini, and Troiano (2016), details in online Appendix B.2. Robust standard errors clustered on the city in parentheses.

Panel A in online Appendix Table C11 reports the results and shows that the overall effects are close to the results for our baseline sample.[35]

Second, to quantify the RD results above, we use the bias-corrected local linear regressions approach suggested by Calonico, Cattaneo, and Titiunik (2014), using a uniform kernel and controlling for cutoff fixed effects and the 2010–2011 baseline value of our outcome variables.[36] We report the formal specification in online Appendix B.1. Column 3 in Table 3 reports our baseline estimate, which uses the optimal bandwidth suggested by the same authors. Results are comparable to our DiD and IV estimates discussed above and suggest a 26 percent increase in enforcement and a 3.9 percent reduction in AOD. We again investigate the robustness of these estimates in the online Appendix. We start with establishing in online Appendix Table C10 that these estimates are also robust to excluding the large provinces Xinjiang and Tibet. Online Appendix Figure D7 shows that estimates are

[35] We are able to include the non-ASIF firms in this analysis since we can match them to cities even if we don't know the exact geographic location within the city.

[36] We explore the implications of these choices in Table C12 in the online Appendix. This shows that the choice of kernel has limited impact on the results, while the inclusion of baseline controls is important for precision but less so for the size of the estimates.

consistent across a range of alternative bandwidths, and online Appendix Table C13 reports that estimates are similar across the two cutoffs.

Finally, we combine the two approaches above and estimate a difference-in-discontinuities regression following Grembi, Nannicini, and Troiano (2016). This approach combines the standard nonparametric RD model with every term being interacted with dummy variables indicating the post-period. We again rely on the optimal bandwidth approach suggested by Calonico, Cattaneo, and Titiunik (2014). We report the formal specification in online Appendix B.2. Estimates from this specification are reported in column 4 of Table 3. Results are comparable to previous estimates but less precisely estimated. Again, we document the robustness of these results with respect to the sample studied and the firms considered and separate by cutoff in online Appendix Tables C10, C11, C13.

City Level: Spillovers.—We now turn to separate the aggregate city-level effect of monitoring into direct and spillover effects. To perform this analysis we exploit the same four identification strategies as above but calculate the outcomes separately for the monitoring station (within 10 km), the city center (10–50 km from a monitor), and the surrounding areas (beyond 50 km). This allows us to observe how enforcement behavior and pollution change in monitored as well as in nonmonitored areas. Observing this is important since officials could potentially reallocate enforcement effort from nonmonitored to monitored areas in response to the policy. If such behavior took place, we would expect to see a reduction in enforcement activity in the nonmonitored areas of cities that installed a greater number of monitors compared to cities that installed fewer monitors. Our approach is similar in spirit to the work by Crépon et al. (2013), which identifies spillovers by investigating the response of nontreated firms in regions with different treatment saturation.

Online Appendix Table C14 reports these results and shows that while we see that a greater number of monitors led to a substantial increase in enforcement close to the monitors, estimates for the city center and surrounding areas are still positive, though smaller and not significantly different from zero for most empirical strategies. This suggests that the increase in enforcement close to the monitors that we observe in the firm-level analysis is not driven by a reallocation of enforcement effort from nonmonitored areas. Estimates for pollution show reductions across the city. While DiD and IV estimates are comparable across areas, discontinuity estimates suggest smaller pollution reductions in surrounding areas (with point estimates ranging from from 49 percent to 78 percent of the size of the estimates for the area close to the monitor). These estimates suggest that there are positive spillovers (reduction in pollution) in areas not covered by a monitor. There are a number of potential explanations consistent with these results. One possibility is that there is a reduction in the spread of pollution from the high-pollution areas close to the monitors, where we document an increase in enforcement, to the nonmonitored areas of the city. An alternative explanation is that the step-up in enforcement in the city center has deterrence effects affecting the pollution behavior of noncovered firms as well. However, we are careful in not interpreting these estimates too strongly since satellite-based measures of pollution are highly spatially correlated. Hence, these effects could partly reflect the inability of satellites to pick up differences in pollution at a high spatial resolution.

To sum up, the results in this section show that an increase in monitoring intensity (i.e., a greater coverage of high polluters in a city) leads to enforcement against a larger number of firms (in particular those close to a monitor) and an overall reduction in AOD in the entire city. Point estimates are consistent across four different empirical strategies but vary in how precise they are. The RD approach has the key advantage of requiring weaker assumptions for causal inference, but the power of this analysis is lower, and it rests on a limited sample close to the threshold. We therefore focus on the DiD specification, which produces the most conservative and precise estimates, in the following mechanism section.

