THREE ESSAYS ON AIR POLLUTION REGULATION

By

Yuxian Xiao

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Economics—Doctor of Philosophy
Environmental Science and Policy—Dual Major

2023

ABSTRACT

Chapter 1: How Effective Are Low-Emission Vehicle Standards?

While mobile emission standards have been used for decades in several countries, their effectiveness is rarely discussed. Using the spatial and temporal variation caused by the implementation of California's Low Emission Vehicle (LEV) program, the paper applies the difference-in-differences method and the event study method for study. It compares the monitor level ozone concentration between states that implemented the LEV program with the states that did not. It is found that the LEV program significantly reduces the ozone concentration. By using the same empirical methods on county level mortality data, the paper also find a significant reduction in respiratory diseases. The results build the first step of evaluating the LEV program, and shed light on the less-studied field of mobile source emission standards.

Chapter 2: Monitoring and Inspections: the Correlation Between Monitor Breakings and Inspection Intensity

This study analyze the relationship between two critical regulation activities in air pollution control: monitoring and inspection. Utilizing a theoretical model, we reveal that the relationship between monitoring and inspection is complicated and not merely substitute. In particular, variations in monitoring intensity can change the marginal benefits of inspection differently across regions, depending on their cleanliness and the associated penalties for non-compliance. Empirical examination on the impact of incomplete monitors uncovers diverse responses from regulatory bodies at federal, state, and local levels. Specifically, while the EPA's actions remain unaffected, state agencies tend to curtail inspection irrespective of regional cleanliness in the face of monitoring reductions. Conversely, local regulators display a marked drop in inspection activities for dirtier regions under similar circumstances. Our findings underscore the importance of refining the monitoring system and allocating resources judiciously, especially in regions with higher pollution.

Chapter 3: Too Small to be Regulated? An Empirical Study on Colorado's Exemption Policy

Policy interventions for the mispriced externalities often have geographical and participant-specific limitations, potentially prompting regulated parties to sidestep compliance. This paper examines

Colorado's "Permit Section (PS) Memo-10-01"—an exemption introduced in 2011 that relaxed compliance requirements for smaller emitters of Nitrogen Oxides (NOx). Using the national emission inventory database, this study assesses whether this policy influenced smaller facilities to adjust their emission patterns. My findings suggest no significant evidence of emission manipulation at the specified 40 tons threshold or any strategic evasion by smaller emitters. Nonetheless, the nuanced complexity of such regulations underscores the necessity for future in-depth investigations.

ACKNOWLEDGEMENTS

My deepest gratitude to my parents. Their boundless love, sacrifices, and faith in me have been my bedrock. My sincere thanks go to my chair, Dr. Soren Anderson. His expertise, patience, and mentorship have been invaluable. I am also immensely thankful to my committee members – Dr. Justin Kirkpatrick, Dr. Prabhat Barnwal, and Dr. Adam Zwickle, for their critical insights, feedback, and encouragement. Also, many thanks to Dr. Jinhua Zhao for the insightful comments and discussions. I also wish to express my gratitude to my fellow graduate students, especially Kaidi Wu, Katie Bollman, and Suhwan Chung. Their support and shared experiences have made this journey both enlightening and enjoyable.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

NMOG Non-methane Organic Gases

LEV Low Emission Vehicle

VOCs Volatile Organic Compounds

NOx Nitrogen Oxides

CHAPTER 1

HOW EFFECTIVE ARE CALIFORNIA'S LOW-EMISSION VEHICLE STANDARDS?

1.1 Introduction

The economic theory of externalities teaches us that a pollution regulation policy should be designed so that the marginal abatement cost equals the marginal damage for all emitters. However, when it comes to ambient air pollution and non-point source emissions, this principle is difficult to follow. It is largely because these sources produce hard-to-measure emissions. It is infeasible to measure real-time vehicle emissions individually (Feng, Fullerton, & Gan, 2013), and modeling them is also challenging, because they are influenced by multiple factors. When one traditional fossil fuel vehicle is on the road, idling will create more pollution, as does traffic congestion. High temperatures and sunlight will cause more pollution. Fuel type also matters. In the absence of precise emissions data, either a tax or cap-and-trade will be ineffective in regulating mobile source pollution.

Mobile sources, however, are responsible for a significant amount of air pollution. As shown in Fig. 1.1, in the United States, the transportation sector contributes a substantial amount of Volatile Organic Compounds (VOCs) emissions and is the single biggest source of Nitrogen Oxides (NOx). These two chemicals are the precursors of ground-level ozone pollution, which is associated with an increased risk of cardiovascular and respiratory mortality (Zhang, Wei, & Fang, 2019). Currently, there are still more than 137 million people living with an unhealthy level of ozone.[1]

Mobile source emissions have an important health impact and are difficult to monitor, which forced policy-makers to rely on mandated policy tools. This includes driving or vehicle purchase restrictions, congestion taxes, mandated smog checks, fuel-content regulations, and this paper's focus, vehicle emissions standards. In 1966, California issued the world's first mandated vehicle emissions standard, largely in response to the notorious Los Angeles smog. Later, federal regulators of the United States borrowed California's standards to develop a nationwide emissions standard.

[1] Source: State of Air 2020 report by American Lung Association (ALA).

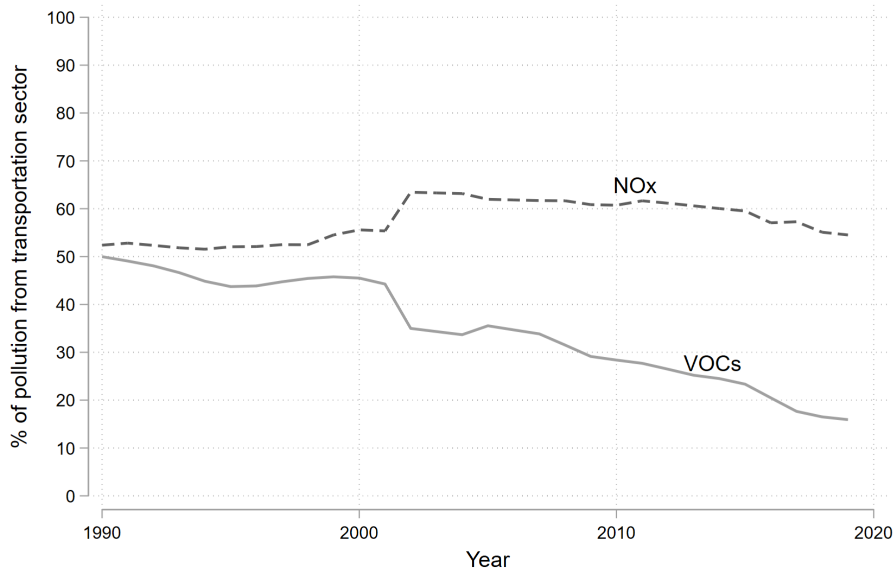


Figure 1.1 Transportation sector's share in U.S.'s emissions of NOx and VOCs

Source: National Emission Inventory (NEI) Air Pollutant Emissions Trends.

Following the United States, countries like Japan, Canada, Australia, and several European nations developed their own emission standards in the 1970s. Developing countries such as China and India also imposed emission standards following the same mode. However, while widely used for decades, little economic research directly evaluated these emissions standards. This disparity of policy practice and lack of attention in academia calls for studies that could provide an economic evaluation of emission standards. This paper is a response to the call.

In this paper, I focus on one important vehicle emission standard, California's Low Emission Vehicles (LEV) program (which also adopted by 12 other states and DC), and seek to answer the following questions: How do stricter state emissions standards affect air pollution? Does this regulation generate observable health benefits? To shed light on the pollution impact of LEV program, I supplement the 7.5 million monitor-day records of ozone pollution between 1990 to 2014 with the daily weather data, and employ the difference-in-differences and the event study methods to reveal the impact of the program. By comparing monitors that are located in the states that employed the LEV program and the states that have never been affected, the DD method gives an estimate of a 5.7% reduction in the ground-level ozone concentration, with more than 7% reduction in summer months and around 6.7% reduction in urban regions. Moreover, the event study shows that the parallel pre-trend holds. Further, the impact of the LEV program seems to follow the shape of the Non-methane organic gases (NMOG) requirement gap between the LEV and federal standards: the impact expands for about six years after the policy intervention and then quickly disappears. The ozone reduction impact is also found to be robust to various confounding factors, such as the change of the monitor system's coverage and the impact of cross-border commuting. Applying similar empirical methods to county-year mortality data, I find a 5.7% reduction in mortality from respiratory diseases and zero impact on mortality for irrelevant reasons, such as mortality from unintended injuries. However, when using the event study tool to estimate the impact on the health outcome, it is shown that the parallel trend assumption is violated and the impact of LEV seems to accumulate with time, which is different from the findings on ozone pollution.

This study contributes to the literature on air pollution regulation, especially the literature relating to mobile source emissions. The use of policies such as drive restrictions, congestion taxes, mandated smog checks, fuel economy standards, and fuel-content regulations has proven to be effective at reducing air pollution (Anderson & Sallee, 2016; Davis, 2008; Marcus, 2017; Oliva, 2015; Sanders & Sandler, 2020; Simeonova, Currie, Nilsson, & Walker, 2021; Viard & Fu, 2015; Xu, Zhang, & Zheng, 2015). Vehicle emission standards, on the other hand, are little studied. To the best of my knowledge, aside from Kahn (1996)'s early work, only Jacobsen, Sallee, Shapiro, and van Benthem (2022) directly addresses the emission standards. Using new vehicles' emission testing records from 1957 to 2019, Jacobsen et al. (2022) shows that the decline in vehicle emission rates over the past fifty years is largely due to the tightening of emission standards, and they calculate that the elasticity of vehicle emission rates with respect to emission standards is between 0.5 and 1.0. With detailed micro data on vehicle emissions, they also confirm that the vehicle emission rate increases rapidly with vehicle age, with vehicles more than ten years old producing the majority of the pollution. Their work provides the first comprehensive causal analysis of the emission standards' impact on real tailpipe emissions. However, it does not examine the impact on ambient air quality or health, which is the focus of this paper. Given the gap in literature, this paper could provide the first step in evaluating emission standards' impact on ambient air quality and health.

Furthermore, this study has significant policy relevance. Fifty years after the Clean Air Act (CAA), the United States has achieved great success in improving air quality (Currie & Walker, 2019), but ground-level ozone pollution is still a problem. In 2019, there are 284 counties around the United States that cannot meet the National Ambient Air Quality Standards (NAAQS),[2] and 220 among them have exceeded the ground-level ozone standards. In contrast, for the other five pollutants with NAAQS, there are no more than sixty nonattainment counties.[3] Given the important role that the transportation sector plays in generating ozone pollution, we need policies that effectively and efficiently control mobile source ozone-related emissions, and we need studies that focus on these policies.

The remainder of this paper proceeds as follows. Section 1.2 provides background on the LEV program. Section 1.3 introduces the data. Section 1.4 describes the empirical methods. Section 1.5 and 1.6 present the pollution results and mortality analysis results. Section 1.7 concludes the study.

1.2 Policy background

1.2.1 "No third car": either federal or California standard

Under the Clean Air Act (CAA), there are only two sets of vehicle emission standards in the United States: California and Federal standards.[4] There is a historical reason for such a situation. Originally, the congress of the United States preferred to have unitary federal vehicle emission standards, because the vehicle emissions would easily go across state borders and vehicle manufacturers prefer not to have multiple production standards. However, due to its historically

[2]NAAQS set standards for six criteria air pollutants: carbon monoxide (CO), lead, ground-level ozone, particulate matter (PM), nitrogen dioxide (NO_2), and sulfur dioxide (SO_2). The nonattainment data is collected from the EPA's Greenbook.

[3]In 2019, there are 13 lead nonattainment counties, 60 PM nonattainment counties, and 51 SO_2 nonattainment counties. All counties have achieved attainment with NO_2 and CO.

[4]Clean Air Act, Pub. L. No. 90-148, § 208(a), 81 Stat. 501, 501 (codified as amended at 42 USC § 7543(a)): "No State or any political subdivision thereof shall adopt or attempt to enforce any standard relating to the control of emissions from new motor vehicles or new motor vehicle engines subject to this part."

severe air pollution problem and early actions to battle these problems, California already had a complete emission regulation system before the federal government took any action. Thus, California became the only state that got the exemption to set its own standards from the beginning of mobile source pollution regulation in 1967 (Holmes, Allen, & Russel, 2007). The exemption allows California to set its vehicle emission standards based on the state's needs, as long as the standards are more strict or as strict as the federal regulation (Carlson, Hankins, & Stein, 2019). Later on, under section 177 of the CAA, which was added in the 1977 amendments of the CAA, other states are allowed to employ California's standards if such standards are found to be necessary for that state to comply with the National Ambient Air Quality Standards (NAAQS).

California Air Resources Board (CARB) exerted the discretion of making its own standards and introduced the Low-Emission-Vehicle Program in 1990, regulating vehicles manufactured in model year 1994 or later. The ultimate goal of the regulation was "to achieve the maximum degree of emission reduction possible from vehicular and other mobile sources".[5] Shortly after California's introduction of LEV, New York and Massachusetts exerted their authority under section 177 and adopted the California LEV program in 1994 and 1995, respectively. Later, Maine and Vermont joined the group in the early 2000s. By 2019, the "LEV states" have expanded to twelve states, located on the east coast and west coast, where the ozone pollution problem is rather severe (Fig. 1.2; hereafter, I use "LEV states" to refer to the states that employed the LEV program, and "control states" to refer to states that have never employed the program in the contiguous United States). Nowadays, any vehicle produced in the United States must comply with either the California or the Federal standards, although some manufacturers' production lines might fit both standards. The temporal and spatial variation within the employment of the LEV program makes it a natural experiment, which allows this study to evaluate its impact on ozone pollution causally.

1.2.2 What is the LEV program?

For the purpose of this research, I want to first introduce a term intensively used in the LEV program the nonmethane organic gas or NMOG. The NMOG, as defined by the California Air

[5] California Clean Air Act of 1988.

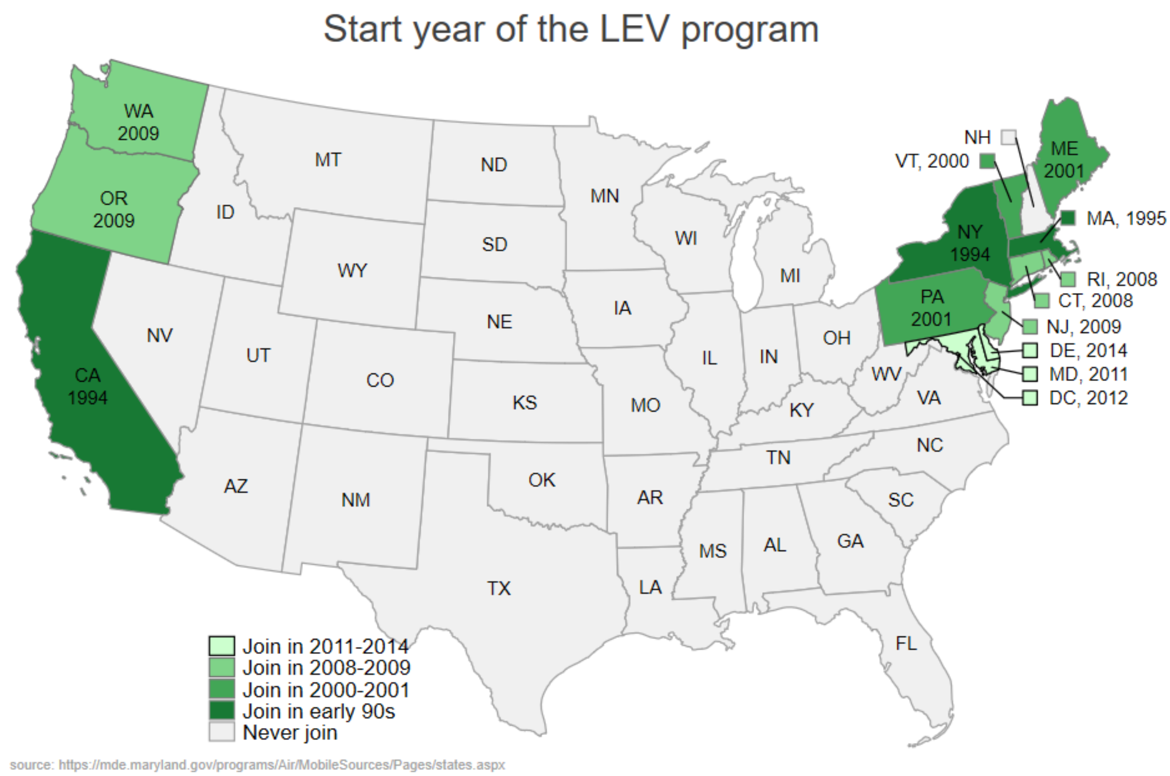


Figure 1.2 The LEV States in 2019

Note: This figure displays the states that have implemented the LEV program before 2019. Source: CARB.

Resources Board (CARB), is “the total mass of oxygenated and non-oxygenated hydrocarbon emissions”(CARB, 1990). It includes the traditionally regulated non-methane hydrocarbons (NMHC) and various types of oxygenated hydrocarbons, which are “highly reactive in forming ozone” (Reed Jr, 1996). The LEV program focuses on NMOG “to fully evaluate the ozone-forming potential of hydrocarbon emission from candidate vehicle/fuel systems” (CARB, 1990).

By applying regulation on the fleet-average NMOG, the LEV program has a similar design to the Corporate Average Fuel Economy (CAFE) standards. As table 1.1 illustrates, the LEV program specifies four types of vehicles: transitional low-emission vehicles (TLEVs), low-emission vehicles (LEVs), ultra-low-emission vehicles (ULEVs), and zero-emission vehicles (ZEVs). Over the years, vehicles needed to meet increasingly stringent emission requirements to be qualified as one of the four types. These certification standards are all stricter than the federal standards during the same period. On the other side, the manufacturers were given the flexibility of selling any combination of TLEVs, LEVs, ULEVs, ZEVs, or even conventional vehicles (which do not receive any of the

Table 1.1 California LEV Program: four category of certifications

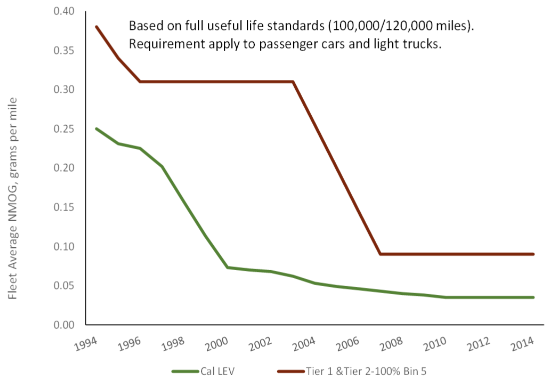
Category	NMOG	CO	NOx
Transitional low-emission vehicle (TLEV)	0.125	3.4	0.4
Low-emission vehicle (LEV)	0.075	3.4	0.2
Ultra-low-emission vehicle (ULEV)	0.040	1.7	0.2
Zero mission vehicle (ZEV)	0.000	0.0	0.0

Note: This table displays the 50,000 mile certification standards for passenger cars with weight less than 3750 lbs. The complete LEV program includes other certification standards based on vehicles' type, weight and durability basis. The unit used in the table is grams per mile. Source: NESCAUM (1991) 1.3.2.

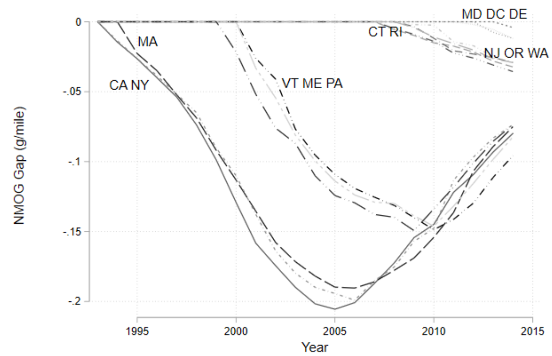
above four types of certifications), as long as sales-weighted emissions of all the vehicles sold meets the fleet average requirement of NMOG in any given year (figure 1.3). For example, as noted in one of the CARB staff reports, in the 1998 model year, the sales of one manufacturer in the California market were composed by 54% of the conventional vehicles certified to a 0.25 g/mi NMHC standard, 8% LEVs, 36% ULEVs and 2% ZEVs, and their fleet average NMOG complied with the LEV regulation (Reed Jr, 1996).[6] If manufacturers are unable to meet the fleet average NMOG requirement in a given year, the LEV program also has a credit borrowing and banking system. To reduce the manufacturers' burden of making the technological adjustment, they are allowed to run a deficit before 1998, though after that, each unfulfilled NMOG credit will cause a civil penalty of $5000 (Reed Jr, 1996).

As shown in Fig. 1.3a, although the gap between federal standards and the LEV's requirement is small for the model year 1994 vehicles, the LEV standards steadily become more stringent over the period of 1996 to 2003, thus the gap widens during that period. After 2004, the gap between the two standards start to shrink and almost disappears after 2008. By focusing solely on this differential in policy and making the assumption that targeted average emission rates are consistently met each year, together with the assumption that different states have identical fleet compositions, I conducted a simulation. This exercise aimed to quantify the potential emission reduction in LEV states, on average, attributable to the LEV policy. The results are depicted in Fig. 1.3b.

[6] The manufacturer's sales-weighted NMOG emissions average will be 0.1554: $(54\% \times 0.25) + (8\% \times 0.075) + (36\% \times 0.040) + (2\% \times 0.0) = 0.1554$. As a result, the fleet average NMOG falls below 0.157, the requirement for model year 1998.



(a) NMOG standards : LEV v.s. Fed Tier I & II



(b) Simulated NMOG Gap: LEV v.s. Fed Tier I & II

Figure 1.3 LEV v.s. Fed Standards

Note: Graph (a) compares the fleet-average NMOG requirement on passenger cars in the LEV program with the certification standards in FED Tier II Bin 5. Graph (b) compares NMOG reduction gains for each LEV state to results if federal standards were used. Source: California LEV program: an overview, Coralie Cooper, NESCAUM.

1.3 Data

This paper analyzes both the pollution impact and the mortality impact of the LEV program. For the pollution analysis, the paper relies on monitor level pollution data and station level weather data. For the mortality analysis, the paper uses county level mortality data and stationary source air pollutants emission data.

1.3.1 Ozone pollution data

Data on ground-level ozone pollution is obtained from the Air Quality System (AQS) database, which collects reports from EPA's air quality monitors. Data from 1990 to 2014 are used in this study, with a special focus on 1990 to 2003, when the LEV program was in its early stages[7]. To make the pollution metrics consistent with the NAAQS, the maximum daily ozone concentration is used, and monitoring days that capture fewer than nine hours of data are dropped.

