

The Environmental Cost of Place-Based Industrial Policy: Evidence from Industrial Zones in Vietnam

Trinh Pham*

This Version: October 28, 2026

Abstract

Special economic zones are widely adopted across countries to boost growth, yet their environmental cost is far less documented. This paper examines the local air pollution footprint of a zone program in Vietnam, which has opened more than 300 industrial zones over three decades. Exploiting their staggered establishment, I first show that a zone's operation increases satellite-detected industrial combustion within a 5 km radius more than twofold. Firm-level data show that this rise accompanies the entry of large and high-output plants, which are likely the baseline foundation of a zone's environmental footprint. Using wind-driven variation in downwind exposure, I find that the zones also raise PM_{2.5} in the surrounding population by an average of 0.70 $\mu\text{g}/\text{m}^3$, a 25-day loss of life expectancy per exposed person. This effect differs substantially between zones of the same age established in different periods. The zones with the larger effect host a heavier mix of industries, and that mix changed little as the zones filled. Under weak enforcement of standards on operating firms, the composition of admitted industries therefore likely plays a decisive role in explaining these differences. A back-of-the-envelope calculation suggests that the implied health cost can be as high as the program's local income gains for the average resident.

Keywords: place-based policy, industrial zones, air pollution, Vietnam

JEL Codes: Q53, O14, R11

*KDI School of Public Policy and Management. Email: tpham@kdischool.ac.kr; 263 Namsejong-ro, Sejong-si, Republic of Korea. I am grateful to the editor, Joshua Linn, and two anonymous reviewers for constructive feedback and suggestions that improved the rigor and clarity of the paper. I also thank seminar participants at the 7th KDI School–World Bank Development Impact Conference for helpful comments. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. The author declares no conflict of interest.

1 INTRODUCTION

Special economic zones are widely adopted across countries to attract investment and boost growth, with more than 5,000 operating worldwide (UNCTAD, 2019). Evaluations of a small set of programs typically find higher firm entry, productivity, employment, wages, and household welfare (Busso et al., 2013; Wang, 2013; Kline & Moretti, 2014; Lu et al., 2019; Lu et al., 2023; Abagna et al., 2025; Tafese et al., 2025; Pham, 2026). Much less is known about their environmental cost, even though zones concentrate the industrial activity that drives local air pollution. This paper estimates how much pollution a zone program imposes on nearby communities and examines the potential mechanisms behind that cost.

I study these questions using Vietnam’s industrial zone program, which has opened more than 300 zones since the early 1990s. The staggered establishment of these zones across the country allows me to identify their causal effect on local air pollution. Furthermore, the program’s three-decade span permits a comparison of the cost across establishment periods and years of operation. Throughout the period, enforcement of environmental standards on firms already operating remained weak, a gap between *de jure* regulation and *de facto* enforcement that is common to the developing countries in which zone policies operate.

To answer the first question, I proceed in two steps that link the source of the pollution to the population exposed to it. In the first step, I establish that industrial zones generate a localized increase in industrial combustion, as indicated by satellite-detected thermal anomalies from NASA’s FIRMS record (NASA FIRMS, 2021). Comparing cells at varying distances from zone centroids with cells 45 to 50 km away shows that the effect is concentrated within 5 km and is indistinguishable from zero beyond 10 km, consistent with combustion as a point-source phenomenon. Using this result to define treatment, I exploit the staggered establishment of zones to trace the effect over time. The effect accumulates over the first decade of operation rather than appearing when a zone opens, and the pattern is similar under alternative heterogeneity-robust estimators (Callaway & Sant’Anna, 2021; de Chaisemartin & d’Haultfoeuille, 2024). Under a Poisson specification (Chen & Roth, 2024), zone operation more than doubles the frequency of detections.

In the second step, I ask whether zone establishment raises ambient air pollution for the surrounding population, treating the 5 km ring from the first step as the source area from which emissions travel downwind. Unlike thermal anomalies, which are confined to the source pixel, fine particulate matter disperses with prevailing winds and accumulates at a distance from where it is emitted. Following a literature that exploits wind direction to identify the effects of spatially fixed pollution sources (e.g., Deryugina et al., 2019; Anderson, 2020), I use variation in downwind exposure to separate zone-induced pollution from background concentrations. Because a cell’s position relative to the wind changes with daily weather while its distance from the zone does not, a cell serves as its own control in the months when it lies mostly upwind of an operational zone. A one-standard-deviation increase in downwind exposure raises PM_{2.5} by 2.7%. The population-weighted zone-attributable PM_{2.5} increment is 0.70 $\mu\text{g}/\text{m}^3$, equivalent to 14% of the WHO annual guideline. Placebo tests before any nearby zone opens and a permutation test that

relocates zones within their provinces support the attribution of the effect to zone establishment.

Translating the pollution estimates into health costs, the population-weighted PM_{2.5} increment implies a loss of approximately 25 days of life expectancy per exposed individual, or roughly 3.2 million person-years across the 46 million residents of the transport ring. A back-of-the-envelope calculation that values a life-year at Vietnam's income level puts the cost at \$8 to \$44 per exposed person per year, against local income gains of roughly \$30 to \$50 (Pham, 2026).

To examine the potential mechanisms behind these effects, I follow the trade-and-environment literature, which distinguishes the scale of activity, its industrial composition, and the technique of production (Grossman & Krueger, 1995; Shapiro & Walker, 2018). I combine the pollution estimates with the firm-level data and energy module of the Vietnam Enterprise Survey. The rise in combustion at the source is consistent with the scale channel, which operates through the entry of firms. Manufacturing output in the districts within 5 km of a zone rises by about 30% over its first decade, and some of the firms in these districts are far larger than the typical manufacturing firm. These large and high-output plants are likely the baseline foundation of a zone's environmental footprint. Within each establishment cohort, the PM_{2.5} effect per unit of exposure remains relatively stable as the zones age. Across cohorts, the effect of zones established between 2003 and 2009 is about twice that of zones established before 2003 and larger at every age. The 2003–2009 zones host a heavier mix of industries: hot-process industries account for about a fifth of their tenants against a tenth in the older zones, and the mix changed little as the zones filled. In a setting where standards on operating firms are weakly enforced, the composition of admitted industries therefore likely plays a decisive role in explaining this difference. However, scale and technique cannot be ruled out: the 2003–2009 zones also have a higher share of very large plants among their tenants, and the data permit neither a decomposition of output scale nor a comparison of production techniques within industries.

This paper contributes to two literatures. The first concerns the environmental consequences of place-based development policies. The evidence concentrates on their economic effects: firm entry, productivity, employment, wages, and household welfare (Busso et al., 2013; Wang, 2013; Kline & Moretti, 2014; Neumark & Simpson, 2015; Lu et al., 2019; Lu et al., 2023; Abagna et al., 2025; Tafese et al., 2025; Pham, 2026). The environmental costs have only recently drawn attention: polluting-industry growth (Fan et al., 2023; Ma et al., 2024), air pollution (Duan et al., 2024), deforestation (Hess et al., 2021), and avoided or reversed damage under programs with environmental safeguards (Garg & Shenoy, 2021; Lin et al., 2026). These studies compare outcomes in areas with and without a program. For air pollution, such a comparison establishes neither where the pollution comes from nor who bears it. I propose a two-step approach that answers both questions: the first step detects the combustion a program generates at its sites, and the second attributes the PM_{2.5} arriving downwind to those sites through wind-driven variation in exposure. Both steps require only satellite and reanalysis data and the location and timing of program sites. The approach therefore applies where ground monitoring is sparse, and it draws on a broader practice of measuring from space what official records miss (Henderson et al., 2012; Burke et al., 2021). I

use it to examine the air pollution externality of Vietnam’s national zone program and the health burden it implies. The estimates measure the program’s local footprint (the pollution its zones impose on the people around them) rather than its net national effect. For the average resident of the nearby communities, the health cost can be as high as the program’s documented income gains.