IV. Mechanisms

In this section, we investigate the potential channels through which the information captured by the monitors strengthens enforcement and reduces pollution. In Section IVA, we explore whether monitors improve top-down and/or bottom-up accountability. Thereafter, we document in Section IVB how a change in the information provision process that separates the responsibility to provide information from the responsibility to enforce regulations affects our results.

A. *Top-Down Accountability: Performance Incentives*

As discussed in Section IA, pollution reduction is one of the criteria that local leaders are evaluated on, and their performance determines their probability of promotion. Hence, a natural interpretation of our main findings is that monitors improve the central government's ability to evaluate how well local officials perform. In this section, we investigate this proposed mechanism more directly by exploiting heterogeneity in the promotion incentives faced by local officials. To get a plausibly exogenous measure of local promotion incentives, we use two unique features of the Chinese political system. First, we use the timing of the National People's Congress (NPC), which is held every five years and determines when political promotions are made in China. As documented in Xi, Yao, and Zhang (2018), the average probability of promotion for a city official in the last year of a political cycle (when the NPC is held) is nearly three times that of the first year in a cycle. We then combine this information with two official requirements for mayors of prefecture-level cities: that they retire at age 60 and serve for at least three years in a post. This means that city officials above the age of 57 at the time of the NPC face a discontinuously lower probability of being promoted and, therefore, weaker performance incentives (as documented in Xi, Yao, and Zhang 2018).

To conduct this analysis, we collect data on the main mayor in office in the year when the monitors were introduced and calculate their age at the thirteenth NPC, which was held in March 2018. If the information provided by the monitors strengthens the ability of the central government to hold local officials accountable, we would expect smaller effects of monitoring for cities with mayors who were above 57 years of age at the time of the congress. Mayors who are not facing promotion incentives are arguably less likely to work to achieve stricter enforcement of regulations.

To test our hypothesis about promotion incentives formally, we add an interaction term between the number of monitors in a city and the age of the mayor at the time of the congress to equation (4). As mayors' work experience might confound our analysis, we use a similar idea to the RD design and plot the differential effects (i.e., the interaction terms) of an additional monitor on both pollution and enforcement by the age of the mayor at the time of the congress in Figure 6. We set the baseline to the mayor who is 58 years old at the time of the congress and estimate other coefficients relative to this. A distinctive feature of both graphs is that the effects are not distinguishable from the baseline if the mayor is older than 58. At age 57, we see a substantial jump of the estimates in both graphs. The fact that estimates jump at 57 and are then consistent for lower ages suggests that our results are indeed driven by performance incentives and not by work experience or other age-related characteristics (for which we would not expect a jump at age 57). To further ensure that this result is not driven by confounding factors, we perform a corresponding analysis of city baseline characteristics in Figure D11 in the online Appendix. This balance test shows that the number of monitors, the size of the city, and economic activity as proxied by lights at night, as well as baseline pollution and enforcement activity, are smooth around the 57 age threshold—suggesting that our results are driven by the difference in promotion incentives.

We report the regression results of a simplified version of the results presented in Figure 6 in Table C15 in online Appendix C, where we instead interact the number of monitors with whether a mayor is below the age cutoff at the time of the NPC. Panel A displays the results for air pollution, and panel B displays the results on enforcement. In the first column, we use the full sample from our main analysis in Section IIIC, and we then subsequently restrict the sample to mayors closer to the performance age cutoff (again following the RD logic). The coefficients for the number of monitors now correspond to the effect for the mayors with weaker performance incentives, while the interaction term shows the additional effect of monitors in the presence of promotion incentives. These terms are all statistically significant and suggest that the performance incentive increases the response by between 65 and 87 percent. A corresponding balance table is reported in online Appendix Table C16. This shows that while there are some imbalances when we compare all cities with mayors above the threshold with all cities with mayors below the threshold, these differences are substantially smaller and no longer statistically significant when we move close to the threshold. We conclude from this analysis that a preexisting incentive scheme similar to those that are typically proposed to address the principal–agent problem is key in order for monitoring to be effective.