In Fig. 1.4, I plot the annual average of ozone concentrations over the period 1990-2014. A comparison is made between LEV states and never-treated states based on the start time cohorts. As can be seen, the LEV states, especially those that joined LEV in 1994 or 1995, share a similar

[7]The first phase of the LEV program ran from 1994 to 2003. LEV programs and federal standards underwent major changes after 2003. The LEV program entered its second phase in 2004 and introduced stricter NMOG standards. Meanwhile, the federal standard (FED tier II) began regulating fleet-average NOx emissions, which is another precursor to ozone.

Figure 1.4 LEV vs. No-LEV: raw ozone trends by program start time

Note: The figure compares the raw ozone trend of LEV states and the never treated states. The red line indicates the year of 1995.

trend in ozone pollution with the rest of the states.

Table 1.2 provides the summary statistics of the sample of monitor-days used in this study. The sample includes 7,533,057 monitor-days, with coverage of 927 counties and 2228 distinct monitors, with the number of active monitors changing over time. Additionally, Table 1.2 shows that while mean ozone pollution have remained stable, the 99% percentiles of pollution decreased, which matches EPA's logic for controlling maximum ozone concentrations.

1.3.2 Weather data

Ground-level ozone pollution is significantly affected by the weather. In hot sunny days, the ozone level is relatively high since high temperatures and sunshine accelerate ozone's chemical formation (Bloomer, Stehr, Piety, Salawitch, & Dickerson, 2009). To control for the weather's impact, I use weather data from the National Climatic Data Center's Cooperative Station Data, and link the weather data with monitor level pollution data following methods described in Auffhammer and Kellogg (2011). In more detail, I gathered station-day records of rain, temperature, and snowfalls, and then applied the following three-step method. First, I connect every ozone monitor to the nearest ten "qualified" weather stations, within 50 miles, with an elevation difference under

Table 1.2 Summary statistics on monitors and ozone pollution

Year	Obs.	County	Monitor number				Maximum daily ozone			
			Total	Rural	Sub.	Urban	Mean	Std.	99%	1%
1990	243951	503	857	325	360	159	0.041	0.019	0.097	0.005
1991	252224	529	901	363	365	157	0.042	0.020	0.099	0.006
1992	262528	550	938	378	384	163	0.040	0.018	0.092	0.005
1993	271283	573	967	398	384	172	0.041	0.019	0.093	0.005
1994	276001	576	976	408	385	171	0.042	0.018	0.093	0.006
1995	278182	593	1010	434	395	170	0.043	0.019	0.096	0.007
1996	275817	601	1019	436	403	169	0.042	0.018	0.092	0.006
1997	286399	622	1047	458	409	171	0.043	0.018	0.090	0.007
1998	289582	628	1067	469	420	169	0.045	0.019	0.095	0.007
1999	295018	641	1091	474	429	178	0.045	0.019	0.094	0.008
2000	303237	658	1116	481	441	184	0.043	0.017	0.088	0.007
2001	314358	675	1157	499	457	191	0.044	0.017	0.088	0.008
2002	320999	686	1165	510	460	186	0.044	0.018	0.094	0.008
2003	325042	696	1180	513	473	186	0.043	0.017	0.087	0.008
2004	328405	709	1183	527	463	186	0.041	0.015	0.080	0.008
2005	324589	706	1169	521	452	189	0.043	0.017	0.084	0.008
2006	326907	708	1175	519	459	190	0.043	0.016	0.083	0.011
2007	332314	721	1196	544	456	189	0.044	0.015	0.082	0.011
2008	334221	726	1206	545	463	192	0.042	0.015	0.080	0.011
2009	339849	739	1215	559	462	188	0.040	0.013	0.074	0.011
2010	349120	751	1231	568	469	189	0.043	0.014	0.076	0.013
2011	286011	661	1060	512	386	157	0.043	0.014	0.077	0.012
2012	299062	669	1071	486	411	169	0.043	0.015	0.079	0.013
2013	310188	686	1104	502	416	182	0.041	0.013	0.072	0.012
2014	307770	690	1086	496	411	175	0.041	0.012	0.070	0.013
Total	7533057	927	2228	1062	771	376	0.042	0.017	0.088	0.008

Note: This table displays the summary statistics of daily maximum ozone records over 1990-2014. The daily ozone pollution records come from EPA's AQS. It should be noted that the sample excludes the monitor-days that lack weather information, as described in section 1.3.2. The measurement for ozone pollution is parts per million (ppm).

300 feet, and able to provide weather data for at least 50% of the monitored days. In the second step, I pick the closest weather station from the ten and impute its missing weather records based on other stations' observations.[8] Finally, I remove the monitor-day observations that don't have qualified weather info available. The process reduces the sample size by 2.7%.

[8] The imputing process involves two steps: First, regressing the data from the closest station on all remaining qualified weather stations. Next, we extrapolate the results to days when there are no data from the nearest station, but there are data from other qualified stations.

1.3.3 County-level mortality data

The paper uses the age-standardized mortality rate from a database created by Institute for Health Metrics and Evaluation (IHME). The database includes the mortality estimation among all counties in the United States from 1980 to 2014, and it is publicly available. The methods used for developing this database are described in Dwyer-Lindgren et al. (2016) and Dwyer-Lindgren et al. (2017). As stated in these two papers, the database's foundation is the de-identified death records from the National Center for Health Statistics, but instead of directly calculating a crude death rate, the authors apply the method of garbage codes redistribution and small area estimation to generate age-standardized mortality. By redistributing the death records related to "garbage codes" (i.e., codes that only record the intermediate death reason or not specific), the authors linked all death records to 21 mutually exclusive causes. By using a Bayesian spatially explicit mixed-effects regression model which incorporates the impact of demographics (race, income, and population density) and county-time trend, the author estimates the mortality for each county-year-age group and aggregate to get the age-standardized mortality rate (hereafter, the rate is referred as "*adjusted mortality rate*"). The adjusted mortality rate thus control for the influence of demographic variations across different counties. Table 1.3 provides a summary of the adjusted mortality rates that are used in the study.

This adjusted mortality rate from IHME is used by papers that are published in top medical journals (Nosrati et al., 2019; O'Connor, Sedghi, Dhodapkar, Kane, & Gross, 2018) but are seldom used in the economics literature. Thus, as a robustness check, I obtain the county-year level death certificate records from the National Center for Health Statistics multiple causes of death public-use data files for 1987–2014 via the CDC WONDER system. The CDC WONDER system provides county-year aggregated death records by the underlying cause of death, together with an estimate of county population. Using this information, an unadjusted mortality rate for respiratory diseases is calculated as deaths related to respiratory diseases divided by total population (hereafter, the rate is referred to as "*unadjusted mortality rate*"). As shown later, while the estimated coefficients are slightly different in magnitude, using an unadjusted mortality rate will deliver similar results.

Table 1.3 Summary statistics on health analysis data

Variable	Obs	Mean	Std	Min	Max
Adjusted Mortality rates (death per 100,000)					
All respiratory diseases	74,275	58.747	14.136	14.272	160.972
COPD	74,275	50.760	13.214	9.941	152.292
PNEU	74,275	0.748	1.624	0.101	46.415
Silicosis	74,275	0.054	0.084	0.001	2.472
Asbestosis	74,275	0.186	0.252	0.008	10.810
CWP	74,275	0.258	1.531	0.001	44.897
Asthma	74,275	1.785	0.631	0.488	7.029
Unadjusted Mortality rates (death per 100,000)					
Unadj. respiratory diseases	74,275	104.613	44.969	0	731.835
Other adj. mortality rate (for placebo test) (death per 100,000)					
Mental	74,275	8.357	5.485	1.278	73.154
Diabetes	74,275	60.453	16.362	11.548	183.462
Neuro.	74,275	90.566	20.752	14.481	212.162
Unint. injuries	74,275	23.434	4.803	7.621	78.018
Stationary source emission					
TRI-VOCs (ton)	74,275	0.176	0.741	0	48.269

Note: COPD - Chronic obstructive pulmonary disease; PNEU - pneumonia; CWP - coal workers' pneumoconiosis, Neuro. - neurological disorders.

1.3.4 County-level VOC emissions

To control for pollution from stationary sources, this study uses the EPA's Toxic Release Inventory (TRI) basic data files, which covers the self-reported annual emissions of each stationary source by mass and chemical. Using the map in Greenstone (2003) which links the TRI chemical to CAA pollutants, and applying the method provided in Gibson (2019), this study obtains the precise quantity of each facility's VOCs emission, and aggregates the emissions into the county-year level. There are two problems with using the emissions from stationary sources. First, only bigger facilities have their records in TRI, and smaller facilities are exempt from the annual self-reporting obligation. Thus, there might exist measurement errors in this variable of county level total VOCs emissions, even if we believe there is no error in the self-reporting process. Second, while both VOCs and NOx are precursors to ozone pollution, the TRI only provides information related to VOCs.

1.4 Empirical strategy

In this section, I describe the econometric method used to identify the impact of the LEV program.

1.4.1 Difference-in-Differences (DD)

Following Auffhammer and Kellogg (2011), I use the DD estimation equation as given below:

$$ln(ozone)_{icst} = \beta \cdot LEV_{st} + \gamma W_{icst} + \alpha_i + \Delta_t + \epsilon_{icst} \tag{1.1}$$

The dependent variable $ln(ozone)_{icst}$ is the natural logarithm of the daily ozone concentration of monitor i at date t in county c and state s. LEV_{st} is a dummy variable that captures whether the state s has implemented the LEV program at date t. The vector W_{icst} includes a set of weather variables, along with their square terms, cubic terms and interactions.[9] Coefficient α_i denotes the monitor fixed effect. The vector Δ_t is the set of time fixed effects, which includes the year dummies, the month dummies, and the day of week dummies. The ϵ_{it} denotes the idiosyncratic error term. The regression is weighted by an adjusted population which equals the total population of county c divided by the total monitors located in that county. The reported standard errors are clustered at the state-year level, so that ϵ_{icst} is allowed to be correlated within the state and year.

For the health analysis, I apply a similar DD estimation as follows:

$$ln(Mortality)_{cst} = \beta \cdot LEV_{st} + \gamma VOC_{cst} + \alpha_c + \Delta_t + \epsilon_{cst} \tag{1.2}$$

Similarly, $ln(Mortality)_{cst}$ is the natural logarithm of the mortality rate caused by certain diseases in county c state s and year t, and LEV_{st} is a dummy variable that equals 1 if state s has implemented the LEV policy in year t. The VOC_{cst} is the VOC emissions from stationary sources in county c state s and year t (obtained from the TRI). The α_c is the county fixed effect and $Delta_t$ is a time fixed effect (year dummy). To account for differences in population density across counties, the regression is weighted by population. The standard errors are clustered at the state level.

[9]Following Auffhammer and Kellogg (2011), the W_{icst} includes: $tmax$, $tmax^2$, $tmax^3$, $tmin$, $tmin^2$, $timn^3$, $rain$, $rain^2$, $rain^3$, $snow$, $snow^2$, $snow^3$, $tmax \cdot tmin$, $rain \cdot tmax$, $rain \cdot tmin$, $snow \cdot tmax$, $snow \cdot tmin$, $L.tmax$, $L.tmin$, $tmax \cdot L.tmax$, and $tmax \cdot L.tmin$.

The legitimacy of the above DD approaches hinges upon the assumption that the unobserved ϵ_{icst} are not correlated with LEV_{st} conditional on all controlled variables. Under this assumption, the estimated coefficient β is interpreted as the percent change in the daily maximum ozone concentration recorded by monitor i at date t relative to its counterfactual when the LEV does not exist.

However, the LEV states were self-selected to use their authority under section-177 and adopted California's LEV program at different years. This raises two issues. First, the problem of endogeneity: a) the LEV states might choose to join because they are relatively dirty; b) the LEV states might choose to join because they can benefit more in health outcomes. Both imply reverse causality and the existence of other unobserved confounding factors. For this, I plan to apply the synthetic control method to construct a counterfactual for each state that chooses to adopt the LEV.

Second, recent literature on staggered DD finds that the naive DD will yield a biased coefficient in a staggered treatment setting when heterogeneous treatment effects exist (Borusyak, Jaravel, & Spiess, 2021; Goodman-Bacon, 2018). To address this concern, I use the method proposed by Borusyak et al. (2021). First, I estimate the model with the control group observations only. Second, I extrapolate the model to the treated group, impute the counter factual outcome for each treated unit, and calculate the treatment effect for each individual monitor. Finally, I average the treatment effects to a single value. Additionally, I provide the estimates from using Wooldridge (2021) in Appendix 2 B.

1.4.2 Event study

I use the following event study models to show the dynamic impact of the LEV program on ozone pollution.

$$ln(ozone)_{icst} = \sum_{\tau=-4}^{10} \beta_\tau \cdot 1(t - start_s = \tau) + \gamma W_{icst} + \alpha_i + \Delta_t + \epsilon_{icst} \tag{1.3}$$

Here, the independent variable of interest LEV_{st} in Eq.(1.1) is replaced with the set of event time indicators $1(t - start_s = \tau)$. These variables indicate the relevant time difference between the current year t and the start year of the LEV program for each state $start_s$. For example, the indicator $1(t - start_s = 0)$ equals one only if year t is the beginning year of state s in joining LEV program. Thus, the β_τ with $\tau < 0$ correspond to the pre-treatment periods, and the β_τ with $\tau > 0$ correspond to the post-treatment periods. In practice, I exclude the year of $\tau = -1$, and collapse the time periods more than 4 years before and more than 10 years after into the "<= −4" and ">= 10" periods.

Similarly, the following equation is used for event study of the mortality impact:

$$ln(Mortality)_{cst} = \sum_{\tau=-4}^{10} \beta_\tau \cdot 1(t - start_s = \tau) + \gamma VOC_{cst} + \alpha_c + year_t + \epsilon_{cst} \tag{1.4}$$

Suppose that the common trend assumption holds, and the LEV states and the control states serve as good counterfactual for each other, then the β_τ should be zero before the enactment of LEV, and deviate from zero after it starts.

1.5 Results: ozone reduction

1.5.1 DD results

Table 1.4 presents the estimation results from the DD method during the period of 1990 to 2003. The column (1) display estimation results from the simplest DD model, and columns (2) to (5) add control variables step by step, moving from the basic DD in column (1) to the model that equation (1.1) specifies.

With only monitor fixed effects and year fixed effects, the estimated coefficient on *LEV* in column (1) is negative and statistically significant at the 1% level. It suggests that the LEV program's implementation reduced the concentration of ground-level ozone pollution of about 8.2%. Since mean ozone is about 0.042 ppm (see Table1.2), the point estimate implies about 0.0034 ppm decrease in ozone daily maximum. Column (2) adds a set of time dummies: the month dummies are used to control the seasonal trends of ozone pollution (Oltmans & Levy, 1992); the day of week dummies are used to control the within week variation of ozone concentration (Gao, 2007). Intuitively, the month dummies should control for variation caused by seasonal weather cycles and seasonal activities, especially agricultural activities, which are the primary source of ozone pollution in rural areas (Almaraz et al., 2018). The seasonal trends of ozone are depicted in

Table 1.4 The Impact of LEV on Log(*Ozone*) [1990-2003]

	(1)	(2)	(3)	(4)	(5)
LEV	-0.082***	-0.067***	-0.058***	-0.057***	-0.057***
	(0.019)	(0.017)	(0.015)	(0.015)	(0.014)
Year FE	Y	Y	Y	Y	Y
Monitor FE	Y	Y	Y	Y	Y
Month, day of wk.	N	Y	Y	Y	Y
Weather(linear)	N	N	Y	Y	Y
Weather(poly.)	N	N	N	Y	Y
Weather*Month	N	N	N	N	Y
Weather*day of wk.	N	N	N	N	Y
Groups					
Obs	3989482	3989482	3989482	3989482	3989482
R^2	0.004	0.309	0.392	0.416	0.441

Note: This table displays the regression results corresponding to Eq.(1.1). The dependent variable is the nature logarithm of the daily maximum ozone concentration. All standard errors are clustered at the state-year level. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Fig. A.1. The within week ozone variation also needs to be controlled because both the emission from stationary polluting facilities and the emission from mobile vehicles could be affected by weekdays/weekends. Adding the time controls reduces the point estimate on LEV variable to -0.067, while the point estimate is still significant at the 1% level. It also increases the regression's R^2 from a mere 0.004 to 0.309.

In column (3), the weather level variables are added, which further reduces the point estimate on LEV to -0.058, and increases the R^2 of the regression to about 0.392. In column (4), the weather polynomials and their interactions are added for better control, but they prove to have little effect on the point estimate of LEV. In column (5), I further allow the impact of weather to vary among months and day of the week. Given the flexibility of controls that are added, the estimates on LEV behave stably. It suggests that the LEV program's introduction reduced the daily maximum ozone concentration by about 5.7% on average, which equals to 0.0023 ppm, or 4.9 $\mu g/m^3$ reduction in absolute ozone concentration.[10]

As a comparison, Table 1.5 provides the results with the sample extended to 2014. As we can

[10] The average daily maximum ozone concentration over 1987-2003 is 0.043 ppm. Thus, a 5.7% reduction equals to $0.043 \times 5.7\% = 0.00245$ ppm. The conversion between ppm and $\mu g/m^3$ is $1ppb = 2.00\mu g/m^3$, and it is based on the technology document from Danish Centre For Environment and Energy.

Table 1.5 The Impact of LEV on Log(*Ozone*) [1990-2014]

	(1)	(2)	(3)	(4)	(5)
LEV	-0.035***	-0.031***	-0.020*	-0.020*	-0.022**
	(0.013)	(0.011)	(0.011)	(0.010)	(0.010)
Monitor FE	Y	Y	Y	Y	Y
Month, day of wk.	N	Y	Y	Y	Y
Weather(linear)	N	N	Y	Y	Y
Weather(poly.)	N	N	N	Y	Y
Weather*Month	N	N	N	N	Y
Weather*day of wk.	N	N	N	N	Y
Groups					
Obs	7526896	7526896	7526896	7526896	7526896
R^2	0.004	0.289	0.368	0.392	0.418

Note: This table displays the regression results corresponding to Eq.(1.1). The dependent variable is the nature logarithm of the daily maximum ozone concentration. All standard errors are clustered at the state-year level. $*$ $p < 0.1$, $**$ $p < 0.05$, $***$ $p < 0.01$.

see, the key coefficients on LEV stay negative and significant, but the magnitude of the coefficients is smaller. The diminished impact of the policy could be attributed to the 'U'-shaped pollution gap illustrated in Fig. 1.3b. This leads us to the event study of the next section.

As noted in section 1.4.1, with the staggered setting and the potential heterogeneity in the treatment effect, the reliability of our DD estimate is challenged. To mitigate this worry, I applied the method proposed by Borusyak et al. (2021), and the results are provided in Table 1.6. The results shows that the BJS estimates are larger than the traditional DD estimates, which indicate the above estimates are biased toward zero. It should be noted that due to the computation complexity, Table 1.6 is provided with the yearly panel, not the daily panel used in Table 1.4 and 1.5.

1.5.2 Event study

The remaining question here is the validity of DD's parallel trends assumption. Fig. 1.5 plots the point estimates of all β_τs in Eq.(1.3), and the shaded area represents the 95% confidence interval of the estimates. As shown in Fig. 1.5, the difference between the LEV states and the control states is small and statistically not different from zero for all years before the LEV program's enactment. After the policy change happens, the gap between the two groups of states starts to deviate from zero, and the gap widens over the next six years then start to shrink. This trend is consistent with Fig. 1.3. The LEV program employs a progressively stringent requirement on NMOG, and the

Table 1.6 BJS Imputation with yearly data

	(1) BJS	(2) DiD
Panel A: 1990-2003		
LEV	-0.069***	-0.055***
	(0.017)	(0.022)
Obs	13,770	13,907
Panel B: 1990-2014		
LEV	-0.052***	-0.038***
	(0.012)	(0.012)
Obs	25,225	25,366
Monitor FE	Y	Y
Year FE	Y	Y

Note: This table displays the regression results corresponding to Eq.(1.1) with yearly panel. The dependent variable is the natural logarithm of the daily maximum ozone concentration. All standard errors are clustered at the state-year level. $* \; p < 0.1$, $** \; p < 0.05$, $*** \; p < 0.01$.

difference between LEV and FED standards shall accumulate given the older vehicles are replaced over the years. But at the same time, given that the federal standards started to catch up the LEV standards around 2008, the program's impact on ozone would ultimately disappear.

1.5.3 Robustness and heterogeneity: ozone reduction

1.5.3.1 Heterogeneity: by month and by region

In this section, I check the heterogeneity of LEV's impact in time and region. First, I split the 1990-2003 data sample into twelve sub-samples for each month, perform the previous DD analysis, and plot the coefficients on LEV and their 95% confidence intervals in Fig. 1.6. It can be seen that the LEV program's impacts vary across months and peak in the summer months. As noted before, the production of ozone needs sunlight and high temperature, which makes the pollution problem much more severe during summer time (Fig. A.1). Thus, the LEV's ability of reducing ozone pollution during the most polluted months could generate more benefits.

For the question of where the pollution reduction happens, Table 1.7 displays the results with three sub-samples: rural, suburban, and urban. It shows that the significant reduction happens in suburban and urban areas, while rural places benefit less from the LEV program and only experience

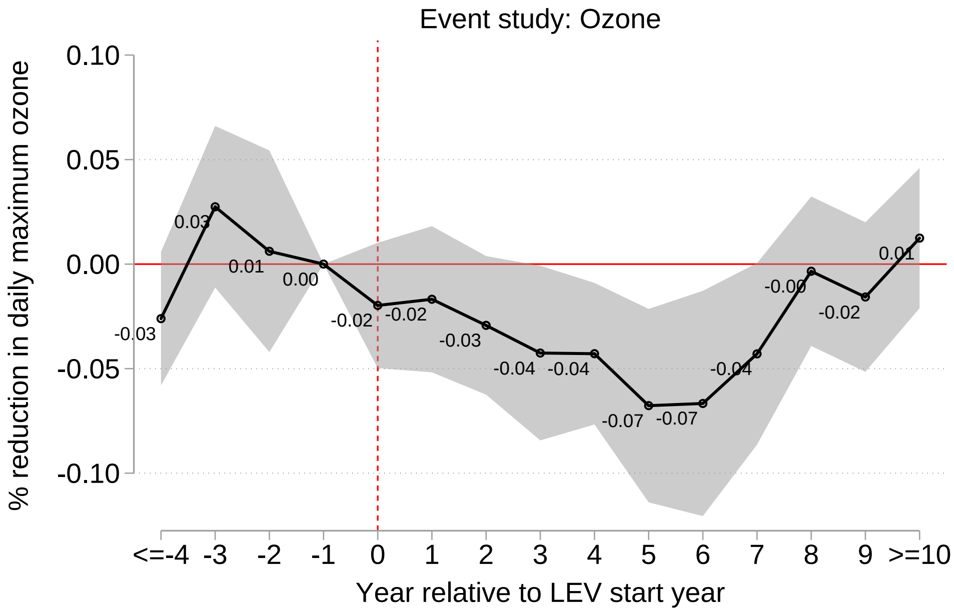


Figure 1.5 The Ozone pollution difference between LEV states and control states by year

Note: This figure plots the regression results obtained using Eq.(1.3). The dots represent the point estimates of β_τ, and the shaded areas represent the 95% confidence intervals. The vertical red line denotes the start of the LEV program.

a 1.2 percent reduction in ozone. However, the ozone reduction could still generate huge welfare effects in rural areas, since ozone alone accounts for 90% of the crop loss in the United States that could be attributed to air pollution (Heck et al., 1982). Applying the results from the most recent research (Metaxoglou & Smith, 2020), one can work out that the LEV program's impact could increase corn yield by 1.2 bushels per acre, which values around \$16 million in the LEV states, or \$410 million for the United States if all adopted.[11]

1.5.3.2 Are the results robust to AQS monitoring systems' expansion?

As discussed in section 1.3.1, over the years, the AQS monitoring systems covers more counties, which makes the sample unbalanced and may bring bias into the estimation. To resolve this concern,

[11] According to Metaxoglou and Smith (2020), 1 ppb increase in ozone concentration could decrease the corn yields by 2.23 bushels per acre ($1\ ppm = 1,000\ ppb$). The average daily maximum ozone concentration in rural during my study period is 0.045 ppm. Thus, a 1.2% reduction equals to $0.045 \times 1.2\% = 0.00054\ ppm = 0.54\ ppb$ reduction. Suppose the impact of ozone pollution is symmetric, the LEV program could increase the corn yields by $0.54 \times 2.23 = 1.2$ bushels per acre. According to The United States Department of Agriculture, in 2003, there are 78,603,000 acres of corn planted in the United States, and 3,114,000 acres are in the LEV states. Thus, the aggregated corn yields increase will be 3,736,800 bushels for the LEV states or 94,323,600 bushels for the United States, which values around \$8.5 million and \$0.21 billion respectively.