The second contribution concerns the potential mechanisms behind the pollution cost of zones, which I examine through the three channels of the trade-and-environment literature. In the United States, cleaner production within industries accounts for most of the decline in manufacturing emissions (Levinson, 2009), attributed to regulation enforced on operating firms (Shapiro & Walker, 2018). In developing countries, where enforcement on operating firms is weaker (Duflo et al., 2013; Greenstone & Hanna, 2014), less is known about how the footprint of a place-based program arises and at what stage it is set. I use the three channels to organize the evidence on the level of combustion at the source and on the variation in the PM_{2.5} effect across zone cohorts, drawing on firm-level data on the industries each cohort of zones admitted and on the fuel that firms burn. The evidence does not separate the three channels, since that would require the output and emissions of each tenant. It nonetheless suggests that, where such enforcement is limited, a zone’s burden is largely decided when its projects are approved. Programs with environmental safeguards elsewhere likewise shaped which industries entered: a tax incentive program in India denied its benefits to environmentally damaging industries (Garg & Shenoy, 2021), and local governments in a new area in China raised entry barriers for polluting firms (Lin et al., 2026).

The remainder of the paper proceeds as follows. Section 2 describes the institutional background, including characteristics of industrial zones and related environmental regulations in Vietnam. Section 3 describes the data and provides summary statistics. Section 4 presents the empirical analysis of industrial combustion at the source. Section 5 presents the wind-based analysis of ambient air pollution. Section 6 discusses the potential mechanisms behind these estimates and translates them into a health burden. Section 7 concludes.

2 BACKGROUND

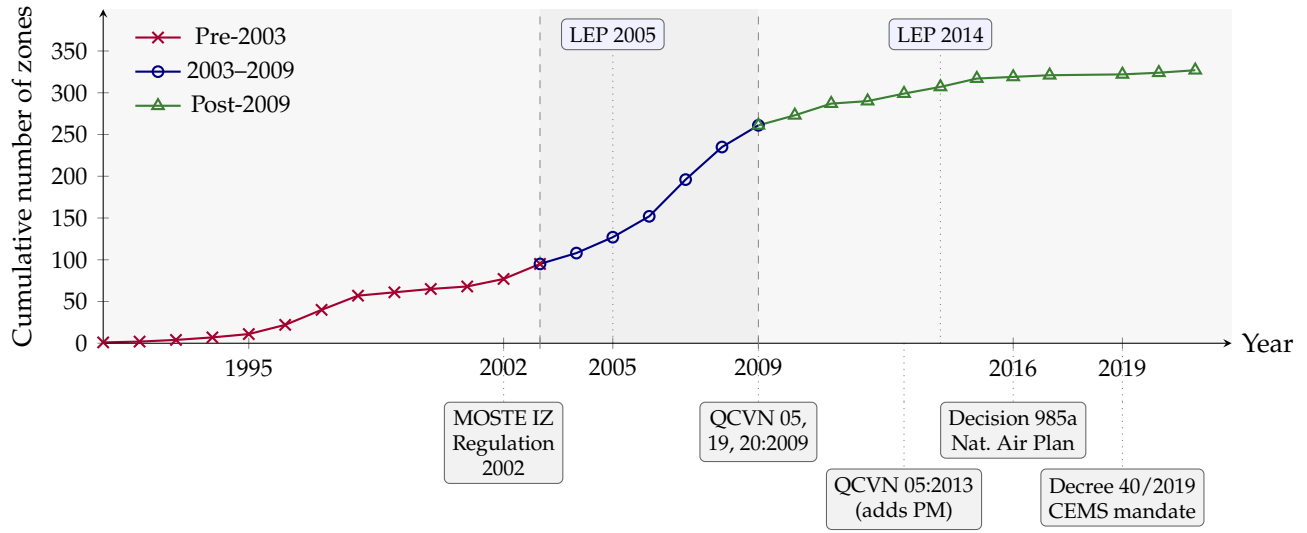
2.1 Industrial Zones and Environmental Regulations

Industrial zones have been a central instrument of Vietnam’s industrial policy since the mid-1990s. The first zones were established following Decree 192/CP of 1994, and by the end of 2020 over 300 zones were in operation across provinces (Figure 1). Zones are bounded areas set aside for manufacturing and related services, approved by the Prime Minister on proposal from the Ministry of Planning and Investment and administered by management boards operating under provincial People’s Committees.¹ Tenants receive publicly built infrastructure together with tax, duty, and land-rent incentives.

The environmental regime governing these zones has evolved through three broad periods,

¹Decree 35/2022/ND-CP on the Management of Industrial Zones and Economic Zones.

Figure 1: Evolution of Industrial Zones and Air Pollution Regulatory Framework in Vietnam



Notes: The figure presents the evolution of Vietnam's air pollution regulatory framework. Laws (Law on Environmental Protection, LEP) are shown above the timeline; technical regulations (QCVN and TCVN standards) and action plans are shown below. The shaded region from 2016 onward marks the period following Prime Minister Decision 985a/QD-TTg, which established a coordinated air quality management regime. MOSTE refers to the Ministry of Science, Technology, and Environment; QCVN denotes National Technical Regulations on air quality and industrial emissions; CEMS denotes continuous emission monitoring systems.

marked by the dashed lines in Figure 1. The first generation of zones, established before 2003, operates under minimal environmental oversight. The Ministry of Science, Technology, and Environment (MOSTE) issues the first regulation specific to industrial zone environmental management in 2002, but monitoring infrastructure and enforcement capacity are limited. The 2003 to 2009 period sees the first comprehensive framework legislation with the Law on Environmental Protection (LEP) of 2005, which formalizes the environmental impact assessment (EIA) process (Decree 80/2006/ND-CP) and is followed by the first national technical regulations on ambient air quality (QCVN 05:2009/BTNMT) and industrial emissions of inorganic dust and organic substances (QCVN 19:2009/BTNMT, QCVN 20:2009/BTNMT). Sector-specific emission caps are established for the highest-emitting industries, including thermal power (QCVN 22:2009/BTNMT), cement (QCVN 23:2009/BTNMT), and steel (QCVN 51:2013/BTNMT). The post-2009 period brings progressively tighter requirements. QCVN 05:2013 adds PM_{2.5} to the ambient air quality standard for the first time, the revised LEP of 2014 strengthens EIA and inspection provisions, and Prime Minister Decision 985a/QD-TTg of June 2016 establishes a coordinated air quality management regime combining emission source control, ambient monitoring, and vehicle and construction dust standards. Decree 40/2019/ND-CP mandates continuous emissions monitoring systems (CEMS) at major industrial sources.

Throughout these three periods, regulation has mainly relied on command-and-control instruments, with market-based instruments playing a minor role. Zone-level environmental infrastructure has been focused on water rather than air: industrial zones are required to install central-

ized wastewater treatment plants before tenant firms can operate, but no equivalent zone-level air treatment requirement exists.² Price instruments during the study period are limited to the Law on Environmental Protection Tax of 2010, which took effect in 2012 and applies upstream to fossil fuels, coal, and selected chemicals rather than directly to emissions at the stack. Vietnam has no operational emissions trading scheme covering industrial air pollutants.³

Most of these instruments apply before a firm operates. Projects in polluting sectors or above size thresholds are required to prepare an EIA, which regulators accept before the project receives an investment certificate. The project design must also meet the sector emission standards, and the heaviest industries need approval under their sector master plans. Once in operation, a firm faces the same standards, enforced through periodic inspections and the continuous emissions monitoring introduced in 2019.