Since the information from the monitors is also provided to the public online, an alternative explanation for our main results is that monitors strengthen bottom-up accountability. To investigate whether this is a plausible driver of our results, we estimate the impact of the number of monitors on online search behavior using equation (4). We report results in online Appendix Table C17, which shows that while the point estimate on all five keywords is positive, the estimates are small, and

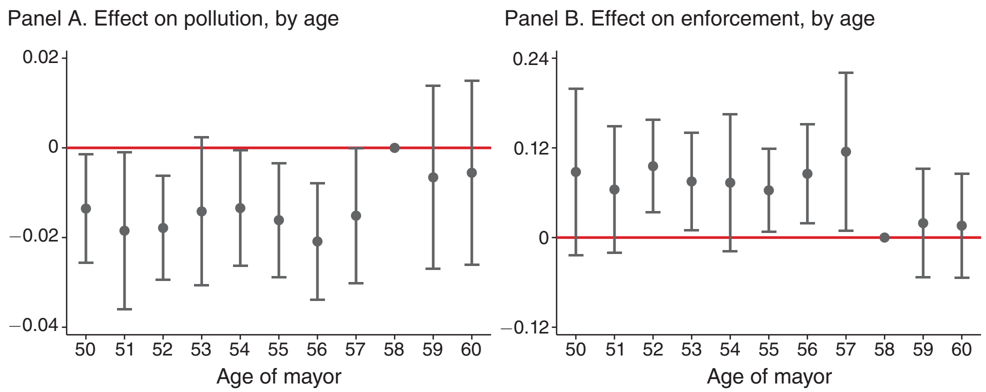

Figure 6. Main Results by Performance Incentives

Notes: This figure displays the effects of an additional monitor on both enforcement and pollution by mayors' age at the time of the NPC. Reported coefficients are relative to the effect for mayors who would be 58 years old at the time of the NPC (i.e., the baseline). Panel A reports estimates for pollution, and panel B reports estimates for enforcement. Error spikes represent 95 percent confidence intervals.

none of them are significantly different from zero. Hence, we do not find any evidence suggesting that the number of monitors improves information acquisition by the local population. This test compares cities with more or less comprehensive monitoring and not cities with and without monitoring. The result therefore does not rule out that citizens with access to monitors may exert more pressure on local leaders than citizens without access to any monitoring. However, the impact of more comprehensive monitoring on enforcement and pollution that we document does not seem to be driven by increased dissemination of information to the public.

B. *Changing Information Provision*

Although providing incentives for performance is a common approach to deal with the principal–agent problem, it has long been recognized that high-powered incentives can also distort the type of effort exerted or even encourage various harmful activities focused on improving indicators of performance (Figlio and Winicki 2005; Banerjee, Duflo, and Glennerster 2008; Fisman and Wang 2017; Acemoglu et al. 2020). Manipulating data on which performance is evaluated is one strategy that has been documented in a series of studies (Jacob and Levitt 2003; Figlio and Getzler 2006; Banerjee, Duflo, and Glennerster 2008; Sandefur and Glassman 2015; Greenstone et al. 2022). In this section, we study whether the structure of the information provision system could mitigate such concerns. In particular, our interest lies in understanding whether a separation of the agent responsible for providing information from the agent responsible for enforcing regulations affects the quality of information, and whether such quality improvements can, in turn, strengthen accountability and government performance (i.e., change behavior of the enforcing agent).

Several media sources have reported on manipulation of the pollution data from the monitors by local government officials.[37] Following this reporting, the central government decided to reassign the control of monitors to external parties, as documented above. In this section, we take advantage of this reassignment policy to see whether increasing the cost of manipulation for the local government is an effective way to improve monitoring, to reduce manipulation, and, through that, to enforce environmental policy.

As discussed in Section I, all monitors in our sample were reassigned to third parties at the same time in 2016. Hence, we are not able to exploit any cross-sectional variation to estimate the causal effect of the information provider. Instead, we focus on a descriptive analysis and discuss potential implications. First, we study how the AOD elasticity of $PM_{2.5}$ changes when the way information is provided changes (I_t).[38] More specifically, we estimate

$$\log\left(PM_{2.5}\right)_{mt} \;=\; \delta_m + \gamma_t + \beta_1 AOD_{mt} + \beta_2 AOD_{mt} \times I_t + \epsilon_{mt}, \tag{5}$$

where $\log\left(PM_{2.5}\right)_{mt}$ is the logarithm of monthly average concentrations of $PM_{2.5}$ reported from monitor m at time t, and δ_m and γ_t represent fixed effects for monitors and time. The variable AOD_{mt} captures the average monthly AOD for pixels covering monitor m.[39] I_t is a dummy variable indicating whether the data are reported after the reassignment. Therefore, the main coefficient of interest is β_2. If information is more accurate when monitors are controlled by the third party, we would expect that AOD and $PM_{2.5}$ measures are more aligned after the reassignment and therefore that $\beta_2 > 0$. Note that this coefficient captures how the alignment between AOD and $PM_{2.5}$ changes over time while still allowing for pollution levels to change over time.