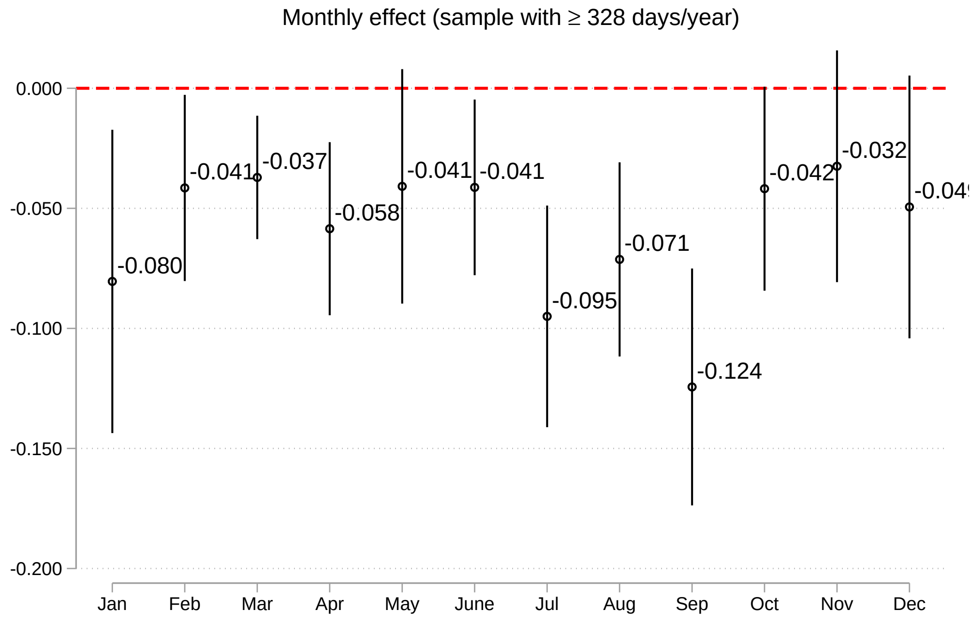


Figure 1.6 The LEV's impact: by month

Note: This figure plots the regression results obtained by using Eq.(1.1) on 12 sub-samples of each month. The dots represent the point estimates of β, and the vertical lines represent the 95% confidence intervals.

Table 1.7 The Impact of LEV on $ln(Ozone)$ [Rural vs. Suburban vs. Urban; 1990-2003]

	(1) Urban	(2) Suburban	(3) Rural
levpostD	-0.067*	-0.053***	-0.012*
	(0.032)	(0.011)	(0.006)
Monitor FE	Y	Y	Y
Month, day of wk.	Y	Y	Y
Weather(linear)	Y	Y	Y
Weather(poly.)	Y	Y	Y
Weather*Month	Y	Y	Y
Weather*day of wk.	Y	Y	Y
Obs	683578	1613574	1651062
R^2	0.459	0.455	0.416

Note: This table displays the regression results corresponding to Eq.(1.1). The dependent variable is the nature logarithm of the daily maximum ozone concentration. All standard errors are clustered at the state-year level. $*$ $p < 0.10$, $*$ $p < 0.05$, $***$ $p < 0.01$.

I select a sub-sample of the data, where all monitors must have been kept working for 14 years from 1990 to 2003 (hereafter referred to as "14-year monitors"), and replicate the main analysis in section 1.5.1. Table 1.8 present the estimation results with this sub-sample. We can see that there

Table 1.8 The Impact of LEV on Log of ozone Pollution (14-year monitors only)

	(1)	(2)	(3)	(4)	(5)
LEV	-0.089***	-0.077***	-0.068***	-0.067***	-0.068***
	(0.020)	(0.018)	(0.016)	(0.016)	(0.015)
Year FE	Y	Y	Y	Y	Y
Monitor FE	Y	Y	Y	Y	Y
Month, day of wk.	N	Y	Y	Y	Y
Weather(linear)	N	N	Y	Y	Y
Weather(poly.)	N	N	N	Y	Y
Weather*Month	N	N	N	N	Y
Weather*day of wk.	N	N	N	N	Y
Groups					
Obs	2137615	2137615	2137615	2137615	2137615
R^2	0.005	0.325	0.401	0.426	0.457

Note: This table displays the regression results corresponding to Eq.(1.1), but only includes monitors which have continuously worked for 14 years. The dependent variable is the natural logarithm of the daily maximum ozone concentration. All standard errors are clustered at the state-year level. $* \ p < 0.10, * \ p < 0.05, * * * \ p < 0.01$.

is no substantial change in the magnitude of the key coefficient.

Table 1.9 The Impact of LEV on $ln(Ozone)$ [drop border monitors; 1990-2003]

	(1)	(2)	(3)	(4)	(5)
LEV	-0.086***	-0.074***	-0.064***	-0.062***	-0.062***
	(0.019)	(0.017)	(0.016)	(0.016)	(0.015)
Monitor FE	Y	Y	Y	Y	Y
Month, day of wk.	N	Y	Y	Y	Y
Weather(linear)	N	N	Y	Y	Y
Weather(poly.)	N	N	N	Y	Y
Weather*Month	N	N	N	N	Y
Weather*day of wk.	N	N	N	N	Y
Groups					
Obs	3680725	3680725	3680725	3680725	3680725
R^2	0.004	0.298	0.385	0.409	0.438

Note: This table displays the regression results corresponding to Eq.(1.1). The dependent variable is the natural logarithm of the daily maximum ozone concentration. All standard errors are clustered at the state-year level. $* \ p < 0.10, * \ p < 0.05, * * * \ p < 0.01$.

1.5.3.3 Are the results robust to dropping border monitors?

The impact of the LEV program, might spillover to nearby locations in a different state but in the same commuting zone. For example, if someone who works in New York city (joined LEV

Table 1.10 The Impact of LEV on SO_2 [1990-2014]

	(1)	(2)	(3)	(4)	(5)
LEV	-0.135	-0.142	-0.040	-0.027	-0.017
	(0.448)	(0.447)	(0.442)	(0.446)	(0.447)
Monitor FE	Y	Y	Y	Y	Y
Month, day of wk.	N	Y	Y	Y	Y
Weather(linear)	N	N	Y	Y	Y
Weather(poly.)	N	N	N	Y	Y
Weather*Month	N	N	N	N	Y
Weather*day of wk.	N	N	N	N	Y
Groups					
Obs	4608230	4608230	4608230	4608230	4608230
R^2	0.055	0.070	0.086	0.099	0.104

Note: This table displays the regression results corresponding to Eq.(1.1). The dependent variable is the daily maximum of 3h SO_2 concentration. All standard errors are clustered at the state-year level. $* \, p < 0.10$, $* \, p < 0.05$, $*** \, p < 0.01$.

since 1994) has a house in New Jersey (joined LEV since 2008) and buys a car there for commuting, then the estimated impact of the LEV in NY might be biased because that car which is used in NY only follows the emission requirement in NJ. To test for such bias, I drop border monitors and conduct the DD analysis.[12] As Table 1.9 shows, the coefficients are slightly larger in the magnitude than the main results, but there is no major change.

1.5.3.4 Placebo test: LEV's impact on sulfur dioxide

In this placebo test, I perform the DD analysis and the event study on the concentration records of SO_2. If the observed changes in $Ozone$ were due to changes in factors other than the LEV regulation, such as improvements in manufacturing activity, burning coal, or electric power generation, then we would also expect some trend in SO2. However, as Table 1.10 and Fig. 1.7 shows, the SO_2 concentration is not affected by the employment of LEV policy. The coefficients in Table 1.10 are all insignificant, and there is no clear trends which deviate from zero in figure 1.7. This provides additional support that the ozone reduction is driven by the LEV program.

[12]Border monitors are defined as monitors located in LEV states' border counties or LEV states' border counties.

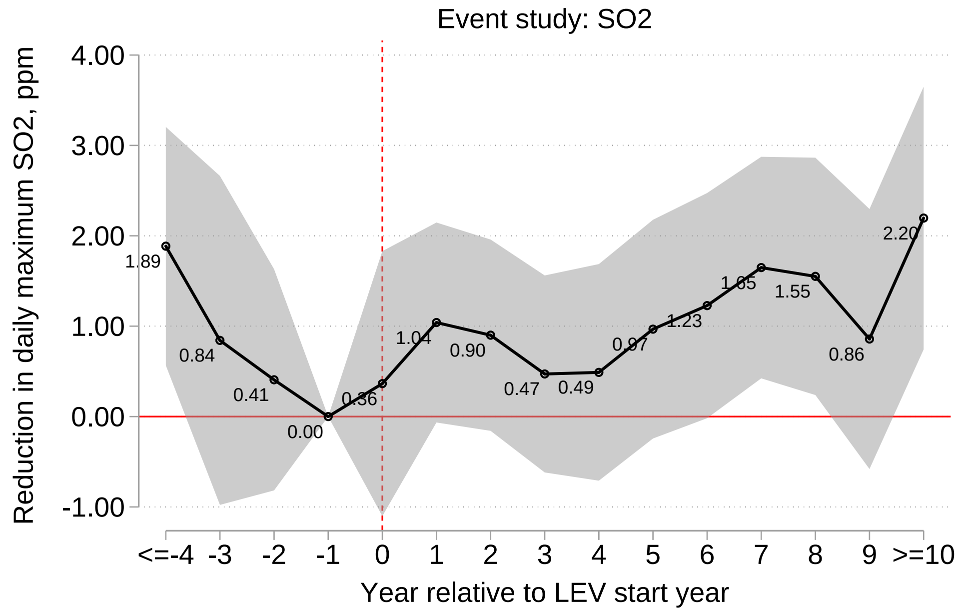


Figure 1.7 The Annual Effect of LEV on SO_2

Note: This figure plots the regression results obtained by applying Eq.(1.3) to the SO_2 concentration. The dots represent the point estimates of β_τ, and the shaded areas represent the 95% confidence intervals. The vertical red line denotes the start of the LEV program.

1.6 Mortality effect

1.6.1 Mortality: DD results

In this part, I present the mortality analysis results. The difference-in-differences results on the health outcomes are displayed in the Table 1.11. In column (1) and (2), it can be seen that the LEV reduces the mortality rate in respiratory diseases by about 5.7%. Given the average mortality rate of 58.7, the LEV policy contributes to a reduction of around 3.3 death per 100,000, or around 10,000 fewer deaths per year assuming all states implemented. Column (3) and (4) present the results when the dependent variable is replaced by the unadjusted mortality rate. The results suggest that the mortality rate is reduced by about 6.3%, which is close to the results from adjusted mortality. To further reveal the impact of LEV on health, table 1.12 provides the analysis on the six subcategories of the respiratory diseases. As we can see, the estimated coefficients are all negative and significant. But some of the coefficients seem to have super large magnitudes, like silicosis, asbestosis, and CWP. One possible explanation might be that we are considering mortality. Thus, the “harvesting” effect might be more significant for relatively severe diseases. As a placebo test, I apply the same Eq. (1) regression to several other causes of death that have not been linked to air pollution: mental illness, diabetes, neurological disorders, and unintended injuries. As shown in Table 1.13, none of these diseases is affected by the introduction of the LEV policy.

Table 1.11 The Impact of LEV on Mortality of Respiratory Diseases

	(1)	(2)	(3)	(4)
	All resp.	All resp.	Unadj. Resp	Unadj. Resp
LEV	-0.057***	-0.057***	-0.063***	-0.063***
	(0.018)	(0.018)	(0.021)	(0.021)
TRI_ozone		0.001		0.001
		(0.002)		(0.004)
County FE	Y	Y	Y	Y
Year FE	Y	Y	Y	Y
N	74275	74275	74027	74027
R-sq	0.92	0.92	0.80	0.80

Notes: The table presents the results from Eq.(2). Regressions are weighted by population. Standard errors are in parentheses and clustered at the state level. The "All resp." data comes from Dwyer-Lindgren et al. (2016), and the "Unadj. resp." data comes from CDC Wonder database. (*p<0.10, ** p<0.05, *** p<0.01)

Table 1.12 The Impact of LEV on Mortality of Respiratory Diseases (by subcategory)

Panel B						
	(1)	(2)	(3)	(4)	(5)	(6)
	COPD	PNEU	Silicosis	Asbestosis	CWP	Asthma
LEV	-0.066***	-0.123**	-0.305***	-0.131***	-0.456***	-0.115***
	(0.021)	(0.061)	(0.066)	(0.049)	(0.056)	(0.035)
TRI_ozone	0.002	-0.003	0.017	0.001	0.017	0.006
	(0.002)	(0.003)	(0.013)	(0.007)	(0.014)	(0.004)
County FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
N	74275	74275	74275	74275	74275	74275
R-sq	0.92	0.97	0.94	0.96	0.97	0.93

Notes: The table presents the results from Eq.(2). Regressions are weighted by population. Standard errors are in parentheses and clustered at the state level. COPD-Chronic obstructive pulmonary disease; PNEU- pneumonia; CWP - coal workers' pneumoconiosis (black lung). (*p<0.10, ** p<0.05, *** p<0.01)

Table 1.13 The impact of LEV on other unrelated Diseases

	(1) Diabetes	(2) Diabetes	(3) Neuro.	(4) Neuro.	(5) Unint. injuries	(6) Unint. injuries
LEV	-0.002	-0.003	-0.011	-0.011	-0.045	-0.046
	(0.026)	(0.026)	(0.021)	(0.021)	(0.028)	(0.028)
TRI_ozone		0.004		-0.002		0.007
		(0.003)		(0.003)		(0.006)
County FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
N	74275	74275	74275	74275	74275	74275
R-sq	0.94	0.94	0.94	0.94	0.94	0.94

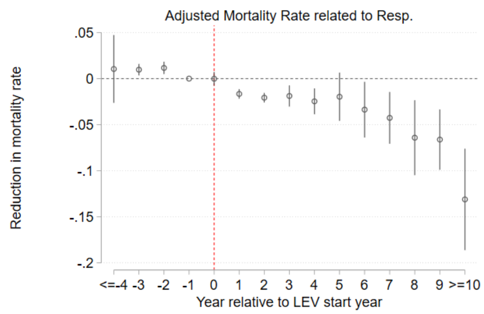
Notes: The table presents the results from Eq.(2). Regressions are weighted by population. The four diseases that are tested as placebo here are: Diabetes, neurological disorders, and unintended injuries. Standard errors are in parentheses and clustered at the state level. (*p<0.10, ** p<0.05, *** p<0.01).

1.6.2 Health: event figures

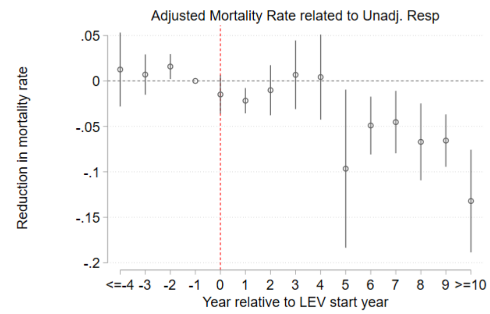
Fig. 1.8 plots the point estimates of all coefficients β_k in Eq.(1.2) and their 95 percent confidence intervals. In general, we cannot see clear evidence of parallel pre-trends, especially for the subcategory of respiratory diseases. Moreover, following the implementation of the LEV, we observe that the magnitude of the coefficients continues to increase, becoming particularly large and negatively significant in the ">= 10" category, ten years after the policy began. Consider the fact that only the early adopters are counted for calculating β_{10}, this might also be the evidence supporting the idea that early adopters choose to employ the policy because they expect to benefit more from the policy. Given such explanations, a synthetic control seems to be necessary for establishing the rigorous causal relationship between LEV and mortality.

1.6.3 Health: synthetic control

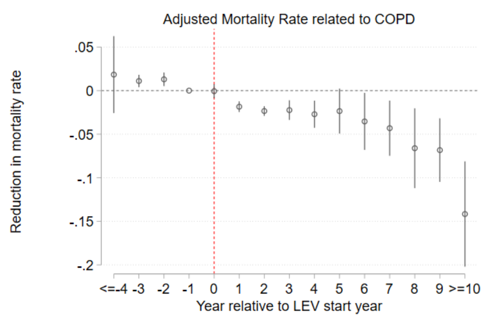
In the previous analysis, I estimated the impact of LEV policy based on DD methods and presented the trend of the outcome variable pre and post-treatment using event study. The method, however, may still suffer from endogeneity problems: LEV states are self-selected to implement the LEV program, so they may be the ones who benefit most from it, either in pollution reduction or mortality reduction. To mitigate such concerns, I implement the synthetic control method described



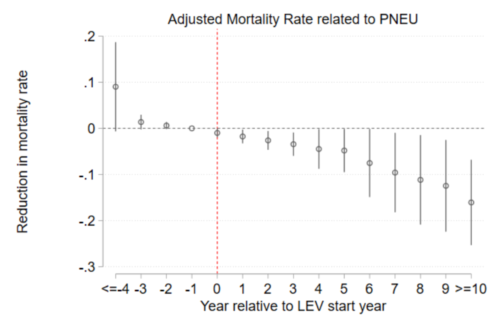
(a) Adjusted mortality of resp.



(b) Un-adjusted mortality of resp.

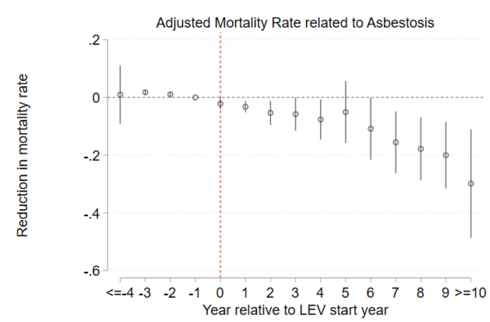


(c) COPD

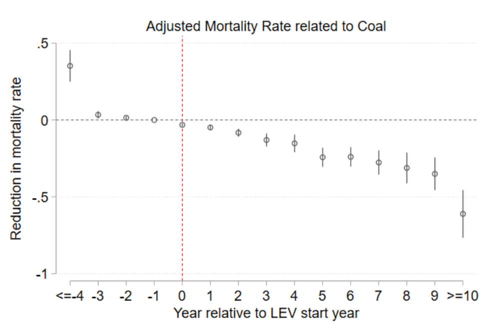


(d) PNEU

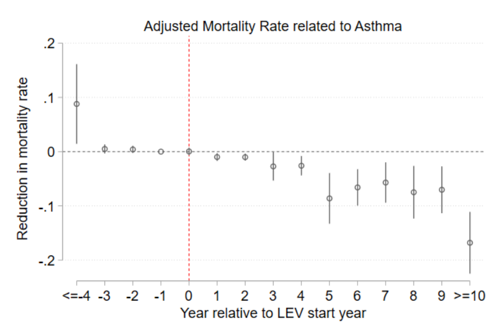
(e) Silicosis



(f) Asbestosis



(g) CWP



(h) Asthma

Figure 1.8 LEV's impact on mortality rate of respiratory diseases (by subcategory)

Note: The unit of the mortality rate is deaths per 100,000 people.

in Abadie and Gardeazabal (2003) and Abadie, Diamond, and Hainmueller (2010), and conduct studies that account for unobserved county-specific factors over time:

$$Y_{cst} = \beta \cdot LEV_{st} + \alpha_c year_t + \epsilon_{cst} \tag{1.5}$$

Here, Y_{cst} is the outcome variable for county c in state s when the year is t. The $\alpha_c year_t$ represents the unobserved county-specific time trend. Since we cannot truly observe the coefficient α_c for each county c, we use the synthetic control methods to construct a "synth" county as counterfactual for each treated county that has employed LEV program. Specifically, I use the similar method that has been developed in recent literature to deal with multiple treated units and multiple treated period (Dickert-Conlin, Elder, & Teltser, 2019; Dube & Zipperer, 2015). First, I denote the 363 treated counties in the LEV states as $j \in \{1, 2, 3, ..., 363\}$, and for each treated county j, a "donor county pool" is created by untreated counties that has outcome variable located in the same quintile as county j. This pool of these donor counties are denoted as P_j. Then, I use the synthetic control method to construct a weighted average by choosing weights such that the weighted combination of the outcome variable in donor county pool P_j could mimic the trends of the treated county j during the pre-treatment period:

$$W_j = \arg \min_{w_p^j \in [0,1]} \sum_{t=\tau_j-4}^{\tau_j-1} \left(Y_{jt} - \sum_{p \in P_j} w_p^j Y_{pt} \right)^2 \tag{1.6}$$

In this expression, the weights w_p^j would be sum to 1 across all donor counties p in P_j. The τ_j denotes the start year of LEV in county j. Thus, the weighting matrix W_j are determined by minimize the differences between the treated county j and the combination of control counties p during the four-year pre-treatment period.