Despite increases in provincial spending on environmental protection per capita and in detected environmental violations (General Statistics Office, 2023), enforcement remains structurally weak through the end of the study period. The physical monitoring infrastructure is sparse and poorly connected: by end-2020, Vietnam operated just 108 ambient air quality monitoring stations nationwide (76 reference-grade) and 522 automatic emissions monitoring systems at polluter facilities (Ministry of Natural Resources and Environment, 2021), yet only 55% of emissions stations and 85% of ambient stations were transmitting data to the Ministry, with eight of 63 provinces not transmitting any emissions data at all.⁴ Even where monitoring systems were installed and transmitting, the data were unreliable. In March 2026, Vietnam’s Investigation Police Agency charged 74 individuals across 59 enterprises with manipulating data from the mandated continuous monitoring systems, with tampering identified at approximately 55% of inspected stations nationwide, including at major thermal power plants and metal producers.⁵ Sparse infrastructure, incomplete data transmission, compromised monitoring integrity, and routine inspections that rely on self-reporting by firms to under-resourced provincial departments together imply that the effective constraint on firm-level air emissions was considerably weaker than the statutory framework alone would suggest. Regulation therefore likely acted mainly on which industries entered a zone rather than on the firms already operating.

2.2 Industrial Composition of Zones

Firms operating inside zones have a distinctive profile. Panel A of Table 1 reports firm and employment composition in the Vietnam Enterprise Survey (VES) 2016 for three groups: firms operating

²This asymmetry is partly technical: wastewater can be aggregated and treated at a common facility, while air emissions must be controlled at each stack.

³A direct fee on industrial exhaust gas emissions was issued only under Decree 153/2024/ND-CP in November 2024, effective January 2025. A domestic carbon market pilot under Decree 06/2022/ND-CP is scheduled for 2025 to 2028.

⁴The 108 ambient stations include 32 lower-cost sensor units. The emissions station count sits alongside 655 automatic wastewater monitoring stations.

⁵<https://vietnamnews.vn/society/1778345/viet-nam-prosecutes-74-over-widespread-manipulation-of-environmental-monitoring-data.html>

inside an industrial zone, firms outside zones but in districts whose territory intersects a 5 km buffer around any zone, and firms outside zones in districts beyond the 5 km buffer. Inside zones, 39% of firms are foreign-invested, 29% employ 100 or more workers, and 62% operate in manufacturing. Manufacturing employment accounts for 95% of total employment inside zones, against 23 to 36% outside. Zones therefore function as concentrated sites of foreign-invested, large-firm, manufacturing activity within a broader Vietnamese economy dominated by small domestic enterprises and services.

Table 1: Firm and Employment Composition, by Proximity to Industrial Zones, Year 2016

Panel A: All Firms, by Proximity				
	Inside Zones		Outside Zones	
			Within 5 km	Beyond 5 km
	(1)		(2)	(3)
Share of foreign-invested firms	0.393		0.018	0.012
Share of firms with 100+ employees	0.292		0.029	0.029
Share of firms in manufacturing	0.619		0.144	0.105
Share of employment in manufacturing	0.946		0.359	0.233

Panel B: Manufacturing Firms, by Pollution-Intensity Classification				
	Inside Zones		Outside Zones	
	Light	Pollution-intensive	Light	Pollution-intensive
	(1)	(2)	(3)	(4)
Share of manufacturing firms	0.810	0.190	0.847	0.153
Share of manufacturing employment	0.905	0.095	0.862	0.138
Share of foreign-invested firms	0.615	0.424	0.035	0.030
Share of firms with 100+ employees	0.446	0.385	0.077	0.094
Mean employment	411	184	53	47
Foreign-invested firms	564	270	516	189
Domestic firms	166	120	36	42

Notes: Panel A reports group means for all firms in the Vietnam Enterprise Survey 2016, by proximity to industrial zones. Inside Zones comprises firms operating inside an industrial zone, as reported by the firm in the VES. Outside Zones (Within 5 km) comprises firms located in districts whose territory intersects a 5 km buffer around an industrial zone centroid but which are not themselves inside a zone. Outside Zones (Beyond 5 km) comprises firms in districts that do not intersect the 5 km buffer. Panel B reports means for manufacturing firms, splitting firms by location (inside a zone or outside) and by the pollution-intensity classification (Appendix B): pollution-intensive sectors are the top quartile of the energy-based intensity ranking constructed from the VES energy module, comprising non-metallic minerals, paper, coke and refined petroleum, basic metals, and textiles; light manufacturing comprises all other manufacturing sectors. The shares of manufacturing firms and employment are computed within each location group. Employment is total labor at the end of the year. Mean employment is reported for all firms and separately for foreign-invested and domestic firms. *Source:* Data from VES 2016.

Classifying manufacturing firms as light or pollution-intensive shows that zones do not screen out the polluting industries: pollution-intensive firms account for 19% of manufacturing firms

inside zones and 15% outside.⁶ Their employment share is lower inside zones than outside, but this reflects firm size rather than composition: light manufacturing in zones is dominated by large foreign-invested export assembly operations, while the pollution-intensive producers are smaller, less often foreign-invested, and concentrated in sectors such as cement and steel. The size gap is concentrated among the foreign-invested firms, whose light-manufacturing operations are more than twice the size of their pollution-intensive counterparts (Table 1, Panel B).

This sorting is consistent with at least three explanations, none mutually exclusive. Investment promotion during the study period (e.g., Decree 24/2000/ND-CP, Investment Law 59/2005/QH11, Decree 108/2006/ND-CP) favored export-oriented production and directed foreign investment toward light manufacturing (Akinci & Crittle, 2008; Athukorala & Tran, 2010). Comparative advantage works in the same direction: foreign investors come to Vietnam largely for labor-intensive export assembly (Hollweg et al., 2017; McCaig & Pavcnik, 2018; Athukorala & Nguyen, 2022), while the pollution-intensive industries sell mostly to the domestic market, where construction demand grew rapidly over the period. Finally, the domestic ownership of the pollution-intensive producers reflects the state-owned legacy of heavy industry, which developed under ministry master plans and state general corporations with foreign participation limited to joint ventures (Nguyen et al., 2016).⁷

3 DATA AND SUMMARY STATISTICS

3.1 Data

To examine the environmental impacts of industrial zones, I combine information on zone locations and opening dates with satellite-derived measures of industrial combustion and air pollution, together with reanalysis wind fields that support the identification strategy. This subsection describes the analysis grid and the construction of the treatment variable and the two outcome variables, along with the underlying data sources. The main study period covers 2000 to 2020.

Spatial unit and administrative geography. The analysis is conducted on a regular grid of 0.05 by 0.05 degree cells covering Vietnam, each approximately 5 by 5 kilometers. Each cell in the 0.05 degree grid contains arrays of finer 0.01 degree sub-cells, which I use for aggregating higher-resolution inputs and for assigning cells to administrative units. Provincial boundary shapefiles come from the Humanitarian Data Exchange.⁸ Because provincial boundaries changed during the study period through splits and mergers, I harmonize administrative units to their 2000s equiv-

⁶The classification assigns each firm by its primary sector. Pollution-intensive sectors are the top quartile of the energy-based intensity ranking constructed from the VES energy module: non-metallic minerals, paper, coke and refined petroleum, basic metals, and textiles. Three of these, non-metallic minerals, basic metals, and coke and refined petroleum, burn fuel in production itself and are referred to below as hot-process industries. Light manufacturing comprises all other manufacturing sectors (Appendix B).

⁷Prime Minister Decisions 90/TTg and 91/TTg (1994) established the framework for the state general corporations, under which VICEM in cement and VNSTEEL in steel were organized. Sector master plans include Decision 108/2005/QD-TTg for cement and Decision 145/2007/QD-TTg for steel.