The results from estimating equation (5) are reported in Table 4. As a point of reference, we start by estimating the elasticity for all monitors without any interaction term (this replicates the results in the first column of online Appendix Table C4). We then restrict our analysis to the incentivized monitors used in our study and find a positive estimate for the interaction term. This shows that the elasticity is 0.10 larger after the third party takes over the monitoring stations (corresponding to a 38 percent increase compared to the pre-period when local governments control the monitors). This evidence is consistent with less manipulation and higher-quality information during the period when the information provision responsibility is separated from the enforcement responsibility.

One alternative explanation for the above results is that the AOD data are better able to capture changes in pollution after the reassignment (this could, e.g., be due to changes in the composition of pollution over time or changes to the satellite

[37] See https://p.dw.com/p/32jqR and http://www.xinhuanet.com/politics/2018-08/09/c_1123244676.htm, for example.

[38] We focus on $PM_{2.5}$ because this is the pollutant most strongly correlated with AOD (see online Appendix Table C4). Martinez (2022) studies the manipulation of GDP data by autocratic leaders using a similar specification.

[39] To deal with the fact that data are sometimes missing for the pixel just above the monitor due to cloud coverage, we use the value from the closest neighboring cell as long as this is within 20 km of the monitor. All results are robust to using data at the city level instead.

Table 4—Monitor Reassignment, Data Quality, and Policy Impact

Sample:	All	Incentivized		Background			
Outcome:	$\log(PM_{2.5})$					*AOD*	$\log(\# \, firms)$
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
AOD	0.30	0.30	0.27	0.39	0.37		
	(0.031)	(0.033)	(0.038)	(0.050)	(0.057)		
AOD × *Reassigned*			0.10		0.050		
			(0.047)		(0.11)		
# Monitors						−0.025	0.12
						(0.0070)	(0.045)
# Monitors × *Reassigned*						−0.014	0.065
						(0.0030)	(0.028)
Mean outcome	3.68	3.71	3.71	3.42	3.42	0.34	0.58
Observations	17,535	15,496	15,496	2,039	2,039	16,322	5,646

Notes: This table reports the AOD elasticity of $PM_{2.5}$. Each column is from a separate regression. Columns 1–5 control for average temperature, rainfall, mayor's age, and fixed effects specific to monitor and time (month by year). Columns 6 and 7 control for city fixed effects, time-by-pollution-reduction-target fixed effect, population and the geographical size of the built-up area at baseline interacted with the post variable, and time-varying controls for total precipitation, average temperature, and the age of the mayor. Robust standard errors clustered on the city in parentheses.

instruments). To make sure that the changes we observe are due to improved monitor data rather than satellite data, we conduct a placebo analysis using the background monitors described in Section IB. The readings from these monitors are not used by the central government to evaluate the performance of the local government. Hence, there are weaker incentives for officials to manipulate this information. Columns 4 and 5 report the results. We notice that the overall elasticity between air pollution measures reported from monitors and satellites is larger for this sample. When looking at the reassignment, we find that the elasticity change is about half in magnitude and not statistically distinguishable from zero. Taken together, this evidence is consistent with less manipulation of the background monitors from the start and no change after the reassignment. This supports our conclusion above that the change in elasticity that we observe for the main sample is driven by changes in the data reported from the monitors. However, we are cautious against drawing too strong conclusions from these patterns since the estimates for the background monitors are imprecisely estimated and not statistically different from those for the main monitors.

The next exercise we carry out is to check whether local governments exert more effort to decrease pollution after monitors have been reassigned, since manipulation is then a less viable option. The results are reported in columns 6 and 7 of Table 4 and show that effects are indeed stronger after monitors have been reassigned. Column 6 shows a 1.4 percentage point greater reduction in pollution and column 7 a 6.5 percentage point larger increase in enforcement per monitor after the retraction. These pieces of evidence are consistent with local governments switching from data manipulation toward exerting more effort to enforce environmental regulations. Again, we emphasize that these results must be interpreted with caution because we are only exploiting temporal variation and thus need to assume that there are no other simultaneous changes causing these results. An alternative interpretation is that these results capture a lagged impact of the introduction of the monitors. However, we see no apparent reasons for why that would affect the relationship between satellite and ground-based measures of pollution discussed above.

V. Discussion and Concluding Remarks

This study uses the introduction of a pollution monitoring program in China to investigate the impact on local government enforcement of environmental regulations. Exploiting georeferenced firm data matched with enforcement records, we find that enforcement is stepped up against firms located within 10 km of a monitor. This resulted in an increase in enforcement in highly polluted areas in the city, which is where the monitors were placed. We further document that monitoring affects enforcement by altering which firms are targeted by local governments and by strengthening the responsiveness to local pollution shocks.