Then I define the treatment effects θ_j for each treated county j as the mean difference between the treated county j and its "synth" counties, before and after the treatment, following the intuition of difference-in-differences method:

$$\theta_j = \frac{1}{\sum_t \mathbf{1}(t \geq \tau_j)} \sum_{t:t \geq \tau_j} \left(Y_{jt} - \sum_{p \in P_j} w_p^j Y_{pt} \right) - \frac{1}{\sum_t \mathbf{1}(t < \tau_j)} \sum_{t:t < \tau_j} \left(Y_{jt} - \sum_{p \in P_j} w_p^j Y_{pt} \right) \tag{1.7}$$

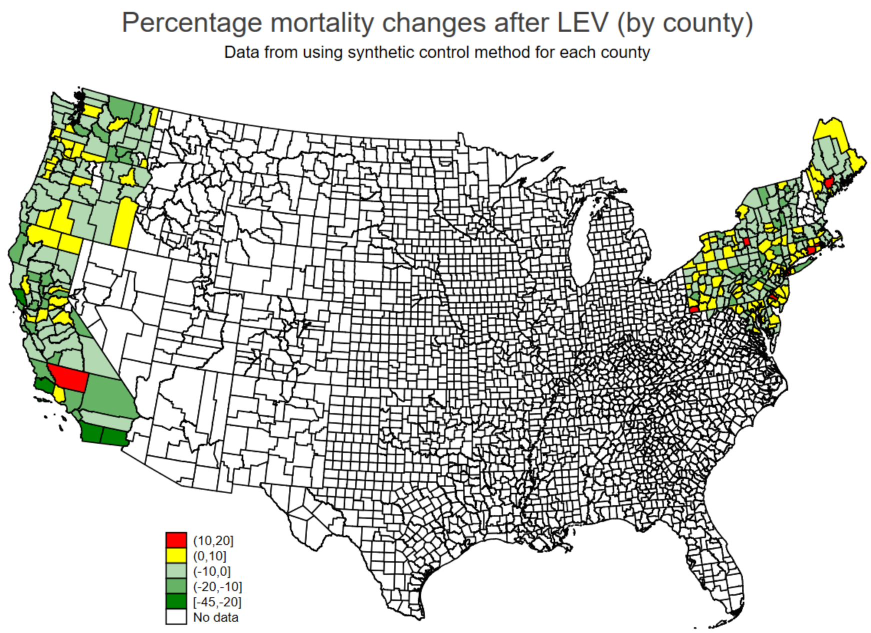


Figure 1.9 The Effect of LEV on mortality rate of respiratory diseases

Note: This map plots estimated change in mortality (θ_j) for each treated county obtained from applying synthetic control methods (Eq:1.7).

Here, $\mathbf{1}(t \geq \tau_j)$ is the dummy indicator which equals to 1 when year t is larger than τ_j, the start year of LEV in treated county j. Similar definition applies to $\mathbf{1}(t < \tau_j)$

In Fig.1.9, I choose respiratory mortality as the outcome variable, and plot the θ_j for each treated county. As the map shows, there exits strong heterogeneity across counties. We observe reductions in the respiratory diseases among the majority of the counties, while some counties seem to experience mortality increase. How should we interpret these θ_js? Would they be significant? To illustrate this, I apply the placebo-based rank test introduced by Dube and Zipperer (2015). In this test, analogous θ_p^j are calculated for each donor county p in the pool P_j. For example, if P_j includes N donor counties, then, when one donor county p is used as the placebo "treated" counties, the rest $N-1$ counties in P_j are used as the donor counties. In this placebo-test, all donor counties' LEV start year are set to be the same as the real LEV start year of the treated county j. After obtaining all N different θ_p^j, the real treatment effects θ_j were compared to the set of $\{\theta_p^j\}$, and the relative rank percentile of θ_j is denoted as r_j. As Dube and Zipperer (2015) argues, we can reject the null hypothesis of $\theta_j = 0$ at the significance level of five percent when the $r_j < 0.025$

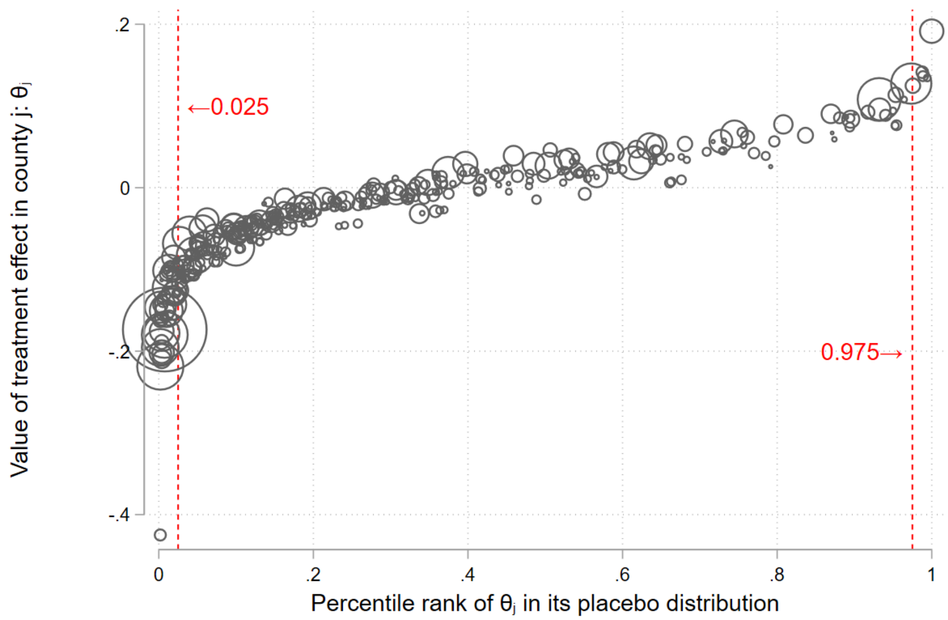


Figure 1.10 Test the significance of all θ_j

Note: This map plots θ_j for each treated county obtained from applying synthetic control methods (Eq:1.7).

or $r_j > 0.975$, that is, when the real θ_j lies in the tail of the corresponding empricial placebo distribution of θ_p^j.

Fig.1.10 plots all 363 different θ_j and their percentile ranks in a bubble chart, where the bubble size is proportional to the average population during 1990-2014 of the county. In this figure, there are more counties located in the lower left corner, especially left of the 0.025 line, which indicates that the large, negative treatment effects estimates of these counties are significant at the five percent level. In the middle of the figure, between the two line of 0.025 and 0.975, counties with estimates close to zero, either positive or negative, are insignificant. In the upper right corner, there are several counties with positive and significant treatment effects.

$$\bar{\theta} = \frac{1}{363} \sum_{j=1}^{363} \theta_j \tag{1.8}$$

$$\bar{r} = \frac{1}{363} \sum_{j=1}^{363} r_j \tag{1.9}$$

As a final step, all 363 different θ_j are averaged to get an aggregated $\bar{\theta}$, which is equal to -0.031 (Eq.1.8), and the corresponding $\bar{r}$ is 0.29 (Eq.1.9). Again, we follow the idea provided in Dube and Zipperer (2015) for the inference of $\bar{\theta}$. Under the null hypothesis of no impact from LEV, every r_j, the percentile rank of θ_j among its placebos, should be a random variable follows the [0,1] uniform distribution. Thus,the mean of r_j, $\bar{r}$, should be close to 0.5. Also, by central limit theorem, $\bar{r}$ shall follow an appropriately scaled normal distribution $N\left(0.5, \sqrt{\frac{12}{363}}\right)$. Given this, we can derive that $\bar{r} = 0.29$ is small enough that we can reject the null hypothesis of $\bar{r} = 0.5$ at the 1% significance level. As a result, we can reject the null hypothesis of $\bar{\theta} = 0$.

By using synthetic control method, I find a significant mortality reduction, which echoes my results in DD analysis. The synthetic results also illustrate heterogeneity, which raises the question of why the same LEV program has different impacts in different places. In the future, I plan to replicate the previous studies using more accurate data, such as infant death records.

1.7 Conclusions

Fifty years after the CAA, ground-level ozone pollution is still a severe problem in the United States. Ozone pollution is closely related to mobile emission, while existing literature lacks discussion about mobile emission standards.

In this study, by using the geographical and temporal variation in California's Low Emission Vehicle program, I develop an ex-post evaluation of the program's impact on ozone pollution. By applying a difference-in-differences method, I compare the ozone concentration changes in state that implemented the LEV program with the states that did not. The results show that the LEV program significantly reduces the ozone concentration by about 5.7%. The observed effect is slightly larger during summer times, and it is relatively smaller in rural regions. My results are robust to two kinds of confounding factors: the expansion of the ozone monitor network, and the impact of monitors located in state borders. Using a similar method, I evaluate the impact of LEV on mortality from respiratory diseases, and I find a reduction of 5.7% in mortality, which is equal to about 3.3 deaths per 100,000 people.

The event study analysis reveals a perplexing result: the program seems to have continuous impact even ten years after the program's start year, even though the pollution analysis reveals gap closes after seven years. This disparity might indicate there are other ways except the ozone channel that also influence the health outcomes. Alternatively, there could be lagged effects from exposure to pollution in previous years. Nevertheless, there are several limitations to this study. First, it relies on the monitor data for pollution analysis, while the reliability of monitor data faces challenges in recent years (Grainger & Schreiber, 2019; Zou, 2021). Second, the event study shows that the pre-trend parallel assumption might not hold, thus, the DD estimates might be biased. Third, the paper does not provide any evidence about the cost of the LEV program. How much does it cost for manufacturers to build a vehicle that fits the stricter California standards? How much was this cost reflected in the vehicle's sale price? Does this change consumers' vehicle preferences? These questions are essential for developing a complete evaluation of the LEV program, but are not answered in this study.

In conclusion, this study adds to the literature on evaluating environmental regulations, by providing the ex-post empirical evidence on California's LEV program. Although the benefit analysis has limitations, the study builds the first step to monetize the impact of one vehicle emission standard that has been implemented in the United States for about half a century and is currently on the battlefield for its legitimacy.

CHAPTER 2

MONITORING AND INSPECTIONS: THE CORRELATION BETWEEN MONITOR BREAKINGS AND INSPECTIONS

2.1 Introduction

Since the enactment of the Clean Air Act (CAA) in 1970, the United States has made significant strides in addressing air pollution, as demonstrated by the progress shown in Figure 2.1. Like other major U.S. environmental regulations, the CAA centralizes legislation but delegates inspection responsibilities to local authorities (Shimshack, 2014). For instance, state regulators conduct 86.8% of inspection activities under the CAA, while county regulators account for approximately 10%, and the EPA directly initiates only 3.2%[1].

The system that depends on state and local regulators presents both advantages and challenges. On the one hand, it enables the use of local information, which may potentially enhance regulatory efficiency within the constraints of available resources (Duflo, Greenstone, Pande, & Ryan, 2018). On the other hand, it raises concerns about possible distortions in regulatory incentives. Recent studies have found that new monitoring stations tend to be situated in relatively cleaner regions (Grainger & Schreiber, 2019; Grainger, Schreiber, & Chang, 2018), and that higher pollution levels coincide with off days for PM2.5 monitors (Mu, Rubin, & Zou, 2021; Zou, 2021). These findings suggest that the current regional air pollution regulation practices may have certain limitations. In recent years, the EPA's budget has been under increasing constraints, leading to the delegation of more regulation responsibilities, as well as greater discretionary power, to regional regulators (Hejny, 2018). As a result, there is a growing need to better comprehend the strategies employed by local environmental regulators.

In this study, I concentrate on relationship between monitoring and inspection allocations in the regulation of ground-level PM2.5 pollution. I developed a theory model to predict the local regulators' behavior under penalty when there is a reduction in monitoring activities. Then I leverage the variation stemming from changes in the data completeness status of PM2.5 monitors.

[1] Calculated by the author, and the data is from EPA's ECHO database, records of inspections.

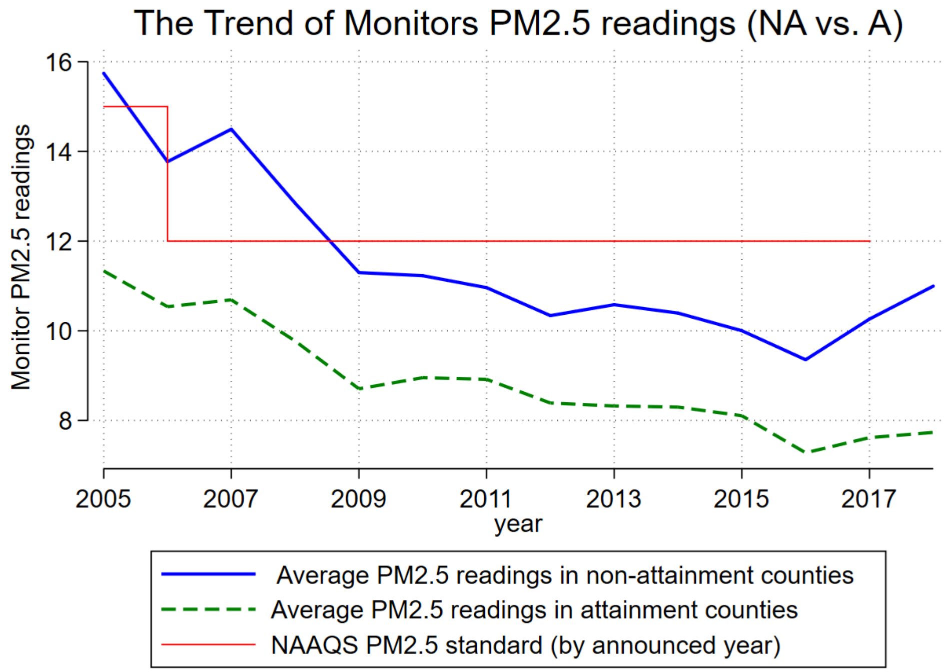


Figure 2.1 The Trends of PM2.5 Pollution and Regulation changes

Note: the data is from EPA's Air Quality System (AQS). The red line represents the change of National Ambient Air Quality Standards in ground-level PM2.5 over the years. The blue line is the average PM2.5 pollution level in non-attainment counties, and the green line is the average PM2.5 pollution level in attainment counties.

It's essential to understand that monitors with incomplete data cannot be used in the EPA's process of determining whether a county or state meets the National Ambient Air Quality Standards (NAAQS). Consequently, areas associated with these monitors are designated as "unclassifiable." In certain instances, an entire state may receive an "unclassifiable" designation due to the absence of complete monitor data, as was the case for Illinois between 2011 and 2013. By examining how regulators respond to these changes in monitoring intensity, we aim to identify the factors influencing regulators' decision-making processes and address the following questions: How would regulators reallocate inspection resources in the absence of federal-level pressure (i.e., when the risk of violating NAAQS decreases due to reduced monitoring intensity)? Are there notable differences between the strategies employed by state regulators and county regulators? Will regulators in different regions employ different strategies?

By examining the both the monitoring and inspection aspects of regulatory strategies and their implications, this study contributes to the broader understanding of the challenges associated with air pollution regulation under the CAA. Additionally, it highlights the importance of addressing the incomplete monitors' problem and sheds light on potential areas for further investigation and improvement in regulatory practices.

The structure of this paper is outlined as follows: Section 2 offers a review of the existing literature concerning air pollution regulation, with a particular focus on monitoring. Section 3 provide the policy background. In Section 4, we introduce a theoretical model designed to analyze the interactions between monitoring and inspection activities. Section 5 outlines the data sources and details the regression models. Section 6 presents and discusses the regression results. Lastly, the final section concludes the study, provides policy implications, and discusses the potential limitations for further research.

2.2 Literature: monitoring and inspection in environmental regulation

Monitoring and inspection are critical components of environmental regulation in general. The economics literature has investigated different aspects of these components, including the effectiveness of monitoring and inspection, the role of regulators, and challenges in the regulation.

The effectiveness of monitoring and inspection in air pollution regulation has been supported by numerous studies, as reviewed in Gray and Shimshack (2011), Shimshack (2014), and Currie and Walker (2019). Specificly, regarding the direct impact of an air pollution monitor itself, Auffhammer, Bento, and Lowe (2009) found that when a monitor's reading exceeded the PM10 national standards in one year, there was a subsequent reduction in pollution levels the following year. The paper also finds that the reduction occurred regardless of whether the area in which the monitor was located was designated as non-attainment or not. For enforcement activities, early research by Gray and Deily (1996) and Gray and Shadbegian (2005) demonstrated that enforcement actions, including inspections, led to enhanced compliance in the steel industry and paper mill industry during the 1980s. Hanna and Oliva (2010) found that EPA air compliance evaluations reduced aggregate Toxic Release Inventory-reported emissions across several manufacturing industries from 1987 to 2001. The inspection activities are also found to have deterrence impact beyond the inspected object itself. For example, Lim (2016) discovered that increasing the average number of inspections at other facilities within the same county by one led to a 2.7% decrease in emissions from that facility. These findings emphasize the importance and impact of both monitoring and inspections activities in air pollution regulation.

When employing monitoring or inspection tools to attain environmental outcomes, the ideal economic model assumes that regulators consistently maximize social welfare. However, recent findings indicate that environmental regulators may not always strictly maximize social welfare, leading to outcomes that deviate from the socially optimal results. This could be attributed to external incentives within the air pollution regulation system. As demonstrated theoretically in Rabassa (2008), when a penalty exists for surpassing the ambient air pollution threshold, local regulators may be incentivized to allocate more regulation resources around the threshold where nonattainment designation occurs. Additionally, political factors could play a role, such as the partisanship of state governors (Bergquist, 2018), the local community's capacity to influence environmental outcomes (Banzhaf, Ma, & Timmins, 2019), or the local regulator's ability to obtain information on pollution sources (Dube & Zipperer, 2015).

Empirically, we see many cases for which the optimal monitoring or inspection is not achieved. For example, Zou (2021) finds that air quality is significantly worse on unmonitored days by combining satellite data with monitor working schedules. Furthermore, Zou's recent paper finds evidence that monitors' sampling rate drops on days when the local government issues pollution alerts, which indicates strategic monitoring behavior (Mu et al., 2021). Also using satellite data, Grainger and Schreiber (2019) find that the ambient pollution monitors are strategically located in cleaner areas within the attainment counties, where there is less regulatory pressure from EPA compared to non-attainment counties.

In conclusion, although monitoring and inspection are effective tools in combating pollution, current literature shows that in practice, factors beyond the pollution itself could influence the allocation of the monitoring or enforcement resources. Building on this literature, this paper will focus on how unexpected changes in the monitoring itself influence the enforcement activities.

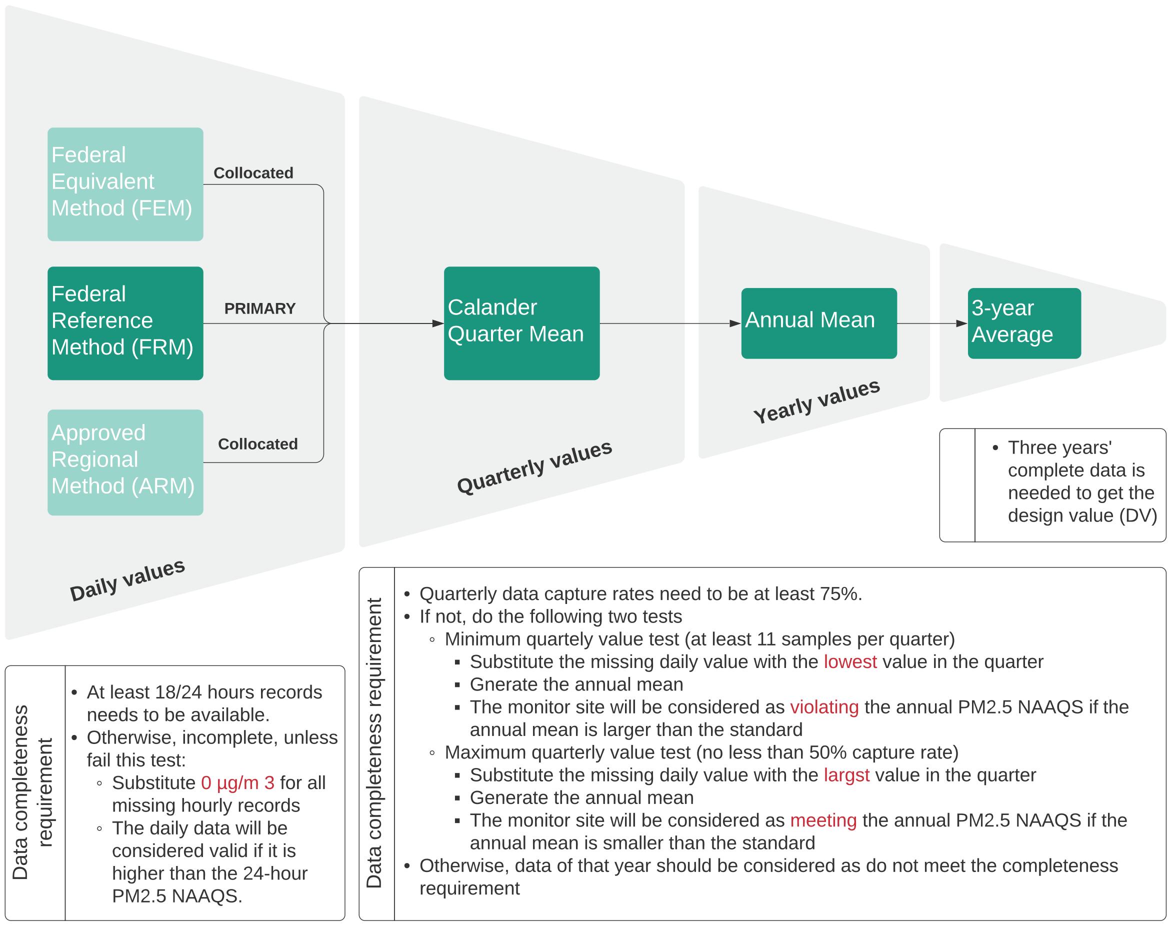


Figure 2.2 Data Completeness Requirement for Annual PM2.5 NAAQS: Getting Design Value (DV) at Monitor Site

Note: The processes in the figure is summarised from the description in government document EPA 40 CFR Part 50, Appendix N to Part 50 - "Interpretation of the National Ambient Air Quality Standards for PM2.5."

2.3 Monitoring and enforcement under the CAA

2.3.1 PM2.5 monitoring and the attainment designation

Under the Clean Air Act (CAA), the EPA is responsible for coordinating the national network of air pollutant monitors and categorizing areas into "attainment" and "non-attainment" groups based on monitor data. The first step in this process involves obtaining qualified data from monitor sites, which must adhere to well-defined data completeness requirements (Figure 2.2). These requirements include daily and quarterly levels to generate a usable annual pollution mean. Moreover, three consecutive years of annual pollution means are needed to generate the three-year average used for the EPA's attainment designation process. Monitors that comply with these requirements are considered "valid monitors," and their data is referred to as "valid data." Using the valid data, the EPA calculates the "design value" for each monitor site and compares it with the NAAQS. If a site has a design value higher than the NAAQS, it is designated as "non-attainment," indicating that pollution levels exceed national standards.[2]

However, not all monitors can provide sufficient valid data to meet data completeness requirements. As seen in Table 2.1, during the period of 2001 to 2018, a large portion of monitors could not provide valid data on a yearly basis. The percentage of monitors unable to meet the standards fluctuates over the years, but typically, around 20% of monitors cannot provide qualified data for the EPA to use.

The existence of a large percentage of incomplete data presents challenges for the EPA in identifying non-attainment areas. Since incomplete data cannot be used to demonstrate compliance with NAAQS, the EPA is required by law to designate monitor sites and related areas with such data as "unclassifiable." This means that the EPA cannot determine, based on available information, whether the area is meeting or not meeting the NAAQS. Sometimes, a large region might be designated as "unclassifiable" due to a lack of sufficient data. For instance, the entire state of Illinois was designated as unclassifiable because there was not enough data meeting the completeness requirement during the relevant period from 2011–2013.