⁸<https://data.humdata.org/dataset/cod-ab-vnm>

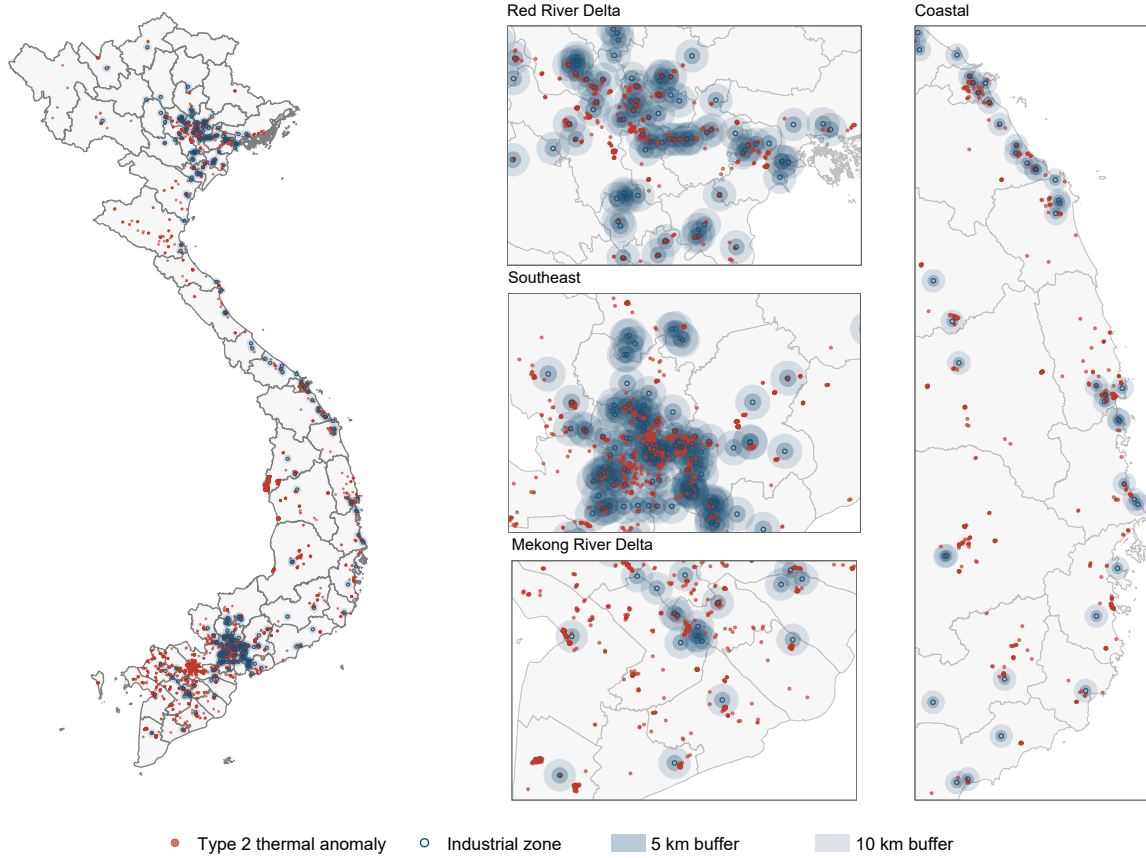
alents using a consistent-identifier mapping. I assign each sub-cell to a province by point-in-polygon and assign the parent cell to the province containing the largest number of its sub-cells, resolving the assignment for cells overlapping multiple provinces. The province is the relevant administrative unit in this setting because environmental regulations during the study period are administered by provincial authorities (Section 2).

Treatment: proximity to an operating industrial zone. A cell is treated when any industrial zone within a given distance of it is in operation, so that treatment varies across cells with their distance to zones and over time as new zones open. Distance to the nearest zone is used only to assign cells to the distance bins of the distance-based analysis. Zone locations come from the Ministry of Planning and Investment database, which records each zone’s name, administrative location, establishment date, and operational status. The database does not report geographic coordinates, and I georeference each zone using Google Maps. The analysis sample comprises zones in operation as of June 2023.

Industrial combustion at the source. The first outcome is the intensity of industrial combustion within a cell, measured from satellite-detected thermal anomalies. Fire detection data come from the NASA FIRMS MODIS Collection 6.1 Active Fire Product (Giglio et al., 2016; NASA FIRMS, 2021). Each detection is recorded with geographic coordinates, date, fire radiative power, and a type classification distinguishing presumed vegetation fires (Type 0), active volcanoes (Type 1), other static land sources (Type 2), and offshore detections (Type 3). Type 2 identifies hotspots detected on sixteen or more days in a calendar year or located in urban areas (Giglio et al., 2021). Because industrial combustion sources such as gas flares, steel mills, and cement kilns tend to operate continuously, this persistence criterion captures most of them. Type 2 has been used in the recent literature as a satellite proxy for local industrial activity (Yang et al., 2024). As seen in Figure 2, Type-2 detections concentrate around the areas with the densest industrial zone footprints in Vietnam, including the Red River Delta in the north, the Southeast and Mekong River Delta in the south, and the central and southern coastal provinces, supporting the interpretation of Type-2 as an industrial proxy for this setting.

The classification does not distinguish industrial from non-industrial persistent thermal sources. In the Vietnamese context, the main non-IZ sources likely captured by Type 2 are informal industrial activities such as brick kilns and rice-husk furnaces, as well as forest clearing in upland areas. The same figure illustrates this: Type-2 detections in the Mekong River Delta and along the interior stretches of the Coastal region extend well beyond the vicinity of industrial zones, reflecting these non-IZ sources rather than IZ-related combustion. Volcanic and offshore detections are negligible in Vietnam and excluded. I aggregate point detections to cell-year counts of Type 2 anomalies and use Type 2 as the measure of industrial activity at the source. As will be shown in Section 4, the estimated increase in combustion is highly concentrated within 5 km of an operational zone. Consequently, my primary on-source specifications define treatment using this 5 km radius, focusing attention on the precise areas where Type-2 detections are most plausibly attributable to

Figure 2: Type-2 Thermal Anomalies and Industrial Zones in Vietnam



Notes: The figure overlays MODIS Collection 6.1 Type-2 thermal anomaly detections (NASA FIRMS, 2021) (pooled over 2004 to 2019) with the locations of industrial zones in operation as of June 2023. Type-2 identifies hotspots detected on sixteen or more days in a calendar year or located in urban areas (Giglio et al., 2021). The left panel shows the national distribution. The four right panels zoom into the Red River Delta, Southeast, Mekong River Delta, and Coastal provinces. Each zone is plotted at its centroid together with 5 and 10 km buffers. Provincial boundaries are harmonized to their 2002 equivalents. *Source:* Industrial zone locations are from the Ministry of Planning and Investment database, geo-referenced using Google Maps.

zone-related activity. Because non-industrial sources of Type-2 detections, such as rice-husk furnaces, are often heavily concentrated in specific regions like the Mekong River Delta, I verify that the estimates are not driven by these localized measurement artifacts by conducting leave-one-out robustness checks that drop one province at a time. Furthermore, the use of cell and year fixed effects helps absorb diffuse, persistent non-industrial combustion across the rest of the sample.

Ambient air pollution. The second outcome is monthly mean surface-level PM_{2.5}, a standard measure of ambient fine particulate matter and the pollutant most commonly used to assess the health and welfare burden of industrial air pollution. Industrial combustion contributes to ambient PM_{2.5} both directly, through primary particulate emissions, and indirectly, through gaseous precursors that form secondary particulates in the atmosphere (Hallquist et al., 2009; Huang et al., 2014; Wang et al., 2016). PM_{2.5} data are drawn from the V6.GL.02.04 satellite-derived product of Shen et al. (2024), which combines aerosol optical depth retrievals from multiple satellite platforms

with GEOS-Chem chemical transport modeling and calibrates the resulting geophysical estimate to ground-based monitors. The product reports monthly surface-level concentrations at 0.01 degree native resolution, approximately one kilometer over Vietnam. I aggregate the data to the 0.05 degree analysis grid by taking cell-level means. Because airborne pollutants are transported by wind, identifying the effect of zone establishment on PM_{2.5} requires accounting for atmospheric circulation. Section 5 describes the wind-based exposure measure used in the empirical analysis.