To study the aggregate impact of the policy, we conduct a city-level analysis and compare enforcement and pollution levels in cities that installed different numbers of monitors and thus introduced differential coverage of local pollution activity. Our baseline analysis shows that one additional monitor (a 30 percent increase in coverage of high-pollution activity) leads to about a 15 percent increase in the number of firms that face regulatory enforcement and a subsequent 3 percent reduction in city-level pollution. The increase in enforcement is focused on the areas close to the monitor, while pollution is reduced across the city. Given that the policy assigned a median of three monitors per city, this corresponds to a substantial reduction in overall pollution. Our estimates suggest a 0.42–0.86 $\mu g/m^3$ reduction in average $PM_{2.5}$ per additional monitor.[40] Previous literature suggests that such a decrease in pollution could have significant health and economic benefits, which would likely exceed the cost of the program in the short run.[41]

An examination of possible mechanisms suggests that the monitoring program is effective because it enables the central government to hold local government officials accountable for their actions. We support this claim by showing that monitoring is substantially more effective in localities where local officials face performance incentives. Finally, we document suggestive evidence showing that monitors deliver more reliable information when local governments are not involved in information reporting and are solely responsible for enforcement. When such an information

[40] We arrive at the estimate of 0.42 (0.86) $\mu g/m^3$, the lower (upper) bound, as follows. We multiply 3.1 percent (4.6 percent) from Table 3 by 0.30 (0.36, the elasticity with truthful reporting) from Table 4 to obtain percentage changes in $PM_{2.5}$ per monitor. We then multiply by 44.8 (52), the average $PM_{2.5}$ in our sample (average $PM_{2.5}$ in 2015, the first year for which we have $PM_{2.5}$ data), to estimate the implied change in $PM_{2.5}$.

[41] For example, Ebenstein et al. (2017) find that a 10 $\mu g/m^3$ increase in PM_{10} reduces life expectancy by 0.64 years in China. The medical costs of air pollution are also substantial—Barwick et al. (2020), e.g., document that a permanent decrease of 10 $\mu g/m^3$ in China leads to annual savings of more than $10 billion in health spending. Another cost of heavy air pollution in developing countries is the loss of productivity—Chang et al. (2016) and He, Liu, and Salvo (2019) find that a 10 $\mu g/m^3$ increase in $PM_{2.5}$ leads to about a 0.5 to 6 percent drop in productivity and labor cost saving. There are two main costs to consider for monitors: the cost of equipment and operation. According to the government procurement website, the cost of equipment/monitor is $200,000–$400,000, while yearly operation is $20,000.

provision structure is in place, the effect of an additional monitor on both enforcement of regulations and the level of pollution is significantly larger.

We believe our findings show not only that pollution monitoring could be an effective policy tool to combat ambient air pollution, but that it also offers some general lessons on how to approach the problem of lacking enforcement of government regulations caused by the principal–agent problem. Our findings suggest that reliable real-time monitoring of policy outcomes at the local level could contribute to closing the enforcement gap as long as local officials face performance incentives. However, the existence of such performance incentives could at the same time distort the behavior of local officials toward data manipulation. Therefore, the information provision system would need to be carefully designed to ensure accurate top-down accountability—e.g., by ensuring that information provision and enforcement responsibilities are sufficiently separated.

REFERENCES

Acemoglu, Daron, Leopoldo Fergusson, James Robinson, Dario Romero, and Juan F. Vargas. 2020. "The Perils of High-Powered Incentives: Evidence from Colombia's False Positives." *American Economic Journal: Economic Policy* 12 (3): 1–43.

Andrews, Steven Q. 2008. "Inconsistencies in Air Quality Metrics: 'Blue Sky' Days and PM10 Concentrations in Beijing." *Environmental Research Letters* 3 (3): 034009.

Auffhammer, Maximilian, Antonio M. Bento, and Scott E. Lowe. 2009. "Measuring the Effects of the Clean Air Act Amendments on Ambient PM10 Concentrations: The Critical Importance of a Spatially Disaggregated Analysis." *Journal of Environmental Economics and Management* 58 (1): 15–26.

Avis, Eric, Claudio Ferraz, and Frederico Finan. 2018. "Do Government Audits Reduce Corruption? Estimating the Impacts of Exposing Corrupt Politicians." *Journal of Political Economy* 126 (5): 1912–64.

Axbard, Sebastian, and Zichen Deng. 2024. "Replication Data for: Informed Enforcement: Lessons from Pollution Monitoring in China." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. https://doi.org/10.3886/E174901V1.

Baidu. 2017. "Search Index." https://index.baidu.com/v2/index.html#/.