The EPA can assign non-attainment status using limited data if it is believed that the incomplete data suggests a potential violation of NAAQS. For example, in the case of PM2.5 monitors, only 11 days from a quarter (approximately 90 days) are needed for the EPA to designate a violation if the average of the 11 days exceeds the annual standard of 15 $\mu g/m^3$, as the limited data already

[2]Following the designation for each monitor site, the EPA takes further steps to determine the attainment status of corresponding areas. Areas contributing to a monitor site's pollution are identified as "non-attainment." This process involves states initiating the identification of areas that contribute to a specific monitor site's NAAQS violation, with the EPA approving the boundary determinations. Typically, the final non-attainment designation occurs at the county level, although there are instances where only a portion of a county is designated as non-attainment. Also, note not all counties have air pollution monitors within their jurisdictions. In fact, only 1,289 of the 3,144 U.S. counties had monitors for any type of criteria air pollutant between 1990 and 2015 (Hsiang, Oliva, & Walker, 2019). The high cost of monitors is one reason for this sparse allocation: the annual cost of a single PM monitor, for example, can reach $41,000 (Zou, 2021).

Table 2.1 PM2.5 Monitor Completeness

	PM2.5 Monitors			
	Complete	Incomplete	Total	%Incomplete
2002	1,019	267	1,286	21%
2003	908	363	1,271	29%
2004	967	217	1,184	18%
2005	954	257	1,211	21%
2006	930	218	1,148	19%
2007	921	197	1,118	18%
2008	867	235	1,102	21%
2009	962	180	1,142	16%
2010	938	201	1,139	18%
2011	829	221	1,050	21%
2012	834	238	1,072	22%
2013	896	236	1,132	21%
2014	937	305	1,242	25%
2015	1,066	235	1,301	18%
2016	1,090	247	1,337	18%
2017	1,125	263	1,388	19%
2018	1,150	278	1,428	19%

suggests a severe pollution problem. If a monitor collects fewer than 11 samples per quarter, the Clean Air Act also grants the EPA the right to use alternative data. We can find several cases of areas with incomplete data being designated as non-attainment, for example, Lincoln County in Montana could not provide complete data for the period of 2009-2011, but Libby city within that county was still designated as non-attainment in 2012. These cases are dealt with on a case-by-case basis according to EPA's announcement.

In conclusion, monitoring data face challenges in data completeness, leaving some flexibility in the designation of non-attainment status, rather than strictly comparing a monitor's three-year averages with the NAAQS.

2.3.2 After non-attainment designation: enforcement activities

Both states and counties are subject to pressure from the federal government to reduce pollution and adhere to the National Ambient Air Quality Standards (NAAQS) requirements. The federal government utilizes various mechanisms to incentivize compliance with these standards, one of which is the allocation of federal funds. Failure to meet NAAQS requirements can lead to the loss of

Table 2.2 Enforcement Actions by Regulatory Authority Levels

		Inspection			Informal			Formal		
Year	# Facilities	EPA	State	Local	EPA	State	Local	EPA	State	Local
2002	7,697	197	6473	784	21	624	214	77	379	142
2003	7,449	205	7416	1017	32	530	216	91	392	138
2004	7,252	206	8095	1095	31	541	230	88	346	162
2005	7,228	383	6868	1229	38	483	205	127	417	112
2006	6,941	469	7512	1191	29	512	200	90	384	115
2007	6,803	455	7226	1274	29	632	209	84	313	132
2008	6,473	564	6807	1128	29	545	241	65	374	163
2009	6,115	506	6246	1179	41	506	219	87	361	176
2010	6,094	554	5850	1238	26	458	143	78	305	97
2011	5,247	408	4845	1312	30	367	182	62	229	109
2012	5,314	366	5023	1225	22	376	187	44	198	131
2013	5,428	444	4819	1350	25	335	195	54	199	149
2014	5,891	436	4673	1172	23	389	136	39	196	125
2015	5,863	357	4109	1057	26	407	183	50	184	130
2016	5,591	205	3864	1042	25	399	169	36	178	111
2017	5,516	203	3504	849	18	392	136	35	190	87
2018	5,092	198	3154	731	14	376	142	38	173	90

federal funding, like funding for highways, which can significantly impact a state or county's ability to support essential programs and services (Greenstone, List, & Syverson, 2012). The potential loss of federal funds can serve as a powerful motivator for states and counties to invest in pollution reduction measures, enforce environmental regulations, and avoid violation of NAAQS.

If a jurisdiction is designated as non-attainment, these authorities must undertake a series of steps to reduce air pollution, improve air quality, and ultimately comply with the NAAQS. This process begins with the creation of a State Implementation Plan (SIP), which delineates the strategies and measures the state will implement to bring non-attainment areas into compliance. Enforcement activities, including inspections, are essential components of these efforts, ensuring that regulations and emission controls outlined in the SIPs are effectively executed.

On-site inspections play a crucial role in enforcement. There are two types of on-site inspections: regular inspections, usually conducted once every two years, and "for cause" inspections, which are triggered by indicators of noncompliance, such as complaints from nearby residents. If inspections reveal that a facility is violating environmental regulations, additional enforcement actions may be taken. These enforcement measures can range from mild informal sanctions, like warning letters, to severe formal actions, such as fines and lawsuits. In both inspection activities, the majority of enforcement actions are conducted by state and county governments (Table 2.2).

2.4 Theory model

Building on the models developed by Rabassa (2008) and Grainger and Schreiber (2019), we analyze local regulators' behavior by dividing their regulatory actions into two groups: monitoring and inspection. For the purposes of this study, we focus solely on point-source pollution regulation, meaning that we assume local regulatory activities influence air quality through their impact on point-source emissions[3]. First, we focus on a model where there is no ambient air pollution standards. Then, we extend the model to a scenario where regulators face penalties if pollution surpasses national standards.

2.4.1 Optimizing regulator actions without external pollution standards

In this model, we assume that local regulators aim to maximize a utility function that includes health costs of ambient air pollution, welfare benefits from polluting industries, and regulation costs associated with monitoring and inspections.

$$U(M, E) = -h(A) + w(e) - c_1(M) - c_2(I)$$

- e represents the point-source air pollution emission, which we assume responds to regulatory activities, denoted as $e = e(I, M)$. Here, I represents the inspection activities, and M represents the monitoring activities. We assume that more resources allocated to monitoring activities will increase the probability of detecting non-compliance activities, and higher inspection intensity directly increases the cost of emissions. As a result, we have $e_M < 0$ and $e_I < 0$, indicating that emissions decrease as there are more monitoring and inspection

[3]While this assumption does not perfectly reflect reality, as local regulators can also influence mobile source pollution (e.g., adopting stricter emission standards like California's Low Emission Vehicle program), our study focuses on inspection activities targeting facilities alone. As a result, we disregard the emission impact from mobile sources. In the empirical analysis, we account for the impact from mobile sources by incorporating a state-specific yearly time trend. This approach helps alleviate concerns about confounding effects from state-level policy changes on mobile emissions.

activities.

We also assume diminishing returns to both monitoring and inspection, with $e_{II} > 0$ and $e_{MM} > 0$. This means that the additional emissions reduction achieved by each subsequent unit of inspcetion or monitoring effort decreases at a diminishing rate.

Furthermore, we assume that the marginal emissions reduction effect of monitoring will not be influenced by inspection activities, and vice versa. This is denoted by $e_{IM} = 0$ and $e_{MI} = 0$. This is a strong assumption, as we are assuming that there are no combined effects of monitoring and inspection that reinforce each other in reducing emissions.

- $-h(A)$ represents the health costs associated with ambient air pollution, where $A = \alpha \cdot e + \epsilon$ denotes the ambient air pollution levels influenced by both point-source emissions (e) and weather shocks (ϵ). The weather shocks follow a normal distribution, $\epsilon \sim N(0, \delta_\epsilon)$. Furthermore, we assume that $h' > 0$ and $h'' > 0$, indicating that the health costs associated with pollution exhibit increasing marginal costs. In other words, as pollution levels increase, the additional health costs incurred for each additional unit of pollution also increase.
- $w(e)$ represents the welfare benefits associated with point-source emissions e. The first-order derivative, $w' > 0$, signifies that there are positive local economic opportunities generated by polluting industries, such as job creation and economic growth. The second-order derivative, $w'' < 0$, captures the concept of diminishing returns in having polluting industries. This means that as the level of point-source emissions increases, the additional welfare benefits gained from each subsequent unit of emissions will decrease. In other words, the positive economic impact of polluting industries diminishes as emissions grow.
- $c_1(M)$ represents the cost associated with monitoring activities. The first-order derivative, $c_1' > 0$, indicates that the cost of monitoring increases as the level of monitoring efforts increases. The second-order derivative, $c_1'' > 0$, signifies that the cost of monitoring activities experiences increasing marginal costs, meaning that the additional cost incurred for each subsequent unit of monitoring effort grows at an increasing rate.
- Similarly, $c_2(I)$ represents the cost associated with inspection activities. The first-order

derivative, $c_2' > 0$, implies that the cost of inspection rises as the level of inspection efforts increases. The second-order derivative, $c_2'' > 0$, conveys that the cost of inspection also experiences increasing marginal costs, with the additional cost for each subsequent unit of inspection effort growing at an increasing rate.

With the assumptions given above, we can rewrite the local regulator's maximization problem as:

$$\begin{aligned}
&\max_{I,M} \mathbb{E}\{-h(\alpha\cdot\ e+\epsilon)+w[e(M,I)]-c_1(M)-c_2(I)\}\\
&=\max_{I,M}\int_{-\infty}^{\infty}\{-h(\alpha\cdot\ e+\epsilon)+w[e(M,I)]-c_1(M)-c_2(I)\}f(\epsilon)d\epsilon\\
&=-\mathbb{E}[h(\alpha\cdot\ e+\epsilon)]+w[e(M,I)]-c_1(M)-c_2(I)
\end{aligned}$$

Thus, the first order condition will be:

$$I: \qquad -\alpha\,\mathbb{E}[h']e_I + w'\cdot e_I - c_2' = 0 \tag{2.1}$$

$$M: \qquad -\alpha\,\mathbb{E}[h']e_M + w'\cdot e_M - c_1' = 0 \tag{2.2}$$

Note that equations (2.1) and (2.2) can also be rewritten as $-\alpha\,\mathbb{E}[h']e_I + w'\cdot e_I = c_2'$ and $-\alpha\,\mathbb{E}[h']e_M + w'\cdot e_M = c_1'$. The left-hand side of these equations represents the marginal benefits of an additional unit of inspection (or monitoring) activities, while the right-hand side signifies the marginal cost of these activities.

Together, these equations yield the optimal monitoring and inspection levels (I^*, M^*). However, our primary interest lies in the relationship between I and M. To illustrate this, we take the total differential of equation (2.1):

$$\begin{aligned}
&\Big[-\alpha^2\,\mathbb{E}(h'')e_I^2-\alpha\,\mathbb{E}[h']e_{II}+w''e_I^2+w'e_{II}-c_2''\Big]dI+\\
&\qquad\Big[-\alpha^2\,\mathbb{E}(h'')e_Ie_M-\alpha\,\mathbb{E}[h']e_{IM}+w''e_Ie_M+w'e_{IM}\Big]dM=0
\end{aligned}$$

Substituting with

$$-\alpha\,\mathbb{E}[h']+w'=\frac{c_2'}{e_I}$$

and

$$-\alpha\,\mathbb{E}[h']+w'=\frac{c_1'}{e_M}$$

which are derived from the first-order conditions (F.O.C), we have:

$$\frac{dI}{dM} = -\frac{-\alpha^2\,\mathbb{E}(h'')e_I e_M + w'' e_I e_M}{-\alpha^2\,\mathbb{E}(h'')e_I^2 + \frac{c_2' e_{II}}{e_I} + w'' e_I^2 - c_2''} < 0$$

This result implies that monitoring and inspection are substitutes for each other, meaning that when one increases, the other decreases, and vice versa.

2.4.2 Regulators' actions with external pollution standards

Now we extend the model to include a penalty Z for exceeding the ambient air pollution threshold $\bar{A}$:

$$U(M,I) = \begin{cases} -h(A) + w(e) - c_1(M) - c_2(I), & A_M < \bar{A} \\ -h(A) + w(e) - c_1(M) - c_2(I) - Z, & A_M \geq \bar{A} \end{cases}$$

where A_M is the monitored ambient air pollution level, $A_M = M \cdot e + \epsilon$. Therefore, the local regulators' maximization problem becomes:

$$\begin{aligned} \max_{E,M} \int_{-\infty}^{\bar{A}-Me} & \{-h(\alpha\cdot\ e+\epsilon) + w[e(I,E)] - c_1(M) - c_2(I)\} f(\epsilon) d\epsilon + \\ & \int_{\bar{A}-Me}^{\infty} \{-h(\alpha\cdot\ e+\epsilon) + w[e(M,I)] - c_1(M) - c_2(I) - Z\} f(\epsilon) d\epsilon \\ & \quad = \mathbb{E}[-h(\alpha\cdot\ e+\epsilon)] + w[e(M,I)] - c_1(M) - c_2(I) - [1 - F(\bar{A} - Me)]Z \end{aligned}$$

And the first order condition will be:

$$I: \qquad -\alpha\,\mathbb{E}[h']e_I + w'\cdot e_I - c_2' - f(\bar{A} - Me)Me_I Z = 0 \tag{2.3}$$

$$M: \qquad -\alpha\,\mathbb{E}[h']e_M + w'\cdot e_M - c_1' - f(\bar{A} - Me)(e + Me_M)Z = 0 \tag{2.4}$$

Similarly, the equation (2.3) and (2.4) could be rewrite to reflect the relations between the marginal benefits and marginal cost of inspection (or monitoring) activities:

$$I: \qquad -\alpha\,\mathbb{E}[h']e_I + w'\cdot e_I - f(\bar{A} - Me)Me_I Z = c_2' \tag{2.5}$$

$$M: \qquad -\alpha\,\mathbb{E}[h']e_M + w' \cdot e_M = c_1' + f(\bar{A} - Me)(e + Me_M)Z \tag{2.6}$$

Now, we can observe that in comparison to the case without the penalty term, the marginal benefits of inspection activities include an additional term, $-f(\bar{A} - Me)Me_I Z$. Given that $f(\bar{A} - Me)$, M, and Z are always positive, and e_I is always negative, we know that $-f(\bar{A} - Me)Me_I Z$ is always positive. Thus, this new marginal benefits term will consistently be larger than the marginal benefits in the case without the penalty term. This conclusion aligns with our intuition: when a penalty is imposed for violating national standards, inspection activities yield additional benefits by reducing the likelihood of violations. Meanwhile, the marginal cost of monitoring increases, as additional monitoring activities heighten the chances of discovering a violation of the national standards[4].

Again, we do the total derivative and get the relation between I and M:

$$\frac{dI}{dM} = -\frac{-\alpha^2\,\mathbb{E}(h'')e_I e_M + w'' e_I e_M + f' e M e_I Z + f' M^2 e_I e_M Z - f e_I Z}{-\alpha^2\,\mathbb{E}(h'')e_I^2 + \frac{c_2' e_{II}}{e_I} - w' e_{II} + w'' e_I^2 - c_2'' + f' M^2 e_I^2 Z} \tag{2.7}$$

For the convenience of analysis, we define:

$$K \equiv -\alpha^2\,\mathbb{E}(h'')e_I e_M + w'' e_I e_M \tag{2.8}$$

$$B \equiv f' e M e_I + f' M^2 e_I e_M - f e_I \tag{2.9}$$

$$C \equiv -\alpha^2\,\mathbb{E}(h'')e_I^2 + \frac{c_2' e_{II}}{e_I} - w' e_{II} + w'' e_I^2 - c_2'' \tag{2.10}$$

$$D \equiv f' M^2 e_I^2 \tag{2.11}$$

[4]In the term $f(\bar{A} - Me)(e + Me_M)Z$, we know that $f(\bar{A} - Me)$ is always positive, as it is the density function, and Z is always positive since it represents the penalty term. The sign of $(e + Me_M)$ is undetermined, but we assume it to be positive. This term is derived from $\frac{dMe}{M}$, which indicates the marginal impact of monitoring activities on the recorded ambient air pollution level. By assuming this to be positive, we posit that increased monitoring activities will consistently lead to a higher recorded ambient air pollution level. This assumption is reasonable, as literature suggests that ozone monitors are typically placed in relatively cleaner regions of a county, and PM2.5 pollution levels are higher on days when PM2.5 monitors are not functioning (Grainger & Schreiber, 2019; Grainger et al., 2018; Zou, 2021).

Here, both K and C are always negative due to the assumptions made earlier. The sign of B should consistently be positive, as it is derived from $\frac{d-f(\bar{A}-Me)Me_I}{dM}$, which represents the marginal impact of monitoring on the additional benefit term in the expression for the marginal benefits of inspection. Essentially, this indicates that as monitoring activities increase, the extra benefits gained from more inspection and avoidance of national standards violations also grow.

The sign of D is determined by the sign of f', which in turn depends on the magnitude of $\bar{A} - Me$. When the expected recorded pollution level (Me) exceeds the national standards $\bar{A}$, we have $\bar{A} - Me < 0$, resulting in $f' > 0$ and $D > 0$. Conversely, when the expected recorded pollution level (Me) is lower than the national standards $\bar{A}$, we have $\bar{A} - Me > 0$, leading to $f' < 0$ and $D < 0$.

When we rewrite equation (2.7) as following, we could see how the magnitude of Z and the sign of f' together impact the sign of $\frac{dE}{dM}$.

$$\frac{dI}{dM} = -\frac{K + BZ}{C + DZ} = (K + BZ)(-\frac{1}{C} - \frac{1}{D}Z) \tag{2.12}$$

Case I: In cleaner region when the expected pollution level is lower than the national standards ($\bar{A} - Me > 0$, $f' < 0$, and $D < 0$).

When $Z > -\frac{K}{B}$, we observe that $\frac{dI}{dM} > 0$ [5]. Unlike the scenario without a penalty, inspection and monitoring are no longer substitutes for each other; instead, they move in the same direction. This implies that the penalty for violating national standards is considerably large, prompting regulators to potentially decrease their inspection activities when monitoring intensity is lower (indicating a lower risk of violating national standards). Conversely, they may increase inspection activities when monitoring intensity is higher (signifying a higher risk of violating national standards).

Case II: In dirtier region when the expected pollution level is higher than the national standards ($\bar{A} - Me < 0$, $f' > 0$, and $D > 0$).

In this case, the quadratic equation (2.12) equating to zero yields two solutions, $-\frac{K}{B}$ and $-\frac{D}{C}$, both of which are greater than zero. We cannot determine which solution is larger, but we can

[5] The other solution point for the quadratic equation (2.12) equal to zero is $Z = -\frac{D}{C}$, which is smaller than zero, thus exclude from the discussion.

Table 2.3 Theory model's prediction on dE/dM

dE/dM	Clean	Dirty
Penalty is small	$-$	$-$
Penalty is medium	$+$	$+$
Penalty is large	$+$	$-$

Note: For both the clean and dirty region, the penalty size definition is for the region itself. We are not intend to do cross region comparison here.

denote them as Z_1^* and Z_2^*, where $Z_1^* < Z_2^*$. Here, if the actual penalty Z is either very large ($Z > Z_2^*$) or very small ($Z < Z_1^*$), inspection and monitoring will still function as substitutes for each other ($\frac{dE}{dM} < 0$); however, when the actual penalty falls between Z_1^* and Z_2^*, they move in the same direction ($\frac{dE}{dM} > 0$). Intuitively, this suggests that for areas that are heavily polluted and expected to violate national pollution standards, thereby almost certainly incurring penalties, the relationship between inspection and monitoring will largely resemble the case without a penalty. The exception occurs when the penalty falls within the range $Z_1^* < Z < Z_2^*$, where reduced monitoring could potentially bring the region into compliance with national standards, making it cost-effective to decrease inspection; or when increased monitoring could potentially cause the region to violate national standards, making it cost-effective to increase inspection to avoid incurring penalties.

In conclusion, when penalties are associated with violating national standards, inspection and monitoring may no longer function as substitutes for each other. The Table 2.3 summarizes their relationship.

2.5 Data and method

2.5.1 Data

We obtain data from several databases and merge them to obtain facility-level data. We provide details on the data used as follows:

Facility-level inspection data from the EPA's Integrated Compliance Information System for Air (ICIS-Air): ICIS-Air provides compliance and inspection data on stationary sources of air pollution, reported at the facility level. It includes the facility's identification number, geographic

Table 2.4 Summary Statistics at the Facility Level (2002-2018)

Variable	Observation	Mean	Std. Dev	Min.	Max.
Inspection					
From EPA	105,994	0.0581	0.4152	0	34
From state	105,994	0.9103	2.5085	0	166
From local	105,994	0.1781	1.1463	0	54
Incomplete	105,994	0.3818	0.4858	0	1
Non-attainment	105,994	0.2411	0.4277	0	1
90%NAAQS	105,994	0.2600	0.4386	0	1
VOC emission	105,994	0.0017	0.0093	0	0.3763
PM emission	105,994	0.0001	0.0064	0	1.3838
Pb emission	105,994	0.0005	0.0113	0	2.7426
Other emission	105,994	0.0001	0.0015	0	0.1751
PM2.5(census tract)	664,012	10.0454	2.9803	0.3347	23

locations, and the number of inspections, formal and informal enforcement activities received by different regulatory agencies. The annual enforcement activities, summarized by agency type, are provided in the previous Table 2.2. As mentioned earlier, the majority of enforcement actions are carried out by state agencies, followed by local agencies, with federal agencies initiating only a small portion of enforcement activities. This distribution of inspection activities by agency type can also be seen in the facility panel's summary statistics in Table 2.4. As shown, the average number of inspections for a facility received by the EPA is 0.06 times per year, while it is 0.91 times by the state per year and 0.18 times by local regulators per year.

PM2.5 pollution monitor data from the EPA's Air Quality System (AQS): The AQS contains comprehensive pollution records for each monitor, along with their location information. It also includes a variable "complete" for each monitor, which, as defined in AQS's description file, indicates whether the regulatory data completeness criteria for valid summary data have been met by the monitor for the year. The occurrences of incomplete data for each year at the monitor level are detailed in the previous Table 2.1. Moreover, the study identifies "above average" monitors by generating a "90%$NAAQS$" variable, defined as a monitor reading higher than 90% of the NAAQS standard for that year. The 90% threshold is chosen arbitrarily, considering that having a reading at 90% of the NAAQS might potentially increase the risk of recording pollution levels that violate the NAAQS[6].

We aggregate the monitor-level data at the county level and assign the variable $Incomplete_{ct} = 1$ when there is at least one incomplete PM2.5 monitor in county c and time t. The same definition also applies to $90\%NAAQS_{ct}$, where we define $90\%NAAQS_{ct} = 1$ when one or more monitors in county c surpass the NAAQS at time t. As summarized in Table 2.4, 38% of the facility-year observations in the panel are located in counties with incomplete monitors, and 26% of the observations are situated in counties where monitors have a reading higher than 90% of the year's NAAQS.