Other data. The PM_{2.5} specifications require controls for meteorological conditions that independently affect ambient concentrations. Precipitation removes suspended particles through wet deposition, and temperature governs boundary-layer mixing and the rate at which gaseous precursors convert into secondary aerosol. I control for both using monthly mean 2-meter temperature (entered as a quadratic) and monthly total precipitation from ERA5-Land (Muñoz-Sabater et al., 2021), which shares the reanalysis platform used for the wind fields. Translating the cell-level PM_{2.5} estimates into population-level health burdens requires spatially resolved population counts, and I use the Gridded Population of the World, Version 4 (Center for International Earth Science Information Network (CIESIN), Columbia University, 2017), which reports year-2000 population at approximately 1 km resolution. Cell-level elevation from the Shuttle Radar Topography Mission (Farr et al., 2007) enters the covariate balance tests. District-level baseline characteristics used in these balance tests and as controls interacted with time fixed effects are constructed from the 2002 Vietnam Household Living Standards Survey (VHLSS). Finally, the analysis of industrial composition and mechanisms draws on the Vietnam Enterprise Survey (VES). I use the full VES panel to characterize the sectoral mix of zone activity and to estimate changes in the number of manufacturing firms operating near zones. The firm-level analysis of fuel use relies on physical quantities of nine fuel types, which are reported only in the 2009 and 2011–2015 rounds of the VES. I use the 2009, 2011, and 2013–2015 rounds, excluding 2012 because of inconsistencies in its energy records, and convert the quantities to tons of coal equivalent using default net calorific values from Garg et al. (2006) and volume-to-mass conversion factors from United Nations Department of Economic and Social Affairs (2022). The same energy records underpin the pollution-intensity classification used throughout the paper, constructed as fossil-fuel energy per unit of output at the industry level; Appendix B details the construction and compares the resulting ranking with the U.S.-based classification of Shapiro and Walker (2018).

3.2 Summary Statistics

Table 2 reports district-level characteristics in 2002 to illustrate that areas with industrial zones differ systematically from areas without. Although the analysis itself is conducted at the grid-cell level, I aggregate to districts here because the household-survey variables that capture local economic conditions are constructed at the district level. Districts are classified into three groups by their exposure history to industrial zones during 2000–2020: never exposed within 5 km of an operational zone (column 1), first exposed via a zone established during 2003–2020 (column 2), and exposed via a zone established in or before 2003 (column 3). For columns (1) and (2), 2002

values are pre-treatment. For column (3), they reflect both pre-existing characteristics and changes since the zone became operational, and these districts are excluded from the cell-level event-study estimation.

Table 2: District-Level Characteristics in 2002, by Zone Establishment Cohort

	No Zones	Zones	Zones	p-values		
		2003–2020	before 2003	(1) = (2)	(1) = (3)	(2) = (3)
	(1)	(2)	(3)	(4)	(5)	(6)
Urban population share	0.164 [0.259]	0.219 [0.286]	0.344 [0.397]	0.045	0.000	0.002
Minority population share	0.364 [0.405]	0.090 [0.190]	0.035 [0.112]	0.000	0.000	0.003
Share working-age with secondary degree	0.146 [0.138]	0.190 [0.135]	0.268 [0.162]	0.001	0.000	0.000
Share working-age in farming	0.720 [0.248]	0.578 [0.229]	0.421 [0.270]	0.000	0.000	0.000
Rice specialization	0.432 [0.247]	0.443 [0.246]	0.391 [0.302]	0.641	0.149	0.102
Type-2 fires (count)	0.078 [0.695]	0.150 [0.604]	0.255 [0.814]	0.259	0.028	0.217
PM2.5 ($\mu\text{g}/\text{m}^3$)	21.164 [4.250]	22.186 [4.183]	22.000 [4.228]	0.015	0.057	0.708
Number of Districts	361	154	146			

Notes: The table reports district-level summary statistics. The sample includes 661 districts, classified by exposure history to industrial zones (IZs) during 2000–2020, where a district is classified as exposed in a given year if its centroid lies within 5 km of an operational IZ. Treatment timing is set to one year before recorded operational status to absorb site-preparation activity. Column (1) reports districts never exposed during 2000–2020. Column (2) reports districts whose first exposure corresponds to an IZ established during 2003–2020. Column (3) reports always-exposed districts, defined as those exposed via IZs established in or before 2003. Each cell reports the group mean, with the standard deviation in brackets. Columns (4) through (6) report p-values from t-tests of equal means between the indicated group pairs, with heteroskedasticity-robust standard errors. *Source:* Urban population share, the share of working-age members with a secondary degree, the share of working-age members in farming, and rice specialization are constructed from the 2002 Vietnam Household Living Standards Survey. Type-2 fire counts are the sum of MODIS Type 2 (static land source) detections in 2002 (NASA FIRMS, 2021), the detection type used for the combustion outcome in the analysis. PM2.5 is the annual mean surface concentration in 2002 from Shen et al. (2024). Type-2 fire counts and PM2.5 are aggregated from cell-level data to the district level.

The three groups differ along nearly every economic dimension. Moving from never-exposed to newly-exposed and early-exposed districts, the urban population share rises from 16% to 22% and 34%, while the share of working-age members with a secondary degree rises from 15% to 19% and 27%. The share in farming falls from 72% to 58% and 42%, and the minority population share from 36% to 9% and 4%. These patterns indicate that zones were placed in districts further along in structural transformation. Type-2 fire counts, the industrial combustion measure used in the analysis, follow the same ordering, 0.08, 0.15, and 0.26 in columns (1) to (3), although only the difference between columns (1) and (3) is statistically significant. For column (3), the count also reflects the combustion of the zones already operating in 2002. Rice specialization is the only dimension on which none of the three pairwise comparisons is statistically significant, suggesting

that selection operated through urbanization and broader sectoral composition rather than agricultural specialization. The remaining pairwise differences are statistically significant at conventional levels, except for PM2.5 between columns (2) and (3). I address this non-random placement in the cell-level analysis through the identification strategies described in the next section.

4 EMPIRICAL ANALYSIS: INDUSTRIAL COMBUSTION AT THE SOURCE

Type-2 thermal anomalies, by construction, register only when high-temperature combustion occurs at the surface within a sensor pixel, and the satellite signal is spatially confined to the cell containing the heat source. Cell-level distance from operational industrial zones therefore provides a direct test of whether zone establishment generates detectable combustion at the source: if so, the effect should be concentrated in cells closest to the zone boundary and attenuate with distance. Because zones became operational over a 25-year period (Figure 1), a standard two-way fixed-effects regression pooling all cohorts may produce biased estimates when treatment effects vary across cohorts and over time (de Chaisemartin & d’Haultfoeuille, 2020; Goodman-Bacon, 2021; Borusyak et al., 2024). I address this in two complementary specifications. The first estimates distance-bin coefficients separately by cohort and aggregates them to characterize the spatial reach of any effect. The second uses the distance bin at which the effect is concentrated to define a binary treatment indicator, restricts the comparison group to never-treated cells, and examines the effect’s evolution over time.

4.1 Distance-based Analysis

I group zone cohorts into two-year bins to ensure adequate cell counts within each cohort regression. Cohort c comprises cells whose nearest industrial zone first becomes operational in year c or $c + 1$, where c is set to one year before recorded operational status to capture site-preparation activity. For each cohort c , I estimate:

$$y_{ipt} = \sum_{b=1}^9 \delta_b^c \cdot \mathbb{I}(t \geq c) \cdot \mathbb{I}(\text{bin}_i = b) + \alpha_i + \mu_t + \varepsilon_{ipt} \quad (1)$$

where y_{ipt} is a log-transformed measure of Type-2 thermal anomaly activity in cell i in province p in year t ; α_i is a cell fixed effect; μ_t is a year fixed effect; and $\mathbb{I}(\text{bin}_i = b)$ indicates cell i ’s static distance bin from its nearest zone in 5 km increments, from $b = 1$ (0–5 km) through $b = 9$ (40–45 km). The omitted category is $b = 10$ (45–50 km). The sample for each cohort regression includes treated cells (those first entering the 0–5 km ring in year c or $c + 1$) together with all cells that are never within 5 km of any operational zone during the study period. These never-treated cells are distributed across distance bins 2 through 10 based on proximity to the nearest zone, so the post-treatment indicator $\mathbb{I}(t \geq c)$ takes value one for all cells after the cohort year. Each coefficient δ_b^c therefore measures the change in fire activity in bin b after zone establishment, relative to the

change in the 45–50 km reference bin. Estimating equation (1) separately by cohort, with treatment timing aligned within each cohort, addresses the cohort-pooling problems that arise in standard two-way fixed-effects estimators under staggered adoption with heterogeneous treatment effects (de Chaisemartin & d’Haultfoeuille, 2020; Goodman-Bacon, 2021; Borusyak et al., 2024).