Baker, George, Robert Gibbons, and Kevin J. Murphy. 1994. "Subjective Performance Measures in Optimal Incentive Contracts." *Quarterly Journal of Economics* 109 (4): 1125–56.

Banerjee, Abhijit V., Esther Duflo, and Rachel Glennerster. 2008. "Putting a Band-Aid on a Corpse: Incentives for Nurses in the Indian Public Health Care System." *Journal of the European Economic Association* 6 (2–3): 487–500.

Banerjee, Abhijit, Esther Duflo, Clément Imbert, Santhosh Mathew, and Rohini Pande. 2020. "E-governance, Accountability, and Leakage in Public Programs: Experimental Evidence from a Financial Management Reform in India." *American Economic Journal: Applied Economics* 12 (4): 39–72.

Barwick, Panle Jia, Shanjun Li, Deyu Rao, and Nahim Bin Zahur. 2018. "The Morbidity Cost of Air Pollution: Evidence from Consumer Spending in China." Unpublished.

Barwick, Panle Jia, Shanjun Li, Liguo Lin, and Eric Zou. 2020. "From Fog to Smog: The Value of Pollution Information." Unpublished.

Besley, Timothy, and Robin Burgess. 2002. "The Political Economy of Government Responsiveness: Theory and Evidence from India." *Quarterly Journal of Economics* 117 (4): 1415–51.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2012. "Creative Accounting or Creative Destruction? Firm-Level Productivity Growth in Chinese Manufacturing." *Journal of Development Economics* 97 (2): 339–51.

Brehm, Stefan, and Jesper Svensson. 2020. "Environmental Governance with Chinese Characteristics: Are Environmental Model Cities a Good Example for Other Municipalities?" *Asia-Pacific Journal of Regional Science* 4 (1): 111–34.

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik. 2014. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica* 82 (6): 2295–2326.

Cattaneo, Matias D., Michael Jansson, and Xinwei Ma. 2018. "Manipulation Testing Based on Density Discontinuity." *Stata Journal* 18 (1): 234–61.

Cattaneo, Matias D., Luke Keele, and Rocio Titiunik. 2021. "Covariate Adjustment in Regression Discontinuity Designs." In *Handbook of Matching and Weighting Adjustments for Causal Inference*, edited by José R. Zubizarreta, Elizabeth A. Stuart, Dylan S. Small, and Paul R. Rosenbaum, 153–68. Boca Raton, FL: CRC Press.

Chang, Tom, Joshua Graff Zivin, Tal Gross, and Matthew Neidell. 2016. "Particulate Pollution and the Productivity of Pear Packers." *American Economic Journal: Economic Policy* 8 (3): 141–69.

Chen, Yuyu, Ginger Zhe Jin, Naresh Kumar, and Guang Shi. 2013. "The Promise of Beijing: Evaluating the Impact of the 2008 Olympic Games on Air Quality." *Journal of Environmental Economics and Management* 66 (3): 424–43.

Chen, Yvonne Jie, Pei Li, and Yi Lu. 2018. "Career Concerns and Multitasking Local Bureaucrats: Evidence of a Target-based Performance Evaluation System in China." *Journal of Development Economics* 133: 84–101.

CMA. 2017. "Meteorological Data." China Meteorological Administration. https://data.cma.cn/index.html.

Conley, T. G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics* 92 (1): 1–45.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, and Philippe Zamora. 2013. "Do Labor Market Policies Have Displacement Effects? Evidence from a Clustered Randomized Experiment." *Quarterly Journal of Economics* 128 (2): 531–80.

Dhaliwal, Iqbal, and Rema Hanna. 2017. "The Devil is in the Details: The Successes and Limitations of Bureaucratic Reform in India." *Journal of Development Economics* 124: 1–21.

Duflo, Esther, Rema Hanna, and Stephen P. Ryan. 2012. "Incentives Work: Getting Teachers to Come to School." *American Economic Review* 102 (4): 1241–78.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. "Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India." *Quarterly Journal of Economics* 128 (4): 1499–1545.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2018. "The Value of Regulatory Discretion: Estimates from Environmental Inspections in India." *Econometrica* 86 (6): 2123–60.

Ebenstein, Avraham, Maoyong Fan, Michael Greenstone, Guojun He, and Maigeng Zhou. 2017. "New Evidence on the Impact of Sustained Exposure to Air Pollution on Life Expectancy from China's Huai River Policy." *PNAS* 114 (39): 10384–89.

ESRI. 2017. "China Prefecture City Boundaries." Environmental Systems Research Institute. https://www.arcgis.com/home/item.html?id=4220d0d2877c46789bfe6baf344cee1f.