County-level attainment status derived from the EPA's Green Book: The EPA's Green Book provides information on areas within the United States that are not in compliance with the NAAQS. We utilize the county-level designated non-attainment status. It is important to note that some counties are only partially designated as non-attainment areas; however, we treat these partially designated counties the same as those that are fully designated. As displayed in Table 2.4, 24% of the facility-year observations within the panel are situated in counties that are either fully or partially designated as non-attainment areas.

Facility-level pollution emissions from EPA's Toxic Release Inventory (TRI) database: To account for emissions from each facility, this study utilizes the TRI's basic data files, which cover the self-reported annual emissions of each stationary source by mass and chemical. The TRI includes hundreds of different pollutants, which can be converted into the precise quantities of VOCs, PM, Pb, and other types of pollutant emissions using the crosswalk introduced in Greenstone (2003) and refined in Gibson (2019). These emission data serve as control variables in the subsequent analysis of the facility-year panel. One remaining question is whether these variables are endogenous, meaning that not only do the facilities' TRI emissions affect the inspection activities they receive, but the inspection activities they receive also affect the TRI emissions. To address the concern regarding endogenous control variables, I report regression results with and without the TRI

[6]Monitors with readings that are lower would be considered "safer" in terms of violating NAAQS; for example, Michigan may consider "eliminating" a monitor if its recorded pollution has consistently been less than $85\% \cdot NAAQS$ for several years.

emission control variables in my analysis.

The summary statistics for facility-level emissions can be found in Table 2.4. It is important to note that the unit of the emission data is 1,000 tons. This value was set explicitly to prevent the coefficients displayed in the following regressions from appearing too small.

2.5.2 Regression Model

This study employs a regression model to investigate the relationship between the incomplete status of monitors and inspection frequency. In line with our theoretical model, we hypothesize that a county with at least one monitor unable to meet the completeness requirement ($Incomplete_{ct} = 1$) will experience a decrease in monitoring activities. We utilize the following regression model to assess whether this reduction in monitoring is associated with a significant change in inspection frequency.

$$Inspection_{ict} = \beta_1 Incomplete_{ct} + \Gamma X_{ict} + \alpha_i + \lambda_t + \lambda_t \cdot \theta_{state} + \varepsilon_{it} \tag{2.13}$$

In this context, $Inspection_{ict}$ denotes the number of inspections facility i in county c receives in year t. $Incomplete_{ct}$ is a dummy variable indicating the intensity of monitoring activities, where $Incomplete_{ct} = 1$ if one or more PM2.5 monitors in county c cannot provide complete data in year t, as described in section 2.5.1. The core coefficient in this analysis is β_1. If we expect monitoring activities and inspection to serve as substitutes for each other, we would anticipate a positive β_1. Conversely, if we believe that monitoring activities and inspection move in the same direction, we would expect a negative β_1.

The variable X_{ict} represents facility i's VOC, PM, Pb, and other air pollution emissions recorded in the TRI database. The equation also includes facility fixed effects (α_i), year fixed effects (λ_t), and year-state fixed effects ($\lambda_t \cdot \theta_{state}$). By incorporating these three types of fixed effects, we effectively control for the influence of any specific year, state, or state-level policy changes, such as change in the state level vehicle emission regulations.

In the theoretical model with penalties, we also discuss the varying impacts of being in a polluted region versus being in a cleaner region. To account for this analysis, we extend model 2.13

to include indicators of whether a region is clean or not. We utilize two different indicators: the first is the non-attainment status at the county level ($Non\text{-}attainment_{ct} = 1$), and the second is having a high-reading monitor at the county level ($90\%NAAQS_{ct} = 1$)[7]. As shown in Table 2.4, there are more observations with high-reading monitors than those located in non-attainment counties. Furthermore, as we will see later, the two definitions of "dirty" yield coefficients with the same direction but different magnitudes. In both equations, the coefficient β_1 captures the relationship between monitoring and inspection for cleaner regions, while the combined coefficients $\beta_1 + \beta_3$ represent the relationship between monitoring and inspection for polluted regions. To articulate this more clearly, the term $-\beta_1$ corresponds to the signs found in the first column labeled as 'clean' in Table 2.3. Similarly, the term $-\beta_1 - \beta_3$ corresponds to the signs found in the second column, labeled as 'dirty', in Table 2.3.

$$\begin{aligned} Inspection_{ict} = {} & \beta_1 Incomplete_{ct} + \beta_2 Non\text{-}attainment_{ct} \\ & + \beta_3 Incomplete_{ct} \cdot Non\text{-}attainment_{ct} \\ & + \Gamma X_{ict} + \alpha_i + \lambda_t + \lambda_t \cdot \theta_{state} + \varepsilon_{it} \end{aligned} \tag{2.14}$$

$$\begin{aligned} Inspection_{ict} = {} & \beta_1 Incomplete_{ct} + \beta_2 \cdot 90\%NAAQS_{ct} \\ & + \beta_3 Incomplete_{ct} \cdot 90\%NAAQS_{ct} \\ & + \Gamma X_{ict} + \alpha_i + \lambda_t + \lambda_t \cdot \theta_{state} + \varepsilon_{it} \end{aligned} \tag{2.15}$$

It is important to note that we cannot claim causal relationships from the above regressions, mainly because $Incomplete_{ct}$ may not be random. In fact, in our theoretical analysis, we assume that both monitoring intensity and inspection are variables subject to an agency's choices.

2.6 Results

Tables 2.6, 2.7, and 2.8 present the results obtained from applying the aforementioned regression models to the inspection activities of EPA, state, and local agencies. For improved readability, we

[7]Please note that it is possible for a monitor to record pollution levels exceeding the NAAQS while still being located in a county with attainment status. This is because the EPA requires three consecutive years of data to designate a county as non-attainment. Therefore, to evaluate whether a region is clean or dirty, we consider not only the attainment/non-attainment status but also the current year's monitor readings.

Table 2.5 Summary: Inspections vs. Incomplete

	Basic model in Eq.(2.13)	Non-attainment		90%NAAQS	
		Clean	Dirty	Clean	Dirty
EPA	0.00244	0.00376	-0.00137	0.00115	0.00694
	(0.00338)	(0.00358)	(0.00707)	(0.00357)	(.00614)
State	-0.0497***	-0.0498*	-0.0493**	-0.0514***	-0.0447**
	(0.0175)	(0.0214)	(0.0231)	(0.0193)	(0.0213)
Local	-0.0188***	-0.0130*	-0.0359*	-0.00720	-0.0689***
	(0.00659)	(0.00619)	(0.0195)	(0.00556)	(0.0184)

Note: all standard errors are clustered at the facility level. $* \, p < 0.10$, $* \, p < 0.05$, $* * * \, p < 0.01$. The content in the first row (EPA) is obtained from Table 2.6, the content in the second row (State) is obtained from Table 2.7, and the content in the third row (Local) is obtained from Table2.8.

further refine the coefficients in Table 2.5.

Prior to delving into the analysis of core coefficients, it is crucial to address two points. Firstly, across all three regression tables, the inclusion of facility-level emissions as a control variable has little impact on the coefficients, thus alleviating concerns about endogenous control variables. Secondly, the total number of observations in the regression is marginally lower than the observation figures in Table 2.4, a consequence of excluding facilities with only one observation in our fixed effects model at the facility level.

Turning attention to Table 2.5, the first column shows the β_1 estimate derived from model Eq.(2.13). Under "non-attainment" in the second and third columns, the β_1 and $\beta_1 + \beta_3$ estimates from model Eq.(2.14) are displayed, representing the relationships between inspection and monitoring in attainment and non-attainment regions, respectively. Under the "90%NAAQS", the fourth and fifth columns present the β_1 and $\beta_1 + \beta_3$ estimates obtained from model Eq.(2.15). These two estimates correspond to the correlations between inspection and monitoring in regions where monitor readings exceed 90% of the NAAQS and those without such readings. The complete regression results can be found in Tables 2.6, 2.7, and 2.8.

In the first row of Table 2.5, we can see that EPA inspection activities exhibit no change when a county experiences an instance of monitors becoming incomplete, regardless of whether it is a clean or dirty region. These findings do not precisely align with our initial predictions based on the theoretical model. As a federal regulator, we might anticipate that the EPA would not need to

Table 2.6 Inspections vs. Incomplete (2002-2018; EPA inspections)

	(1) EPA	(2) EPA	(3) EPA	(4) EPA	(5) EPA	(6) EPA
Incomplete	0.00241	0.00371	0.00112	0.00244	0.00376	0.00115
	(0.00338)	(0.00358)	(0.00357)	(0.00338)	(0.00358)	(0.00357)
Non-attainment		0.0202^{*}			0.0202^{*}	
		(0.00922)			(0.00922)	
90%NAAQS			0.00302			0.00301
			(0.00508)			(0.00508)
Incomplete · Non-attainment		-0.00507			-0.00513	
		(0.00756)			(0.00757)	
Incomplete · 90%NAAQS			0.00579			0.00579
			(0.00627)			(0.00628)
VOC emission				0.151	0.155	0.154
				(0.766)	(0.765)	(0.766)
PM emission				-0.0677	-0.0678	-0.0672
				(0.0440)	(0.0444)	(0.0439)
Pb emission				-0.0300	-0.0274	-0.0302
				(0.0602)	(0.0588)	(0.0608)
Other emission				-1.590	-1.608	-1.585
				(3.229)	(3.230)	(3.230)
Facility FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Year × State FE	Y	Y	Y	Y	Y	Y
Obs	105202	105202	105202	105202	105202	105202
R-square	0.354	0.354	0.354	0.354	0.354	0.354

Note: all standard errors are clustered at the facility level. $* \ p < 0.10$, $* \ p < 0.05$, $* * * \ p < 0.01$.

take external penalties into account. Consequently, their optimal choice should adhere to the model without penalties, in which inspection acts as a substitute for monitoring activities, suggesting a positive coefficient. While the obtained coefficient of 0.00244 is indeed positive, it is very close to zero, and the large standard error of 0.00338 leads us to interpret this as indicating no significant impact. However, it is important to consider that the EPA's involvement in inspections is relatively limited. Federal regulators typically intervene only in serious cases or situations with ambiguity, such as those near state borders, necessitating their involvement. Therefore, it is not unreasonable to observe these no-impact results, as EPA activities may not be sensitive to events like a monitor's incomplete status, which occurs in approximately 20% of all monitoring systems. In light of these contextual factors, the current findings can still be deemed consistent with our expectations.

In the second row of Table 2.5, it is clear that state-level inspection activities decrease significantly when monitors become incomplete. For the basic model, which does not distinguish between clean and dirty regions, we observe a statistically significant coefficient of -0.0497 at the 1% level. This coefficient implies that the number of inspections a facility receives from the state agency is reduced by approximately 0.0497 times per year. In comparison to the average number of inspections received by a facility from the state, this represents a 5.46% (=0.0497/0.910) reduction in inspection frequency.

Moreover, when we expand the model to account for heterogeneity between clean and dirty regions, the coefficient's magnitude remains largely unchanged, ranging from -0.0447 to -0.0514 across the four columns, and maintaining high significance. In line with our theoretical model, one possible explanation for this pattern is that state regulators face medium-level penalties, leading them to opt for fewer inspection activities in both clean and dirty regions when monitors are incomplete and the risk of violating NAAQS is low. Beyond the theoretical model, alternative explanations for this reduction may exist. For instance, limited data could hinder local communities' access to air quality information and decrease the likelihood of them filing complaints with regulators. This, in turn, could result in fewer "for cause" inspections, which are ultimately reflected in our data.

In the third row of Table 2.5, the results pertaining to local regulators' inspection activities are displayed. The table also indicates that inspections significantly decrease when monitors become incomplete. In the basic model, which does not distinguish between clean and dirty regions, the estimated coefficient is -0.0188 and is significant at the 1% level. This coefficient implies that the number of inspections a facility receives from the local agency will be reduced by approximately 0.0188 times per year. Compared to the average number of inspections that a facility would receive from the local agency, this amounts to a 10.56% (=0.0188/0.178) reduction in inspection

Table 2.7 Inspections vs. Incomplete (2002-2018; State inspections)

	(1) State	(2) State	(3) State	(4) State	(5) State	(6) State
Incomplete	-0.0496***	-0.0498*	-0.0512***	-0.0497***	-0.0498*	-0.0514***
	(0.0174)	(0.0213)	(0.0193)	(0.0175)	(0.0214)	(0.0193)
Non-attainment		0.0295			0.0297	
		(0.0295)			(0.0295)	
90%NAAQS			0.0104			0.0106
			(0.0310)			(0.0310)
Incomplete · Non-attainment		0.000743			0.000500	
		(0.0301)			(0.0302)	
Incomplete · 90%NAAQS			0.00644			0.00683
			(0.0245)			(0.0245)
VOC emission				3.245	3.250	3.250
				(3.178)	(3.178)	(3.176)
PM emission				1.922	1.920	1.924
				(1.915)	(1.913)	(1.914)
Pb emission				-0.122	-0.117	-0.121
				(0.481)	(0.480)	(0.481)
Other emission				5.668	5.651	5.678
				(23.28)	(23.29)	(23.28)
Facility FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Year × State FE	Y	Y	Y	Y	Y	Y
Obs	105202	105202	105202	105202	105202	105202
R-square	0.565	0.565	0.565	0.565	0.565	0.565

Note: all standard errors are clustered at the facility level. $* \, p < 0.10$, $* \, p < 0.05$, $* * * \, p < 0.01$.

frequency. As a percentage reduction, this change is more substantial than what we observed in state inspections.

However, when examining the difference between clean and dirty regions, we observe a considerable disparity. Firstly, we find that the reduction in dirty regions is much greater than the reduction in clean regions. In fact, when we define clean/dirty using the variable $90\%NAAQS$, we discover that a monitor's incomplete status has no impact in clean regions but causes a substantially

Table 2.8 Inspections vs. Incomplete (2002-2018; Local inspections)

	(1)	(2)	(3)	(4)	(5)	(6)
	local	local	local	local	local	local
Incomplete	-0.0188***	-0.0130*	-0.00717	-0.0188***	-0.0130*	-0.00720
	(0.00659)	(0.00619)	(0.00555)	(0.00659)	(0.00619)	(0.00556)
Non-attainment		0.0311*			0.0312*	
		(0.0161)			(0.0161)	
90%NAAQS			0.0464***			0.0464***
			(0.0152)			(0.0152)
Incomplete · Non-attainment		-0.0229			-0.0230	
		(0.0208)			(0.0208)	
Incomplete · 90%NAAQS			-0.0619***			-0.0617***
			(0.0172)			(0.0172)
VOC emission				1.851	1.858	1.843
				(1.191)	(1.195)	(1.187)
PM emission				-0.0176	-0.0137	-0.0103
				(0.0334)	(0.0341)	(0.0347)
Pb emission				-0.106	-0.102	-0.0957
				(0.163)	(0.162)	(0.161)
Other emission				1.642	1.598	1.620
				(1.470)	(1.468)	(1.476)
Facility FE	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y
Year × State FE	Y	Y	Y	Y	Y	Y
Obs	105202	105202	105202	105202	105202	105202
R-square	0.691	0.691	0.691	0.691	0.691	0.691

Note: all standard errors are clustered at the facility level. $* \, p < 0.10$, $* \, p < 0.05$, $*** \, p < 0.01$.

larger impact in dirty regions. The impact on dirty regions is 0.0689 times per year, representing a 39% reduction compared to the average level.

To explain this using the theoretical model developed earlier, we might consider that for local regulators in areas with low monitor readings (clean), the absence of data in monitors does not change the status quo. As a result, maintaining the inspection level as normal is the regulator's best choice. However, when the area is dirty with a high potential of violating national standards and incurring relatively large penalties, the potential benefits of having an incomplete monitor increase. Especially, when the penalty involved is not too large that regulator wants to avoid violations at all cost, that is, when the penalty is "medium," reducing inspection at the same time becomes a reasonable choice. Consequently, regulators are not incentivized to bear the cost of imposing inspections and ultimately reduce the inspection frequency we observed in data. This corresponds to the positive sign (+) in Table 2.3, associated with the medium penalty and dirty region.

2.7 Conclusion

In this paper, the regulation efforts in air pollution control are categorized into two distinct activities: monitoring and inspection. A theoretical model is developed to analyze the relationship between these activities. It is assumed that both activities can reduce air pollution; however, only monitoring affects the recorded ambient pollution levels which can trigger penalties associated with non-compliance to national standards.

We find that the relationship between monitoring and inspection is not always a simple substitution. In relatively clean regions with substantial penalties, a decrease in monitoring can significantly reduce the marginal benefits of inspection. This may lead regulators to potentially reduce their inspection activities. On the other hand, in dirtier regions, inspection tends to move in the same direction as monitoring, provided that penalties fall within a specific range. This range is defined as being neither too small for regulators to ignore nor too large for them to allocate all resources to avoid non-compliance with national standards. In such cases, when monitoring intensity decreases, inspection intensity also diminishes. This study highlights the nuanced relationship between monitoring and inspection activities in the context of air pollution control.

We also provide empirical evidence by examining the decrease in monitoring activity that results from incomplete monitors being unable to provide quality data for use by the EPA in designating non-attainment status. Our empirical findings partially support the theoretical predictions derived from the model. We observe that the federal regulator, EPA's inspection activities, show no correlation with the reduction in local monitoring intensity. For state regulators, we find a consistent reduction in inspection for both clean and dirty regions when monitors become incomplete. In the case of local regulators, our results indicate that inspection in cleaner regions remains unaffected by changes in monitoring, while inspection in dirtier regions is significantly reduced by a decrease in monitoring intensity.

These findings further emphasize the complex relationship between monitoring and inspection activities, shedding light on how regulatory bodies at different levels may respond to changes in monitoring intensity. Also, this paper has clear policy implications. Firstly, it highlights the importance of improving the monitoring system to reduce the number of incomplete monitors. A robust monitoring system would ensure accurate data collection and facilitate appropriate regulatory responses to pollution levels. Secondly, for dirty regions with incomplete monitors, our findings suggest that the inspection is reduced. However, we believe the opposite is needed to help the regions to improve air quality. This indicates that more resources should be allocated to address this issue and effectively manage air pollution in these areas. Lastly, it is worth noting that the model used in this study may be an oversimplification, as it assigns a fixed penalty for violating NAAQS. Additionally, it assumes that monitoring and inspection do not amplify each other's effects ($e_{ME} = 0$ and $e_{EM} = 0$). These assumptions could potentially limit the accuracy and applicability of the model. Future research could explore more nuanced models and expand upon these findings to better inform air pollution control policies and strategies.

CHAPTER 3

TOO SMALL TO BE REGULATED? AN EMPIRICAL STUDY ON COLORADO'S EXEMPTION POLICY

3.1 Introduction

Conventional economic theory suggests that when there exist mispriced externalities, the market system alone cannot achieve optimal resource allocation. This forms the fundamental justification for policy interventions: through taxes, cap-and-trade systems, or technology standards, policymakers design regulations to ensure that market participants internalize externalities, thereby achieving a more efficient resource allocation. However, real-world regulations are rarely perfect. They typically apply to a limited region and a specific group of participants, which creates incentives for the regulated parties to evade compliance, ultimately calling into question the net impact of these regulations.

Colorado's "Permit Section (PS) Memo-10-01" is a good example of incomplete regulation with a threshold. Starting at the beginning of 2010, Colorado's environmental regulation agency release a new guidance, PS Memo-10-01, for the air permit issuance process. This memo relaxed regulations, allowing smaller emitters producing less than 40 tons of Nitrogen Oxides (NOx) per year to bypass demonstrating compliance with national air quality standards. This practice endured for approximately a decade until insiders exposed the issue publicly in 2020.

What impact might such policy practices have? Numerous studies across various domains of economic research have evidenced manipulation at regulatory boundaries. This includes individual and corporate behavior at tax thresholds (Harju, Matikka, & Rauhanen, 2019; Liu, Lockwood, Almunia, & Tam, 2021), as well as pollution substitution across jurisdictional borders or regulatory threshold (Cai, Chen, & Gong, 2016; Chen, Chen, & Liu, 2021; Rijal & Khanna, 2020). In the context of incomplete environmental regulations with a defined threshold, how do emitters in Colorado react? This study aims to harness the national emission inventory database to test a specific hypothesis: Does PS Memo 10-01 incite smaller facilities to alter their emission behaviors? In response to this question, my findings do not support the existence of pollution manipulation at the 40 tons cutoff. I also fail to observe strategic regulatory evasion behaviors, such as an increase in pollution from smaller emitters or an uptick in the establishment of smaller emitters. I do, however, recognize the need for additional studies in the future to further explore this topic.

In the following sections, I will provide a literature review of related studies in Section 3.2, clarify the policy background in Section 3.3, conduct a concise theoretical analyses in Section 3.4, present the relevant data in Section 3.5, perform several analysis in Section 3.6, and finally, draw conclusions in Section 3.7.

3.2 Literature

There are two streams of literature related to the spillover effects of incomplete regulation. One literature is on regulation-induced pollution substitution. Previous studies show that under the pressure of environmental regulation, pollution spillovers could happen at different dimensions. First, there is a lot of discussion of the pollution substitution across national or regional borders, which is usually referred to as the "pollution heaven" literature. For instance, Fowlie (2009) discovered that the greenhouse emission reductions brought about by California's unilateral action in the energy sector could be neutralized by leakage to unregulated regions. Cai et al. (2016) discovered that the regulation of water-polluting industries led to pollution spillovers affecting their downstream neighboring counties.

Second, the substitution could exist inside firms at the pollutants level, which usually happens due to changes in abatement costs for different production processes. While the earlier study of Greenstone (2003) found no evidence of air-to-water pollution substitution in the U.S. iron and steel industry, more recent research suggests a different trend. Specifically, facilities under air emissions regulation were found to increase their water-to-air emissions ratio by 177% and their water emissions level by 105%. This was observed when the researchers narrowed down to facilities located within a 1.07-kilometer radius of an air monitor that exceeded the NAAQS (Gibson, 2019).

Third, pollution substitution could arise across facilities within the same parent companies. Gibson (2019) provides an insightful observation of intrafirm leakage. The study indicates that when a firm has one or more facilities located proximate to a Particulate Matter (PM) non-attainment monitor, the facilities of the same firm in attainment counties—those counties that meet the air quality standards—tend to increase their air emissions by approximately 11%. Similarly, using a high priority violation as an exogenous shock, Rijal and Khanna (2020) show that a compliant facility will increase its air emissions by about 43 percent if it had at least one other sister facility, within the same 6-digit NAICS industry code and belonging to the same parent firm, under violation. Outside of the United States, Chen, Chen, Liu, Suárez Serrato, and Xu (2021) find that China's Top-1000 energy saving program causes conglomerates to move production away from regulated firms to unregulated ones, which diminishes the energy saving achievement of the program. Xiao, Yin, and Moon (2023) also find evidence of a significant slowdown in the market share growth of the targeted Top-1000 enterprises compared to similar untargeted enterprises.