To obtain an overall estimate, I aggregate the cohort-specific coefficients using a weighted average:

$$\bar{\delta}_b = \frac{\sum_c n_c \cdot \hat{\delta}_b^c}{\sum_c n_c} \quad (2)$$

where n_c is the number of treated cells in cohort c . I report two sets of standard errors. The first clusters at the province level within each cohort regression and treats the cohort estimates as independent. Because the cohort-specific regressions share a common pool of never-treated cells, this approach understates uncertainty. I therefore also report standard errors from a stratified cluster bootstrap that resamples provinces with replacement across all cohort regressions jointly, preserving both within-province correlation and the cross-cohort dependence induced by the shared comparison group.⁹

4.2 Event-Study Specification

Let $\bar{d}$ denote the distance threshold within which the distance-based effect is concentrated, identified from the analysis above. I define treatment as cell i being within $\bar{d}$ of an operational zone, with T_i denoting the year prior to the first recorded year of operational status. The pre-operation timing definition captures site-preparation activity (land clearing, construction, infrastructure installation) that typically precedes the official operational date by approximately one year. I restrict the sample to never-treated cells and cells that first become treated during the analysis window, excluding always-treated cells (those within $\bar{d}$ before the panel begins), and estimate the event-study:

$$y_{ipt} = \sum_{\ell=-5, \ell \neq -1}^{L=9} \beta_\ell \cdot \mathbb{I}(t - T_i = \ell) + \alpha_i + \mu_t + \varepsilon_{ipt} \quad (3)$$

where $\mathbb{I}(t - T_i = \ell)$ is an indicator for event time ℓ relative to first treatment and all other terms are defined as in equation (1). The omitted period is $\ell = -1$, so coefficients β_ℓ measure the change in outcome relative to the year before treatment. I estimate equation (3) by restricting the comparison group to cells never within $\bar{d}$ of an operational zone during the study period, and report province-clustered standard errors.

Identification assumptions. For the coefficients in equations (1) and (3) to be interpreted as causal effects of industrial zone establishment, the following assumptions must hold.

First, the *parallel trends* assumption requires that, conditional on the fixed effects, Type-2 ther-

⁹In each of 200 replications, I resample provinces with replacement within strata defined by cohort, re-estimate equation (1) separately by cohort, and recompute $\bar{\delta}_b$. Cohorts 2011, 2013, and 2015 are pooled with cohort 2009 into a single resampling stratum because of small cell counts. The bootstrap standard error is the standard deviation of $\bar{\delta}_b$ across replications.

mal anomaly activity in cells near operational zones would have evolved similarly to that in cells farther from any zone in the absence of treatment. For the distance-based specification, this requires parallel evolution within each cohort between cells in any treated bin and cells in the 45–50 km reference bin. For the dynamic specification, it requires parallel evolution between treated and never-treated cells. Although not directly testable, the assumption can be partially assessed through pre-treatment coefficients. The pre-treatment estimates in the dynamic specification (Figure 3, Panel B) are small in magnitude and statistically indistinguishable from zero. To assess sensitivity to bounded violations, I follow Rambachan and Roth (2023) and construct confidence sets under a smoothness restriction by bounding how much the differential trend’s slope can change between consecutive periods.

Second, the *no-anticipation* assumption requires that outcomes in cells approaching treatment are not affected by the future arrival of a zone. The relevant concern is combustion associated with site preparation: land clearing, construction, and access road development could generate detectable thermal anomalies before the zone becomes operational. Treatment is defined to activate in the year prior to recorded operational status, which absorbs site-preparation activity occurring in the immediate pre-operation year. The identifying no-anticipation assumption therefore requires that site preparation does not materially begin two years prior to official operation, preserving the omitted baseline ($\ell = -1$) as a clean representation of the untreated state. I evaluate this by testing for earlier anticipation across event times $\ell \leq -2$. Pre-treatment coefficients in this range are small in magnitude and statistically indistinguishable from zero, consistent with a clean baseline and no detectable anticipation beyond the year already absorbed into the treatment definition (Figure 3, Panel B).

Third, the *stable unit treatment value assumption* (SUTVA) requires that treatment of one cell does not affect outcomes in other cells. With treatment defined as being within 5 km of an operational zone, the concern is that industrial activity in the 0–5 km ring could induce or displace combustion activity in nearby comparison cells, contaminating the comparison group. The distance-based estimates in Panel A of Figure 3 provide a first check. The effect is concentrated in the 0–5 km bin, and the coefficients for the bins beyond 10 km are small and indistinguishable from zero. This comparison, however, only detects displacement that differs across bins. If the treated ring drew activity evenly from all comparison bins, every bin (including the reference bin) would fall by the same amount, and the pattern across bins would look no different. This scenario is unlikely to matter for two reasons. First, even if the entire increase in the treated ring had come from the comparison cells, the estimate would overstate the effect by about a tenth, because the treated ring holds one cell for every 22 comparison cells.¹⁰ Second, the manufacturing firms that appear in newly treated districts are almost all new to the survey rather than relocated from other districts (Appendix Table A1).¹¹ However, relocation from beyond 50 km of a zone cannot be detected

¹⁰Under uniform displacement the estimate equals the true effect multiplied by $1 + (N_T/N_C)(\bar{y}_T/\bar{y}_C)$, where N denotes the number of treated and comparison cells and $\bar{y}$ their mean pre-treatment fire counts. In the estimation sample $N_T/N_C = 312/6,819$ and $\bar{y}_T/\bar{y}_C = 0.059/0.023$. The proportional estimate would therefore be overstated by about 12%.

¹¹The enterprise tax code in the VES is constant across rounds, which allows me to track firms across districts from

with this design, and I return to it in Section 7. Finally, the 5–10 km bin shows a small positive coefficient, suggesting that some combustion activity extends just beyond the treatment ring. As a robustness check, I estimate a donut specification that excludes this bin from the sample, and the results are similar (Appendix Table A2, Column 2).

Fourth, under staggered treatment with heterogeneous effects, the TWFE estimator can place negative weights on some cohort-specific treatment effects (de Chaisemartin & d’Haultfoeuille, 2020; Goodman-Bacon, 2021; Roth et al., 2023; Borusyak et al., 2024). The distance specification is estimated cohort by cohort and is not exposed to this problem. For the event study in equation (3), the heterogeneity-robust estimators of Callaway and Sant’Anna (2021) and de Chaisemartin and d’Haultfoeuille (2024) give nearly the same event-time path and post-treatment average as TWFE (Appendix Figure A1).