Fetzer, Thiemo. 2020. "Can Workfare Programs Moderate Conflict? Evidence from India." *Journal of the European Economic Association* 18 (6): 3337–75.

Figlio, David N., and Joshua Winicki. 2005. "Food for Thought: The Effects of School Accountability Plans on School Nutrition." *Journal of Public Economics* 89 (2–3): 381–94.

Figlio, David N., and Lawrence S. Getzler. 2006. "Accountability, Ability, and Disability: Gaming the System." *Advances in Applied Microeconomics* 14: 35–49.

Fisman, Raymond, and Yongxiang Wang. 2017. "The Distortionary Effects of Incentives in Government: Evidence from China's 'Death Ceiling' Program." *American Economic Journal: Applied Economics* 9 (2): 202–18.

Freeman, Richard, Wenquan Liang, Ran Song, and Christopher Timmins. 2019. "Willingness to Pay for Clean Air in China." *Journal of Environmental Economics and Management* 94: 188–216.

Ghanem, Dalia, and Junjie Zhang. 2014. "'Effortless Perfection': Do Chinese Cities Manipulate Air Pollution Data?" *Journal of Environmental Economics and Management* 68 (2): 203–25.

Grainger, Corbett, and Andrew Schreiber. 2019. "Discrimination in Ambient Air Pollution Monitoring?" *AEA Papers and Proceedings* 109: 277–82.

Greenstone, Michael, and Rema Hanna. 2014. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." *American Economic Review* 104 (10): 3038–72.

Greenstone, Michael, and Patrick Schwarz. 2018. *Is China Winning its War on Pollution?* Chicago, IL: Energy Policy Institute at the University of Chicago.

Greenstone, Michael, Guojun He, Ruixue Jia, and Tong Liu. 2022. "Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution." *American Economic Review: Insights* 4 (1): 54–70.

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano. 2016. "Do Fiscal Rules Matter?" *American Economic Journal: Applied Economics* 8 (3): 1–30.

Gupta, Pawan, Sundar A. Christopher, Jun Wang, Robert Gehrig, Y. C. Lee, and Naresh Kumar. 2006. "Satellite Remote Sensing of Particulate Matter and Air Quality Assessment over Global Cities." *Atmospheric Environment* 40 (30): 5880–92.

He, Guojun, Shaoda Wang, and Bing Zhang. 2020. "Watering down Environmental Regulation in China." *Quarterly Journal of Economics* 135 (4): 2135–85.

He, Jiaxiu, Haoming Liu, and Alberto Salvo. 2019. "Severe Air Pollution and Labor Productivity: Evidence from Industrial Towns in China." *American Economic Journal: Applied Economics* 11 (1): 173–201.

Holmström, Bengt. 1979. "Moral Hazard and Observability." *Bell Journal of Economics* 10 (1): 74–91.

Holmström, Bengt, and Paul Milgrom. 1991. "Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design." *Journal of Law, Economics, and Organization* 7: 24–52.

Hsiang, Solomon M. 2010. "Temperatures and Cyclones Strongly Associated with Economic Production in the Caribbean and Central America." *PNAS* 107 (35): 15367–72.

Huang, Zhangkai, Lixing Li, Guangrong Ma, and Lixin Colin Xu. 2017. "Hayek, Local Information, and Commanding Heights: Decentralizing State-Owned Enterprises in China." *American Economic Review* 107 (8): 2455–78.

IPE. 2017. "Environmental Supervision Records." Institute of Public and Environmental Affairs. https://www.ipe.org.cn/index.html.

Jacob, Brian A., and Steven D. Levitt. 2003. "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating." *Quarterly Journal of Economics* 118 (3): 843–77.

Jia, Ruixue. 2017. "Pollution for Promotion." Unpublished.

Jiang, Junyan. 2017. "Chinese Political Elite Database." Junyan Jiang. https://www.junyanjiang.com/data.html.

Kahn, Matthew E., Pei Li, and Daxuan Zhao. 2015. "Water Pollution Progress at Borders: The Role of Changes in China's Political Promotion Incentives." *American Economic Journal: Economic Policy* 7 (4): 223–42.

Kosack, Stephen, and Archon Fung. 2014. "Does Transparency Improve Governance?" *Annual Review of Political Science* 17 (1): 65–87.

Lee, David S., and Thomas Lemieux. 2010. "Regression Discontinuity Designs in Economics." *Journal of Economic Literature* 48 (2): 281–355.

Martinez, Luis R. 2022. "How Much Should We Trust the Dictator's GDP Estimates?" *Journal of Political Economy* 130 (10): 2731-69.

McCrary, Justin. 2008. "Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test." *Journal of Econometrics* 142 (2): 698–714.