This paper also relates to the broader literature on bunching, the behavioral responses to thresholds (Kleven, 2016). Economic theory predicts that rational agents would optimize their behavior to maximize utility when certain interventions distort the incentives at a threshold. For example, Saez (2010) finds that taxpayers bunch right below the line where the marginal tax rate jumps from zero to positive in response to a kink in the personal income tax schedule induced by the Earned Income Tax Credit (EITC). Meanwhile, several papers find bunching below the threshold at which the value added tax (VAT) takes effect in different countries, including Japan (Onji, 2009), Finland (Harju et al., 2019), and the United Kingdom (Liu et al., 2021).

3.3 Background: 40 ton per year as the threshold for "de minimis" pollution

In Colorado, any business that emits air pollution needs to report its emissions and/or apply for a permit. In the application process, business owners are required to submit an Air Pollutant Emission Notice(APEN), which describes in detail the emissions from the business's operation. The modeling staff of the state's environmental agency then uses the information from APEN to predict how much the business would contribute to local ambient air pollution. Finally, the air permit is granted based on the predictions. This permit granting process is in line with the environmental regulation at the federal level in the United States, which requires the state agency to "determine whether the construction or modification of a facility, building, structure or installation, or combination of these will result in... interference with attainment or maintenance of a national standard in the State in which the proposed source (or modification) is located or in a neighboring State".

In 2011, Colorado's environmental regulation agency introduced a rule called "Permit Section (PS) Memo 10-01" which sets a threshold for "de minimis" pollution[1]. According to PS Memo 10-01, all sources with NO_2 emission of 40 tons per year will be granted the air pollution permit without demonstrating compliance with the corresponding 1-hour National Ambient Air Quality Standards (NAAQS). After ten years of adopting this policy, in 2020, it was raised to public notice by a group of employees from the Air Pollution Control Division, Colorado Department of Public Health and Environment (CODPE). These employees claimed that such action deteriorated the already bad air pollution in many regions of Colorado. In April 2021, the director of CODPE told the media that Memo 10-01 had been revoked and would not be used for guiding the allocation of air pollution permits in the future, but the decade of practicing the "too small to be regulated" principle remains to be studied.

3.4 A simple theory model

Consider a facility operating under environmental constraints, where the government imposes a cost of size τ for every unit of pollution emitted. In this case, the facility's total pollution-related cost under such regulation is represented by the following equation:

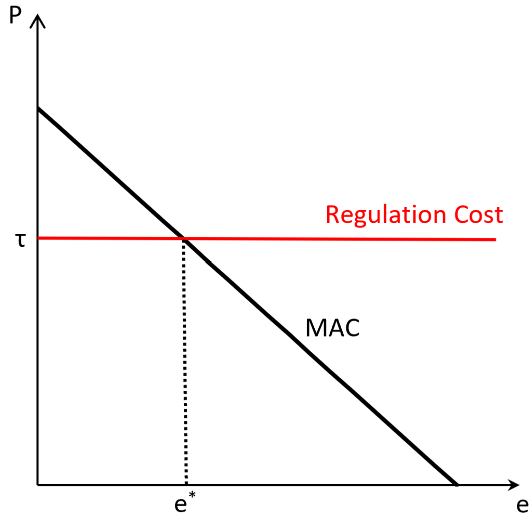
$$TC(e) = C(e) + \tau \cdot e$$

Here, e is the emission level, and $C(e)$ represents the abatement cost. For cost minimization, the facility's optimal choice would be to emit at a level e^*, which satisfies the first-order condition, in which marginal abatement cost (MAC) equals the external regulation cost, τ:

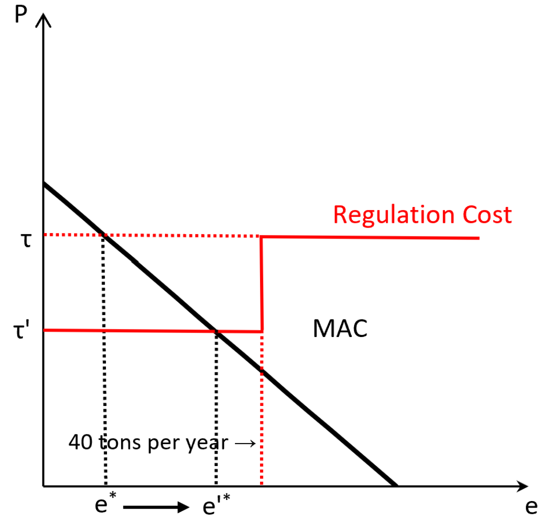
$$-C'(e^*) = \tau$$

I now allow the regulatory cost, τ, to vary according to a facility's annual emissions. If a facility's yearly emissions falls below 40 tons, τ is reduced to a smaller value, τ', where $\tau' < \tau$.

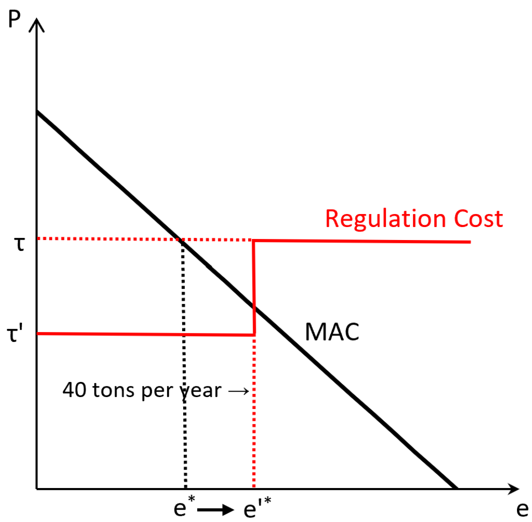
[1] This rule is labelled as nonbinding guidance, but were used as the "supersede statutory /regulatory requirements" since April 2011, according to the public letter of insider of Colorado's Department of Public Health and Environment.



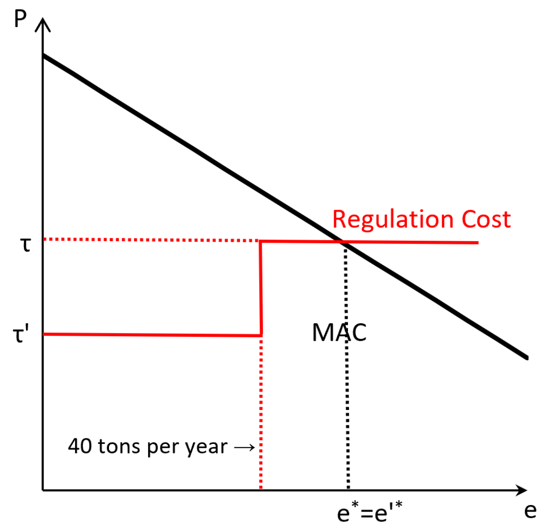
(a) Universal Regulatory Cost



(b) Emitters Away From Cutoff



(c) Emitters Close to Cutoff



(d) Large Emitters

Figure 3.1 Optimal Emission with Regulatory Cost Change

Graph (a) illustrates a scenario where all emitters are uniformly regulated. Graph (b) depicts the behavior changes of a small emitter whose emission is significantly below the cutoff. Graph (c) depicts the action changes of a small emitter whose emission is slightly below the cutoff. Graph (d) depicts the change in behavior for a large emitter whose original optimal emission significantly surpasses the cutoff.

As illustrated in Figure 3.1, the change in regulatory cost triggers a change in the optimal emission level, depending on the original optimal emission level of the facility. For those emitters whose initial emission level is sufficiently far from the cutoff, their optimal emission level will noticeably increase from e^* to e'^*, as displayed in Figure 3.1 (b). For emitters whose original emission level is below but close to the cutoff, their emissions will rise to meet the cutoff level (Figure 3.1 (c)). Conversely, for large emitters, this relaxed regulation aimed at smaller emitters will have no impact (Figure 3.1 (d)).

The analysis above concentrates on the impact on individual facilities. However, there may also be effects on firms operating multiple facilities. Given that the regulatory cost decreases for facilities emitting less than 40 tons per year, multi-facility firms might consider redistributing a portion of their emissions from higher-emission facilities to lower-emission ones, or even building new facilities whose emission is smaller than the cutoff. By doing so, both types of facilities could potentially benefit from reduced regulatory costs. Nevertheless, this spillover effect would hinge on the associated costs of transferring emissions or production. In the event that such an emissions transfer occurs, not only might I observe an increase in emissions from smaller emitters, but I might also witness a growth in the number of smaller emitters.

In conclusion, a decrease in regulatory costs for emitters below a certain cutoff may lead to increased emissions from smaller emitters. Additionally, it could potentially result in a rise in the overall number of emitters, as demonstrated in Figure 3.2.

3.5 Data

In this study, I utilize data from the National Emissions Inventory (NEI) database to gather facility-level emissions information. I also obtain county-level non-attainment designation data from the EPA's Green Book to proxy for the general regulation intensity for each county. Additionally, I attempt to use the Facility Registry Service (FRS) system to acquire facility-level information on firm structure, to test for the redistribution of emissions within firms. Unfortunately, many Colorado facilities do not report the parent company in FRS. I discuss this issue further in Appendix 3C.

Facility-level Data: National Emissions Inventory (NEI) database The NEI database is developed and maintained by the EPA, and it is published every three years. This comprehensive database includes details such as types of pollutants emitted, quantities of each pollutant released, and the geographic location of each facility. In this paper, our focus is specifically on NOx pollution

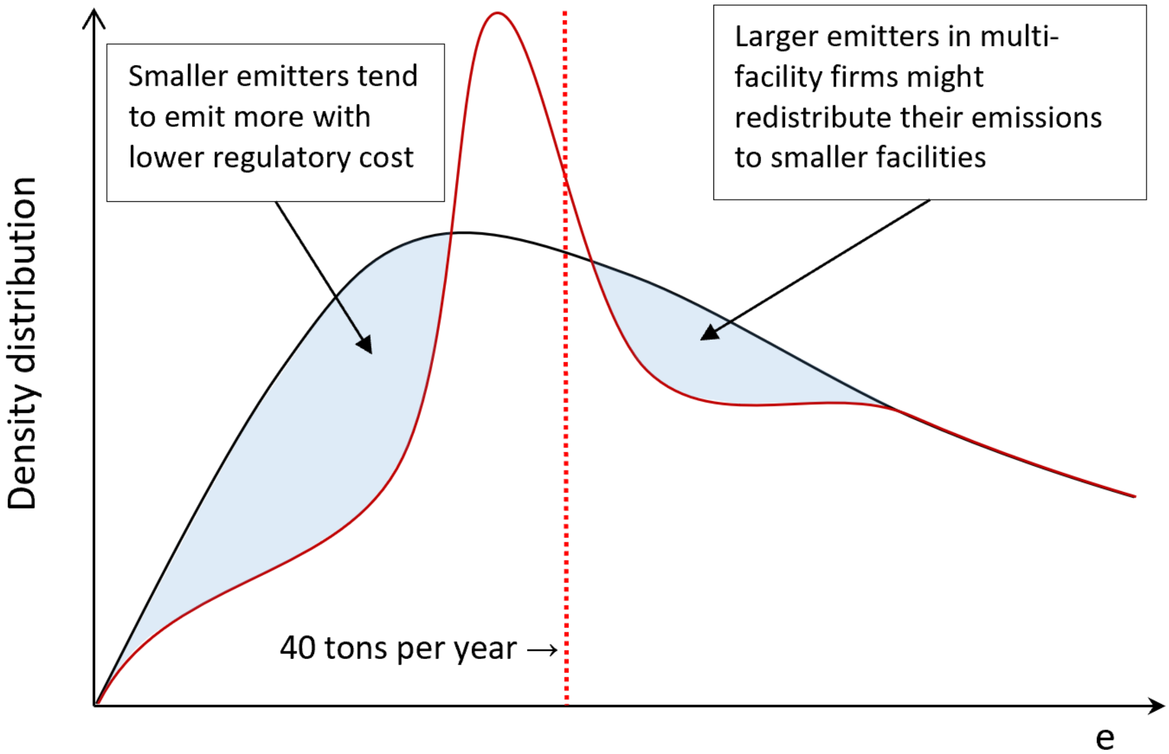


Figure 3.2 Facility-level emission distribution change

Note: This figure depicts the potential shift in distribution that could be triggered by changes in emission regulatory costs. The black line represents the density distribution prior to the policy alteration, while the red line represents the density distribution after the policy change.

Table 3.1 Summary Statistic: NOx Emissions by Facility

Year	Obs	Mean	Std. dev.	Min	Max
2008	57,766	92.219	881.405	0	45518.5
2011	61,938	65.171	581.125	0	38729.12
2014	58,330	59.857	529.001	0	22909.25
2017	57,671	46.120	378.597	0	16291.54
Distinct Facilities:	80,617				
Total Observations:	235,705				

Note: Data is obtained by using all NEI facilities with emissions classified as “Nitrogen Oxides” under the variable “pollutant_desc”.

emission data. I utilize information on facilities for the years 2008, 2011, 2014, and 2017 who reported NOx emissions. As shown in Table 3.1 there are 80,617 different facilities recorded during the period of 2008-2017. On average, the mean emission of NOx is lower than 100 tons per year,

Table 3.2 Summary Statistic: County-level Data

Variable	Obs	Mean	Std. dev.	Min	Max
Emission%	12,618	49.289	45.686	0	100
Facility%	12,618	86.946	17.674	0	100
Non-attainment					
Ozone	12,618	0.0755	0.2643	0	1
SO2	12,618	0.0058	0.0758	0	1
CO	12,618	0.0002	0.0154	0	1
PM10	12,618	0.0501	0.2181	0	1
Pb	12,618	0.0045	0.0671	0	1

Note: The variable "Emission%" denotes the share of emissions from smaller emitters that discharge less than 40 tons annually. The "Facility%" variable represents the percentage of facilities with annual emissions under 40 tons.

but the data has a large range. The emissions data further reveals a median of 0.24 tons and a 75th percentile at 5.61 tons. Specifically for Colorado, the figures are slightly higher with a median of 0.91 tons and a 75th percentile of 6.37 tons.

County-level Data Our theory predicts that the policy could lead to an increase in number of small facilities, in addition to causing an increase in emission from existing small facilities. Thus, I consolidate the facility-level data into a county-level dataset. This aggregation yields two variables: one reflecting the proportion of emissions from facilities under the policy cutoff, and another representing the percentage of facilities falling below the emission cutoff. To account for the potential changes in regulation intensity, I utilize the EPA's Green Book to determine each county's non-attainment status, where a value of one indicates that a county is not meeting the NAAQS standards for a particular pollutant. A non-attainment status implies that the areas need to take actions to lower their pollution level, and face stricter regulations. The county-level data are presented in Table 3.2. As shown in the data, on average, smaller emitters comprise 87% of the total facility count, while only accounting for 49% of total emissions.

3.6 Analysis

The first analysis employed in this paper involves a straightforward descriptive examination of facility-level emissions data, with a particular focus on emissions around 40 tons per year in Colorado. Figure 3.3depicts the density distribution of facility-level emissions in Colorado for different years. It reveals a generally smooth pattern without noticeable bunching. Actually, I observe that there is an increased count of facilities with emissions slightly above 10 tons but below the 40-ton cutoff in 2008, prior to the initiation of Colorado's PS Memo 10-01. This pattern suggests that there are more facilities exist below the 40-ton cutoff before the policy takes effect, and fewer facilities after. This preliminary descriptive analysis suggests that no significant changes in facilities' emission behavior are to be expected.

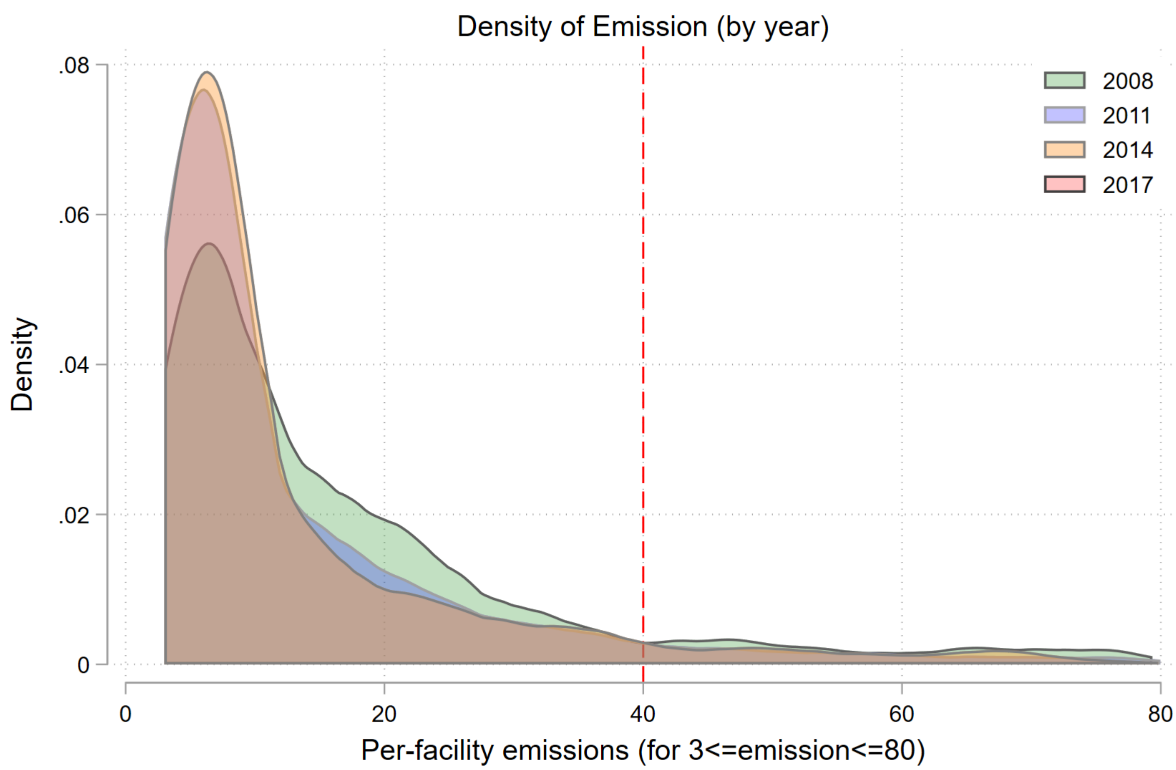
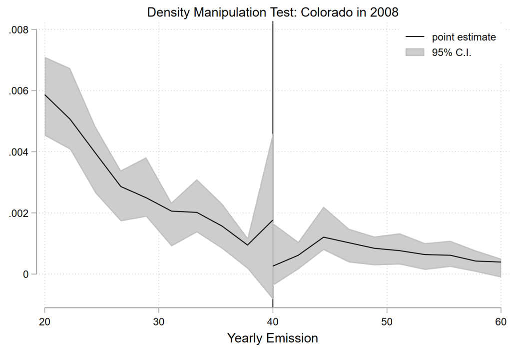


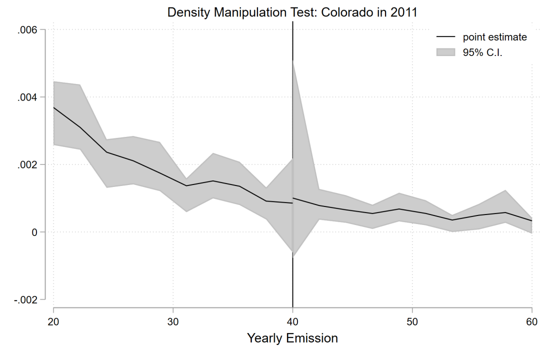
Figure 3.3 Facility-level emission distribution in Colorado

Note: The density distribution is plotted with the bandwidth as 2.5.

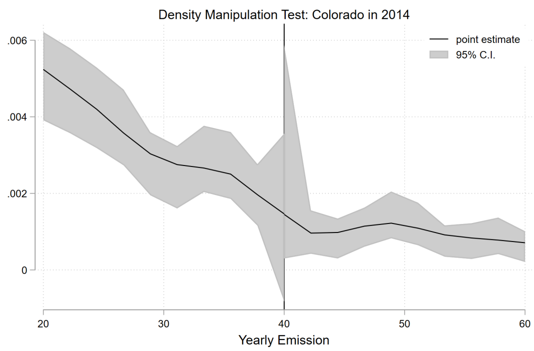
Secondly, I incorporate the manipulation testing technique, as developed by Cattaneo, Jansson, and Ma (2020). This technique aims to identify any concentration of data points, or "bunching", near the 40 tons threshold. As discussed in Section 3.4, I would expect to see an increase in the number of facilities with emissions approaching the cutoff level following the policy change, if smaller facilities respond to the policy by increasing their emissions. This bunching test comprises



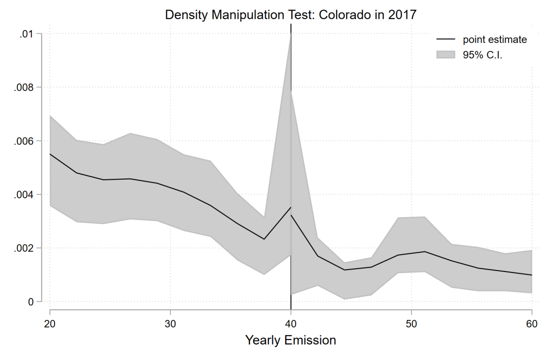
(a) 2008



(b) 2011



(c) 2014



(d) 2017

Figure 3.4 Manipulation Test Around the 40-Ton Threshold

Note: Graphs are obtained by using the Stata package "rddensity".

three steps. First, it creates a smooth local approximation of the empirical distribution function using a polynomial expansion, and derives the density from the slope coefficient in the local polynomial regression. Then, this density estimator is applied to data both below and above the cutoff point, thus generating two separate density estimates at the boundary. Finally, I assess the difference between these two estimates, examining whether the difference significantly deviates from zero. Figure 3.4 presents the results of our manipulation test. I find no evidence suggesting manipulation in proximity to the annual cutoff of 40 tons of NOx emissions, either before or after the policy change.

The third analysis I use is the difference-in-differences (DID) method to contrast emissions from facilities located in Colorado with other states. The application of the DID method allows us to analyze two metrics: the proportion of emissions originating from facilities that fall below the annual cutoff of 40 tons, and the proportion of total facilities whose emissions remain under this cutoff. These metrics have been chosen based on our prediction that the regulation will lead to an increase in emissions from facilities remaining below the cutoff, as well as an uptick in the number of facilities with emissions under 40 tons. This is based on the theory that companies operating multiple facilities would opt to redistribute their emissions to new facilities to stay beneath the cutoff.