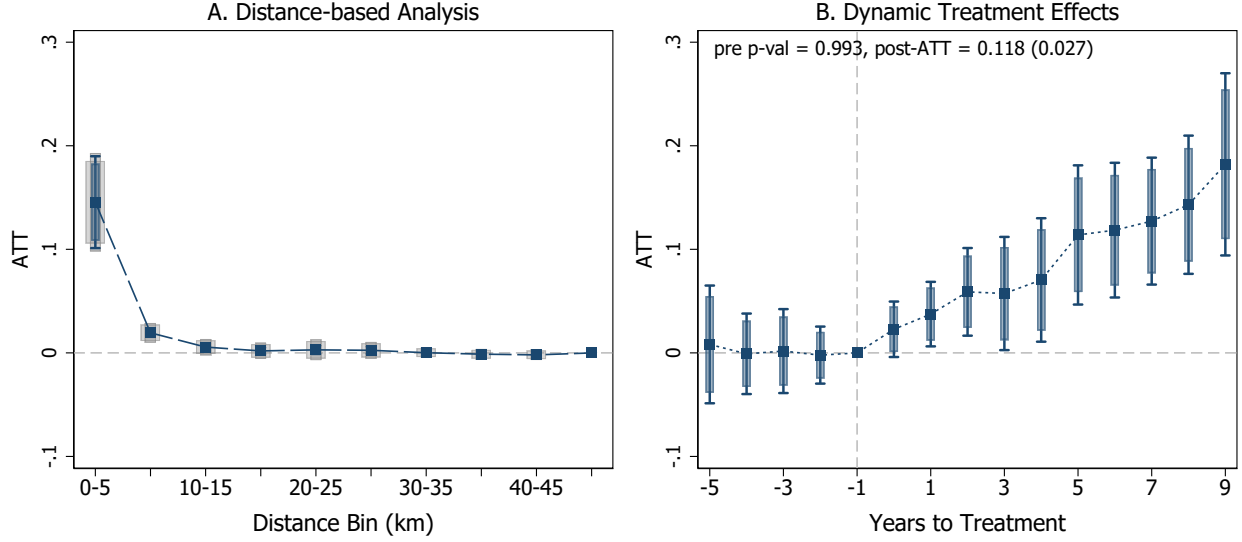
4.3 Results

Figure 3 presents the two sets of estimates, where the outcome is $\ln(1 + \text{Type-2 anomaly})$. The left panel reports the weighted-average distance-bin coefficients $\bar{\delta}_b$ from equation (1). Type-2 thermal anomaly counts in cells within 5 km of an operational zone are approximately 0.14 log points higher than in the 45–50 km reference cells following zone establishment. The effect attenuates sharply with distance. The coefficient for the 5–10 km bin is approximately 0.02 log points ($p < 0.05$ with both analytical province-clustered and bootstrap standard errors), an order of magnitude smaller. All remaining bins between 10 and 45 km are small, precisely estimated, and indistinguishable from zero, with an average coefficient of 0.001 (bootstrap s.e. 0.003) across the seven bins. This pattern is consistent with industrial combustion as a spatially concentrated, point-source phenomenon and differs markedly from the gradual decay observed for dispersed pollutants such as PM_{2.5} (Section 5).

The right panel reports the dynamic effects $\hat{\beta}_\ell$ from estimating equation (3), using the 0–5 km binary treatment indicator. Pre-treatment coefficients are small and statistically indistinguishable from zero ($p = 0.99$ for the joint test), supporting the parallel trends and no-anticipation assumptions. Following zone establishment, Type-2 thermal anomaly counts increase progressively rather than shifting discretely: the effect is close to zero in the year of operation, reaches approximately 0.1 log points by the fifth year, and 0.17 log points by year nine. The average post-treatment effect, computed as a weighted average of the event-time coefficients, with weights proportional to the number of treated cells contributing to each event time, is 0.12 log points ($p < 0.01$). This trajectory is consistent with industrial zones filling in with tenants over time, so that combustion activity

2001 to 2018. Of the manufacturing firms arriving in newly treated districts after zone establishment, 9% of pollution-intensive and 5% of light manufacturing firms had been observed in another district before, and fewer than 2% came from comparison districts within 50 km. The VES enumerates all state and foreign-invested enterprises and all private enterprises with 10 or more workers in every round, while smaller private enterprises are covered on a sample basis in some rounds. Among firms with 10 or more workers, for which absence from earlier rounds means that the enterprise did not yet exist, the shares are 14% and 7%, with fewer than 3% from comparison districts within 50 km, no higher than among arrivals in districts never near a zone. These shares are lower bounds, because a firm that closes in one district and re-registers under a new tax code in another is counted as a new firm rather than a mover.

Figure 3: Industrial Zones and Type-2 Thermal Anomalies: Linear Estimates



Notes: The left panel plots the weighted-average distance-bin coefficients δ_b from equation (1), with 0–5 km through 40–45 km bins and 45–50 km as the reference. The right panel plots the dynamic treatment effects β_ℓ using a binary indicator for whether a cell is within 5 km of an operational industrial zone, with event time ℓ relative to first treatment. Outcome is $\ln(1 + \text{Type-2 anomaly})$ at the cell-year level. Navy markers are point estimates, bars 90% and caps 95% confidence intervals from analytical province-clustered standard errors. Light shaded bands in the left panel are 90 and 95% confidence intervals from a stratified cluster bootstrap with 200 replications. *Source:* Data are from the NASA FIRMS MODIS Collection 6.1 Active Fire Product, 2000–2020.

within the 0–5 km ring accumulates as the zone matures.¹²

Interpretation and a count-data alternative. The estimates above use $\ln(1 + \text{Type-2 anomaly})$ as the outcome. Chen and Roth (2024) show that this transformation does not deliver a percentage-change interpretation when the outcome contains many zeros: the estimated coefficient depends on the units of the outcome and on the constant added inside the log, which can both be problematic when zeros are common. Indeed, the Type-2 detection data have substantial mass at zero, with the share of zero cell-years rising from 84% in the 0–5 km ring to 97% in the 5–10 km ring and 99% or more in all bins beyond 10 km. The coefficients in Figure 3 should therefore be read as evidence that detections are higher near operational zones, not as estimates of a percentage increase.

To obtain estimates with a proper percentage-change interpretation, I re-estimate the dynamic specification (equation 3) using Poisson pseudo-maximum likelihood, which is unit-invariant, accommodates zeros without an ad hoc transformation, and remains consistent under arbitrary distributional misspecification provided the conditional mean is correctly specified (Silva & Tenreiro, 2006; Chen & Roth, 2024). The Poisson specification mirrors equation (3) with the same set of fixed effects, but models the conditional mean of fire counts directly as $E[y_{ipt} | \cdot] = \exp(\cdot)$ rather than

¹²The effects remain stable when time fixed effects are interacted with district-level baseline characteristics, allowing cells in districts with different pre-treatment compositions to follow flexible trends (Appendix Figure A2).

log-transforming the outcome.¹³

Table 3 reports estimates of the average effect (post-ATT) of industrial zones on Type-2 thermal anomalies along three dimensions of combustion activity. The post-ATT is computed as the weighted average of post-treatment event-time coefficients (marginal effects), with weights proportional to the number of treated cell-years contributing to each event-time estimate. Column (1) reports a logit model on $\mathbb{I}(\text{TAP})$, an indicator for any Type-2 detection within the cell-year. Columns (2)–(3) use the count of detections, and columns (4)–(5) use the number of distinct days with any detection. All specifications include cell and year fixed effects; columns (3) and (5) add linear interactions of year fixed effects with district baseline characteristics measured in 2002 from the VHLSS.

Table 3: Industrial Zones and Type-2 Thermal Anomalies: Poisson Estimates

	$\mathbb{I}(\text{TAP})$	Type-2 fires (count)		Type-2 fires (days)	
	(1)	(2)	(3)	(4)	(5)
Post-ATT	1.689 (0.332)	1.040 (0.428)	0.837 (0.265)	0.870 (0.347)	0.766 (0.259)
Post-ATT (marginal effect)	0.395 (0.073)	0.566 (0.233)	0.456 (0.144)	0.422 (0.169)	0.372 (0.126)
Pre-treatment mean	0.106	0.233	0.233	0.229	0.229
pre-trend p-value	0.514	0.701	0.820	0.812	0.991
Observations	11286	11362	11362	11362	11362
Cell FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Year FE $\times$ Baseline Controls			Yes		Yes
Model	Logit	Poisson	Poisson	Poisson	Poisson

Notes: The table reports the effects of industrial zone establishment on Type-2 thermal anomalies, estimated on a sample of cells within 5 km of an operational industrial zone (treated) and cells never within 5 km of any operational zone during the panel (controls); always-treated cells are excluded. Column (1) is a logit regression where the outcome is an indicator equal to one if any Type-2 detection occurs in the cell-year. Columns (2)–(5) are Poisson PML regressions; columns (2)–(3) use the count of Type-2 detections, and columns (4)–(5) use the number of distinct days with any Type-2 detection. The treatment in all columns is a binary indicator for whether the cell is within 5 km of an operational industrial zone. Columns (3) and (5) add interactions of the year fixed effects with four district characteristics measured in the 2002 VHLSS: the urban population share, the ethnic minority share, the share of working-age members with a secondary degree, and the share of working-age members in farming. The post-ATT is the weighted average of post-treatment event-time coefficients (event times 0 through 9), with weights proportional to the number of treated cell-years contributing to each event-time estimate. Marginal effects in column (1) are on the probability scale; marginal effects in columns (2)–(5) are on the count or days scale. The pre-trend p-value is from a joint test that all pre-period coefficients (event times -5 through -2) equal zero. Standard errors are clustered at the province level for columns (2)–(5); column (1) reports unclustered standard errors from `xtlogit, fe`. *Source:* Data are from the NASA FIRMS MODIS Collection 6.1 Active Fire Product, 2000–2020.