MEP. 2011. "Model Cities." MEP. https://www.mee.gov.cn/gkml/hbb/bgt/201101/W020110125328042389677.pdf (accessed January 6, 2022).

MEP. 2013. "Air Ten." MEP. http://www.gov.cn/zwgk/2013-09/12/content_2486773.htm (accessed January 6, 2022).

MEP. 2014. "Monitor Program - Wave 3." MEP. https://www.mee.gov.cn/gkml/hbb/bwj/201405/t20140509_273595.htm (accessed January 6, 2022).

MEP. 2017. "Mornitor Reading." MEP. http://106.37.208.233:20035 (accessed January 6, 2022).

Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2016. "Building State Capacity: Evidence from Biometric Smartcards in India." *American Economic Review* 106 (10): 2895–2929.

Nakagaki, Maiko, Anna Kompanek, and Marko Tominic. 2012. *Improving Public Governance: Closing the Implementation Gap Between Law and Practice*. Washington, DC: Center for International Private Enterprise.

National Bureau of Statistics. 2013. "Annual Survey on Industrial Firms (ASIF)." Ruisi Company. http://www.resset.cn/ied (accessed September 1, 2017).

Neidell, Matthew, and Janet Currie. 2005. "Air Pollution and Infant Health: What Can We Learn from California's Recent Experience?" *Quarterly Journal of Economics* 120 (3): 1003–30.

NEO. 2017. "Aerosol Optical Thickness." NASA Earth Observations. https://neo.gsfc.nasa.gov/view.php?datasetId=MYDAL2_M_AER_OD.

Oliva, Paulina. 2015. "Environmental Regulations and Corruption: Automobile Emissions in Mexico City." *Journal of Political Economy* 123 (3): 686–724.

Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200–49.

Reinikka, Ritva, and Jakob Svensson. 2005. "Fighting Corruption to Improve Schooling: Evidence from a Newspaper Campaign in Uganda." *Journal of the European Economic Association* 3 (2/3): 259–67.

Reinikka, Ritva, and Jakob Svensson. 2011. "The Power of Information in Public Services: Evidence from Education in Uganda." *Journal of Public Economics* 95 (7–8): 956–66.

Sandefur, Justin, and Amanda Glassman. 2015. "The Political Economy of Bad Data: Evidence from African Survey and Administrative Statistics." *Journal of Development Studies* 51 (2): 116–32.

Schlenker, Wolfram, and W. Reed Walker. 2016. "Airports, Air Pollution, and Contemporaneous Health." *Review of Economic Studies* 83 (2): 768–809.

Shimshack, Jay P. 2014. "The Economics of Environmental Monitoring and Enforcement." *Annual Review of Resource Economics* 6 (1): 339–60.

Snyder, James M., and David Strömberg. 2010. "Press Coverage and Political Accountability." *Journal of Political Economy* 118 (2): 355–408.

Song, Zheng, Kjetil Storesletten, and Fabrizio Zilibotti. 2011. "Growing Like China." *American Economic Review* 101 (1): 196–233.

Tanaka, Shinsuke. 2015. "Environmental Regulations on Air Pollution in China and Their Impact on Infant Mortality." *Journal of Health Economics* 42: 90–103.

United Nation. 2019. *Environmental Rule of Law: First Global Report*. New York, NY: United Nations.

Wang, Jun, and Sundar A. Christopher. 2003. "Intercomparison between Satellite-Derived Aerosol Optical Thickness and PM2.5 Mass: Implications for Air Quality Studies." *Geophysical Research Letters* 30 (21).

Wang, Shaoda, and David Y. Yang. 2022. "Policy Experimentation in China: The Political Economy of Policy Learning." Unpublished.

WHO. 2016. *Ambient Air Pollution: A Global Assessment of Exposure and Burden of Disease*. Geneva: World Health Organization.

World Bank. 2017. *World Development Report 2017: Governance and the Law*. Washington, DC: World Bank.

Xi, Tianyang, Yang Yao, and Muyang Zhang. 2018. "Capability and Opportunism: Evidence from City Officials in China." *Journal of Comparative Economics* 46 (4): 1046–61.

Zheng, Siqi, and Matthew E. Kahn. 2013. "Understanding China's Urban Pollution Dynamics." *Journal of Economic Literature* 51 (3): 731–72.

Zheng, Siqi, and Matthew E. Kahn. 2017. "A New Era of Pollution Progress in Urban China?" *Journal of Economic Perspectives* 31 (1): 71–92.

Zou, Eric Yongchen. 2021. "Unwatched Pollution: The Effect of Intermittent Monitoring on Air Quality." *American Economic Review* 111 (7): 2101–26.