Comparing to the sharp bunching method above, this DID method would utilize the data from other states and rely on them to construct a counterfactual. Furthermore, Colorado might be more comparable to other states that also experience the shale gas boom [2]. Thus, I also compare Colorado to other states with shale oil production, such as Texas, North Dakota, and Wyoming. The following equation is the DID model that I use:

$$y_{ct} = \alpha_c + \beta D_{ct} + \Gamma \cdot Non\text{-}attainment_{ct} + Year_{rt} + \epsilon_{ct},$$

where y_{ct} represents the two key metrics that I choose to use: the proportion of emissions originating from facilities below the cutoff, and the proportion of facilities whose emissions are under the cutoff. These metrics are provided at the county level for each county c and year t. The α_c is the county-level fixed effect, which could absorb the time-invariant differences between counties. D_{ct} indicates whether county c is located in Colorado while the policy change is in effect. $Non\text{-}attainment_{ct}$ is the non-attainment status for each county c, while might vary in different years t. The variable $Year_{rt}$ represents year-region dummies, where r represents the different EPA regional offices. This allows us to have separate time trends for locations that are under different EPA regional offices. The ϵ_{ct} is the unobserved error term.

Table 3.3 provides the results of implementing the aforementioned DID model. Notably, the key treatment variable's coefficient remains insignificant across all regressions. There is neither a change in the proportion of emissions from facilities below the cutoff nor an increase in the number of facilities with emissions below this threshold, as clearly illustrated in Table 3.3. As

[2]When the PS Memo 10-01 was brought to the public's attention, concerns were raised about whether the policy was inadvertently encouraging increased pollution by attracting more companies to the gas and oil industry during the shale boom.

Table 3.3 The Impact of PS Memo-10-01: CO vs All Other States

	(1) ln(Emission%)	(2) ln(Facility%)	(3) ln(Emission%)	(4) ln(Facility%)
Treat · Post	0.152	0.0267	0.151	0.0267
	(0.218)	(0.0171)	(0.218)	(0.0171)
Constant	2.503***	4.453***	2.519***	4.453***
	(0.0125)	(0.000979)	(0.0162)	(0.00127)
County FE	Yes	Yes	Yes	Yes
EPA regional time trend	Yes	Yes	Yes	Yes
Non-Attainment			Yes	Yes
Obs	12426	12467	12426	12467
R-square	0.806	0.834	0.806	0.834

Note: all standard errors are clustered at the county level. $* \; p < 0.10$, $* \; p < 0.05$, $* * * \; p < 0.01$.

Table 3.4 The Impact of PS Memo-10-01: CO vs TX, WY, and ND

	(1) ln(Emission%)	(2) ln(Facility%)	(3) ln(Emission%)	(4) ln(Facility%)
Treat · Post	-0.402	-0.0234	-0.388	-0.0217
	(0.276)	(0.0229)	(0.276)	(0.0230)
Constant	2.290***	4.422***	2.248***	4.417***
	(0.0494)	(0.00411)	(0.0664)	(0.00552)
County FE	Yes	Yes	Yes	Yes
EPA regional time trend	Yes	Yes	Yes	Yes
Non-Attainment			Yes	Yes
Obs	1570	1570	1570	1570
R-square	0.817	0.853	0.817	0.853

Note: all standard errors are clustered at the county level. $* \; p < 0.10$, $* \; p < 0.05$, $* * * \; p < 0.01$.

columns (3) and (4) reveal, when controlling for county-level non-attainment status, the results display negligible change.

In Table 3.4, I narrow our comparison to three states with greater shale gas production, Texas, Wyoming, and Nevada, as opposed to the entirety of the United States. Despite this modification, the estimates remain close to zero. Consequently, the regression results yield no evidence to suggest that the “too small to regulate” approach significantly influences facility emission behavior.

3.7 Conclusion

In this paper, I examine the impact of a policy change in Colorado that exempts smaller pollution emitters from stringent environmental regulations. My theoretical analysis suggests that lowering the marginal cost for these smaller emitters should incentivize increased emissions. To test this theory, I use the NEI's facility-level emissions data. Contrary to my hypothesis, I find no evidence suggesting increased emissions from smaller facilities.

A few factors could potentially account for this result: First, the policy shift in Colorado might not effectively diminish marginal costs for smaller emitters. PS Memo 10-01 relieves these smaller facilities from the obligation of demonstrating their minimal impact on local air quality, but this might not create sufficient incentive to modify their behavior. Second, the NEI database might not be the optimal source for studying the effects on smaller emitters. The NEI data is sourced from emissions estimates submitted to the EPA by state, local, and tribal air agencies, and I cannot ascertain if the data procurement process or pollution estimation does not change over time. These potential influences necessitate further investigation.

REFERENCES

Abadie, A., Diamond, A., & Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program. *Journal of the American statistical Association*, *105*(490), 493–505.

Abadie, A., & Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American economic review*, *93*(1), 113–132.

Almaraz, M., Bai, E., Wang, C., Trousdell, J., Conley, S., Faloona, I., & Houlton, B. Z. (2018). Agriculture is a major source of nox pollution in california. *Science advances*, *4*(1), eaao3477.

Anderson, S. T., & Sallee, J. M. (2016). Designing policies to make cars greener: A review of the literature.

Auffhammer, M., Bento, A. M., & Lowe, S. E. (2009). Measuring the effects of the clean air act amendments on ambient pm10 concentrations: The critical importance of a spatially disaggregated analysis. *Journal of Environmental Economics and Management*, *58*(1), 15–26.

Auffhammer, M., & Kellogg, R. (2011). Clearing the air? the effects of gasoline content regulation on air quality. *American Economic Review*, *101*(6), 2687–2722.

Banzhaf, S., Ma, L., & Timmins, C. (2019). Environmental justice: The economics of race, place, and pollution. *Journal of Economic Perspectives*, *33*(1), 185–208.

Bergquist, P. (2018). *Controlling the regulators: How party control of government shapes environmental regulation in the 21st century* (Tech. Rep.). Working paper.

Bloomer, B. J., Stehr, J. W., Piety, C. A., Salawitch, R. J., & Dickerson, R. R. (2009). Observed relationships of ozone air pollution with temperature and emissions. *Geophysical Research Letters*, *36*(9).

Borusyak, K., Jaravel, X., & Spiess, J. (2021). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint arXiv:2108.12419*.

Cai, H., Chen, Y., & Gong, Q. (2016). Polluting thy neighbor: Unintended consequences of china's pollution reduction mandates. *Journal of Environmental Economics and Management*, *76*, 86–104.

CARB. (1990). *Proposed regulations for low-emission vehicles and clean fuels.* State of California, Air Resources Board Sacramento, CA.

Carlson, A. E., Hankins, M. J., & Stein, J. E. (2019). *Shifting gears: The federal government's reversal on california's clean air act waiver* (Report).

Cattaneo, M. D., Jansson, M., & Ma, X. (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association*, *115*(531), 1449–1455.

Chen, Q., Chen, Z., & Liu, Z. (2021). *Regulating conglomerates: Evidence from an energy conservation program in china.* Unpublished.

Chen, Q., Chen, Z., Liu, Z., Suárez Serrato, J. C., & Xu, D. Y. (2021). Industrial energy regulation: The role of business conglomerates in china. In *Aea papers and proceedings* (Vol. 111, pp. 396–400).

Currie, J., & Walker, R. (2019). What do economists have to say about the clean air act 50 years after the establishment of the environmental protection agency? *Journal of Economic Perspectives*, *33*(4), 3–26.

Davis, L. W. (2008). The effect of driving restrictions on air quality in mexico city. *Journal of Political Economy*, *116*(1), 38–81.

Dickert-Conlin, S., Elder, T., & Teltser, K. (2019). Allocating scarce organs: How a change in supply affects transplant waiting lists and transplant recipients. *American Economic Journal: Applied Economics*, *11*(4), 210–39.

Dube, A., & Zipperer, B. (2015). Pooling multiple case studies using synthetic controls: An application to minimum wage policies.

Duflo, E., Greenstone, M., Pande, R., & Ryan, N. (2018). The value of regulatory discretion: Estimates from environmental inspections in india. *Econometrica*, *86*(6), 2123–2160.

Dwyer-Lindgren, L., Bertozzi-Villa, A., Stubbs, R. W., Morozoff, C., Kutz, M. J., Huynh, C., . . . others (2016). Us county-level trends in mortality rates for major causes of death, 1980-2014. *Jama*, *316*(22), 2385–2401.

Dwyer-Lindgren, L., Bertozzi-Villa, A., Stubbs, R. W., Morozoff, C., Shirude, S., Naghavi, M., . . . Murray, C. J. (2017). Trends and patterns of differences in chronic respiratory disease mortality among us counties, 1980-2014. *Jama*, *318*(12), 1136–1149.

Feng, Y., Fullerton, D., & Gan, L. (2013). Vehicle choices, miles driven, and pollution policies. *Journal of Regulatory Economics*, *44*(1), 4–29.

Fowlie, M. L. (2009). Incomplete environmental regulation, imperfect competition, and emissions leakage. *American Economic Journal: Economic Policy*, *1*(2), 72–112.

Gao, H. O. (2007). Day of week effects on diurnal ozone/nox cycles and transportation emissions in southern california. *Transportation Research Part D: Transport and Environment*, *12*(4), 292–305.

Gibson, M. (2019). Regulation-induced pollution substitution. *Review of Economics and Statistics*, *101*(5), 827–840.

Goodman-Bacon, A. (2018). *Difference-in-differences with variation in treatment timing* (Tech. Rep.). National Bureau of Economic Research.

Grainger, C., & Schreiber, A. (2019). Discrimination in ambient air pollution monitoring? In *Aea papers and proceedings* (Vol. 109, pp. 277–82).

Grainger, C., Schreiber, A., & Chang, W. (2018). Do regulators strategically avoid pollution hotspots when siting monitors? evidence from remote sensing of air pollution. *University of Madison–Wisconsin working paper*.

Gray, W. B., & Deily, M. E. (1996). Compliance and enforcement: Air pollution regulation in the us steel industry. *Journal of environmental economics and management*, *31*(1), 96–111.

Gray, W. B., & Shadbegian, R. J. (2005). When and why do plants comply? paper mills in the 1980s. *Law & Policy*, *27*(2), 238–261.

Gray, W. B., & Shimshack, J. P. (2011). The effectiveness of environmental monitoring and enforcement: A review of the empirical evidence. *Review of Environmental Economics and Policy*.

Greenstone, M. (2003). Estimating regulation-induced substitution: The effect of the clean air act on water and ground pollution. *American Economic Review*, *93*(2), 442–448.

Greenstone, M., List, J. A., & Syverson, C. (2012). *The effects of environmental regulation on the competitiveness of us manufacturing* (Tech. Rep.). National Bureau of Economic Research.

Hanna, R. N., & Oliva, P. (2010). The impact of inspections on plant-level air emissions. *The BE Journal of Economic Analysis & Policy*, *10*(1).

Harju, J., Matikka, T., & Rauhanen, T. (2019). Compliance costs vs. tax incentives: Why do entrepreneurs respond to size-based regulations? *Journal of Public Economics*, *173*, 139–164.

Heck, W. W., Taylor, O., Adams, R., Bingham, G., Miller, J., Preston, E., & Weinstein, L. (1982). Assessment of crop loss from ozone. *Journal of the Air Pollution Control Association*, *32*(4), 353–361.

Hejny, J. (2018). The trump administration and environmental policy: Reagan redux? *Journal of Environmental Studies and Sciences*, *8*(2), 197–211.

Holmes, K. J., Allen, D., & Russel, M. (2007). *The roles of state and federal mobile-source emissions standards.* ACS Publications.

Hsiang, S., Oliva, P., & Walker, R. (2019). The distribution of environmental damages. *Review of Environmental Economics and Policy*, *13*(1), 83–103.

Jacobsen, M. R., Sallee, J. M., Shapiro, J., & van Benthem, A. (2022). Regulating untaxable externalities: Are vehicle air pollution standards effective and efficient?

Kahn, M. (1996). New evidence on trends in vehicle emissions. *The RAND Journal of Economics*, 183–196.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, *8*, 435–464.

Lim, J. (2016). The impact of monitoring and enforcement on air pollutant emissions. *Journal of Regulatory Economics*, *49*, 203–222.

Liu, L., Lockwood, B., Almunia, M., & Tam, E. H. (2021). Vat notches, voluntary registration, and bunching: Theory and uk evidence. *Review of Economics and Statistics*, *103*(1), 151–164.

Marcus, M. (2017). On the road to recovery: Gasoline content regulations and child health. *Journal of health economics*, *54*, 98–123.

Metaxoglou, K., & Smith, A. (2020). Productivity spillovers from pollution reduction: Reducing coal use increases crop yields. *American Journal of Agricultural Economics*, *102*(1), 259–280.

Mu, Y., Rubin, E. A., & Zou, E. (2021). *What's missing in environmental (self-) monitoring: Evidence from strategic shutdowns of pollution monitors* (Tech. Rep.). National Bureau of Economic Research.

NESCAUM, N. S. f. C. A. U. a. M. (1991). Adopting the california low emission vehicle program in the northeast states- an evaluation. executive summary [Journal Article].

Nosrati, E., Kang-Brown, J., Ash, M., McKee, M., Marmot, M., & King, L. P. (2019). Economic decline, incarceration, and mortality from drug use disorders in the usa between 1983 and 2014: an observational analysis. *The Lancet Public Health*, *4*(7), e326–e333.

Oliva, P. (2015). Environmental regulations and corruption: Automobile emissions in mexico city. *Journal of Political Economy*, *123*(3), 686–724.

Oltmans, S. J., & Levy, H. (1992). Seasonal cycle of surface ozone over the western north atlantic. *Nature*, *358*(6385), 392–394.

Onji, K. (2009). The response of firms to eligibility thresholds: Evidence from the japanese value-added tax. *Journal of Public Economics*, *93*(5-6), 766–775.

O'Connor, J. M., Sedghi, T., Dhodapkar, M., Kane, M. J., & Gross, C. P. (2018). Factors associated with cancer disparities among low-, medium-, and high-income us counties. *JAMA network open*, *1*(6), e183146–e183146.

Rabassa, M. J. (2008). *Ambient air pollution and the allocation of environmental enforcement effort*. University of Illinois at Urbana-Champaign.

Reed Jr, L. H. (1996). California low-emission vehicle program: forcing technology and dealing effectively with the uncertainties. *BC Envtl. Aff. L. Rev.*, *24*, 695.

Rijal, B., & Khanna, N. (2020). High priority violations and intra-firm pollution substitution. *Journal of Environmental Economics and Management*, *103*, 102359.

Saez, E. (2010). Do taxpayers bunch at kink points? *American economic Journal: economic policy*, *2*(3), 180–212.

Sanders, N. J., & Sandler, R. (2020). Technology and the effectiveness of regulatory programs over time: Vehicle emissions and smog checks with a changing fleet. *Journal of the Association of Environmental and Resource Economists*, *7*(3), 587–618.

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

Simeonova, E., Currie, J., Nilsson, P., & Walker, R. (2021). Congestion pricing, air pollution, and children's health. *Journal of Human Resources*, *56*(4), 971–996.

Viard, V. B., & Fu, S. (2015). The effect of beijing's driving restrictions on pollution and economic activity. *Journal of Public Economics*, *125*, 98–115.

Wooldridge, J. (2021). Two-way fixed effects, the two-way mundlak regression, and difference-in-differences estimators. *Available at SSRN 3906345*.

Xiao, Y., Yin, H., & Moon, J. J. (2023). Coping with externally imposed energy constraints: Competitiveness and operational impact of china's top-1000 energy-consuming enterprises program. *The Energy Journal*, *44*(2).

Xu, Y., Zhang, Q., & Zheng, S. (2015). The rising demand for subway after private driving restriction: Evidence from beijing's housing market. *Regional Science and Urban Economics*, *54*, 28–37.

Zhang, J., Wei, Y., & Fang, Z. (2019). Ozone pollution: a major health hazard worldwide. *Frontiers in immunology*, *10*, 2518.

Zou, E. Y. (2021, July). Unwatched pollution: The effect of intermittent monitoring on air quality. *American Economic Review*, *111*(7), 2101–26. Retrieved from `https://www.aeaweb.org/articles?id=10.1257/aer.20181346` doi: 10.1257/aer.20181346

APPENDIX A

MONTHLY TRENDS OF OZONE POLLUTION

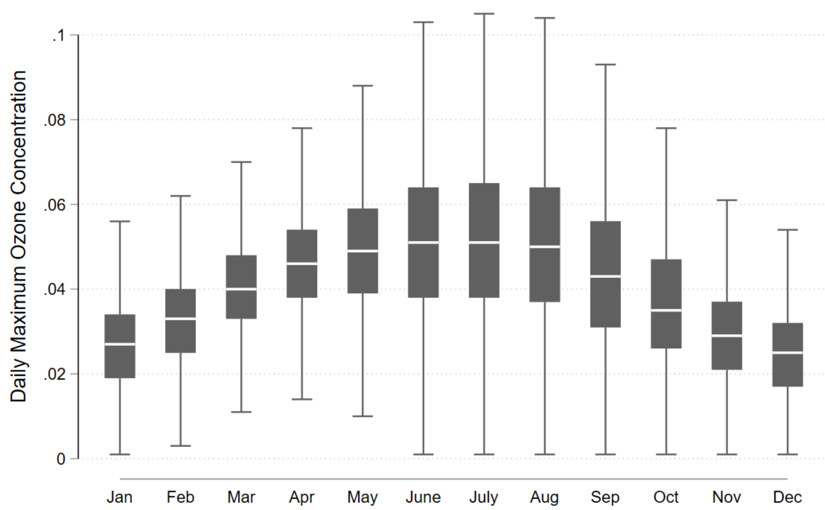


Figure A.1 Monthly trends of ozone pollution

APPENDIX B

APPLYING TWO-WAY MUNDLAK REGRESSION

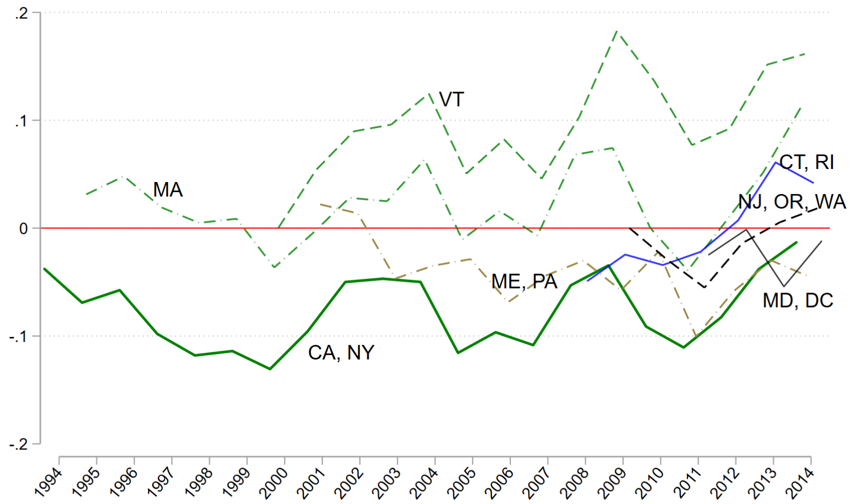


Figure B.1 Two-way Mundlak regression results

Note: The figure present the average treatment effects of LEV for different cohort group

As shown in Fig. B.1, I present the results by applying the two-way Mundlak regression method proposed in Wooldridge (2021). This method identifies the average treatment effects for different cohort groups by adding both unit-specific time series averages and period-specific cross-sectional averages in a POLS regression. The results show that there is a strong heterogeneity in the treatment effect of LEV, and California and New York may have contributed most to the impact found in the DD analysis.

APPENDIX C

THE PROBLEM OF MERGING FRS AND NEI DATABASE

The Facility Registry Service (FRS) system, developed by the EPA, serves as a centralized repository of information on all facilities subject to environmental regulation. The "organization" part of the FRS provides information regarding the parent companies of each facility.This includes two variables: the Data Universal Number System (DUNS) number, which is a numeric identifier assigned by Dun and Bradstreet, and the name of the organization to which the facility belongs[1].

To identify the parent firm, I utilize both variables. Facilities with the same DUNS number or the same organization name are considered as under the same company. In cases where facilities share the same DUNS number but a different company name (or vice versa), they are still treated as part of the same company.

However, after combining data on parent companies, I find that information regarding the parent companies of many facilities is not available, with availability varying considerably by state. As provided in Table C.1, I observe that information about the parent company is available for merely 84,241 observations (although 230,408 out of 235,705 observations in the NEI successfully matched with records in the FRS), with a significant variation in data availability across different states. For instance, in California, parent company information is available for 42,000 out of 48,565 observations, which corresponds to an availability rate of 86.48%. However, for Colorado, the primary focus of our study, the quality of data is considerably lower, with parent company information available for only 69 out of 15,284 observations. Consequently, the organization data in FRS could not provide valuable insights into the structure of firms' facilities.

[1] The organization information utilized in this study is extracted from the "Organization File" available on the EPA's Facility Registry Service (FRS) information website.

Table C.1 Obtaining the Parent Company's Infor. for NEI facilities

State	# Facilities	Facility w/ Parent Company Info.	% Parent Company Info.
AK	3,760	526	13.99%
AL	2,793	109	3.90%
AR	1,924	624	32.43%
AZ	1,672	371	22.19%
CA	48,565	42,000	86.48%
CO	15,284	69	0.45%
CT	769	316	41.09%
DC	186	84	45.16%
DE	574	383	66.72%
DM	4,543	8	0.18%
FL	5,859	270	4.61%
GA	2,853	54	1.89%
HI	306	101	33.01%
IA	2,392	1,046	43.73%
ID	1,317	14	1.06%
IL	12,970	194	1.50%
IN	4,165	1	0.02%
KS	3,605	1,960	54.37%
KY	5,130	312	6.08%
LA	4,747	2,738	57.68%
MA	3,955	3,001	75.88%
MD	1,990	700	35.18%
ME	1,317	40	3.04%
MI	6,757	133	1.97%
MN	7,658	163	2.13%
MO	3,741	123	3.29%
MS	2,045	1,267	61.96%
MT	1,990	861	43.27%
NC	6,137	3,746	61.04%
ND	1,344	187	13.91%
NE	2,205	114	5.17%
NH	740	16	2.16%
NJ	3,350	171	5.10%

Table C.1 (cont'd)

State	# Facilities	Facility w/ Parent Company Info.	% Parent Company Info.
NM	1,298	598	46.07%
NV	1,712	895	52.28%
NY	4,045	1,527	37.75%
OH	5,813	133	2.29%
OK	5,267	3,686	69.98%
OR	2,250	430	19.11%
PA	7,673	3,428	44.68%
PR	654	18	2.75%
RI	1,373	20	1.46%
SC	1,892	88	4.65%
SD	825	7	0.85%
TN	2,951	221	7.49%
TX	15,682	6,874	43.83%
UT	939	307	32.69%
VA	3,011	1,170	38.86%
VI	32	0	0.00%
VT	371	4	1.08%
WA	2,709	82	3.03%
WI	5,998	166	2.77%
WV	1,163	615	52.88%
WY	2,833	2,270	80.13%
Total	235,134	84,241	35.83%

Note: The "# Facilities" column is derived from the NEI database for the years 2008, 2011, 2014, and 2017. Information pertaining to parent companies is extracted from the "Organization File" in the FRS database, which includes all facilities associated with the "Emissions Inventory System (EIS)" program.