Column (1) shows that the probability of any Type-2 detection rises by 39.5 percentage points ($p < 0.01$) following zone establishment, off a pre-treatment baseline of 10.6%. The preferred Poisson specifications yield post-ATT coefficients of 1.04 for detection counts (column 2) and 0.87 for detection-days (column 4), both significant at the 1% level. Exponentiating these coefficients, zone

¹³Estimated using the `ppmlhdfe` Stata package (Correia et al., 2020), which handles high-dimensional fixed effects and drops cells whose outcome is always zero throughout the panel, as their fixed effect is not identified.

establishment multiplies the conditional mean of detections by approximately 2.8 and the conditional mean of detection-days by approximately 2.4. In level terms, the corresponding marginal effects imply increases of 0.57 detections and 0.42 detection-days per cell-year, against pre-treatment means of 0.23 for both measures. The similarity of the count and days marginal effects indicates that the response is concentrated on the extensive margin: zone establishment primarily raises the probability that a cell registers any combustion on a given day, rather than generating multiple detections on already-active days.¹⁴ Pre-trend tests yield p-values between 0.51 and 0.99 across specifications, providing no evidence against parallel trends (Appendix Figure A3).

Robustness. The results are robust to a series of checks. Allowing cells in districts with different pre-treatment characteristics to follow their own time paths leaves the effects modestly smaller (columns (3) and (5) of Table 3). Excluding the 5–10 km ring from the comparison group, dating treatment from the year of recorded operation rather than the year before, and expanding the treatment ring to 10 km give similar estimates, and pre-trend tests do not reject parallel trends in any of them (Appendix Table A2). Because 14% of treated cell-years lie within 5 km of two or more operating zones, Appendix Table A3 restricts treated cells to those with a single zone within 5 km and comparison cells to those with no zone within 10 or 25 km. The estimates are similar, and a second zone within 5 km adds nothing detectable to the first. Dropping one province at a time leaves every estimate statistically significant (Appendix Figure A4), and the post-treatment effect remains positive and statistically significant at the 10% level under deviations from linear pre-trends up to $M \approx 0.005$ (Rambachan & Roth, 2023) (Appendix Figure A5).

The analysis so far establishes that industrial zone establishment generates a localized increase in industrial combustion within 5 km of the zone centroids. This result serves two purposes in the analysis that follows. First, it identifies zones as combustion sources and dates their activation. The 5 km radius encompasses zone-registered firms and the supporting activity that clusters near the zone boundary, and the PM2.5 analysis in Section 5 treats this ring as the source area from which emissions are transported downwind. Second, it measures how combustion at the source evolves in aggregate: it accumulates as zones fill. Section 6.1 relates this accumulation to the manufacturing activity that zones attract.

The thermal data do not, however, measure how much fuel an individual zone burns, because the MODIS algorithm registers a fire only when its radiant signal stands out from the surrounding pixels on a cloud-free overpass. At the 1 km resolution of the sensor, this requires a source of the size and temperature of a gas flare, a furnace, or a kiln (Giglio et al., 2016). Most fuel burned in a zone does not meet this condition. The hot-process industries that such sources belong to account for 9% of the manufacturing firms inside zones and 3% of their employment in 2016 (Appendix Table D1, Panel A); the remaining firms, in textiles, food, paper, electronics, and assembly, burn coal and oil for steam and process heat at temperatures well below the detection threshold. As a

¹⁴The ratio of the count marginal effect to the days marginal effect, $0.57/0.42 \approx 1.3$, gives the average number of detections per newly active day. A ratio close to one indicates that additional combustion takes the form of new days with a single detection rather than multiple detections on the same day.

result, a zone with no detections is not necessarily a zone without combustion, and the number of detections at a zone indicates the presence of high-temperature sources rather than the quantity of fuel burned. The exposure measure in Section 5 accordingly counts every operating zone equally, and Appendix E shows that the per-exposure effect is unrelated to the thermal activity observed at the source. The welfare-relevant question is whether the combustion that zones generate translates into ambient pollution exposure for populations living beyond the immediate source area.

5 EMPIRICAL ANALYSIS: AMBIENT AIR POLLUTION

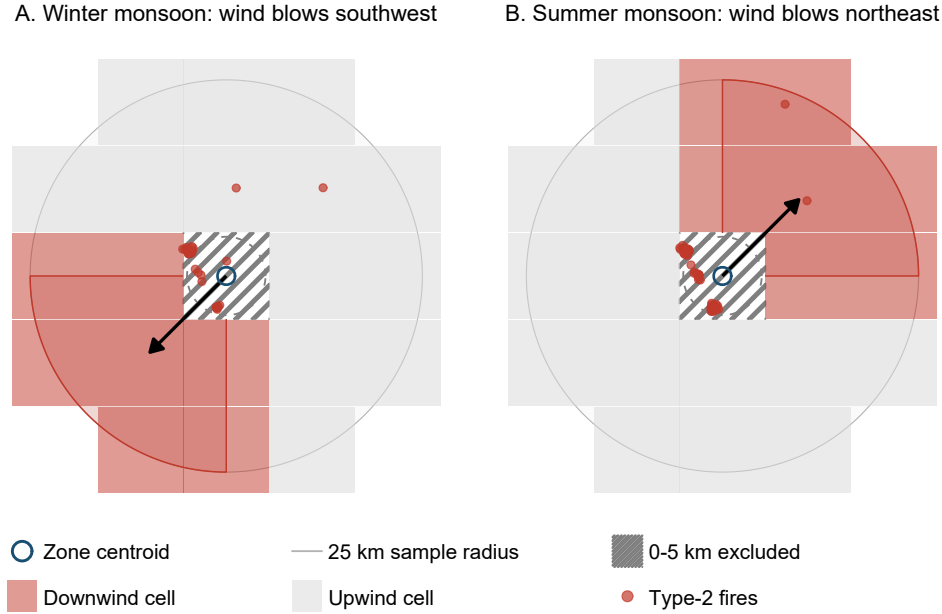
Particulate matter, unlike thermal anomalies confined to the source pixel, disperses with prevailing winds and accumulates at some distance from where it is emitted. Industrial activity contributes to ambient PM_{2.5} through both direct emissions of primary particles and emissions of gas-phase precursors that form secondary aerosol over hours to days as the plume travels downwind (Hallquist et al., 2009; Huang et al., 2014). These mechanisms generate documented PM_{2.5} spillovers across district, regional, and national boundaries (e.g., Jia & Ku, 2019). The PM_{2.5} signal recorded in a given cell therefore reflects emissions from sources that may lie kilometers away in the upwind direction, and distance from an industrial zone is, by itself, a weak proxy for exposure to zone-related pollution. Wind direction breaks the spatial symmetry: if a zone emits meaningful quantities of fine particulates or precursor gases that transform into fine particulates downwind, cells lying downwind would record higher PM_{2.5} than cells lying upwind of the same zone. A growing literature exploits this asymmetry to identify the effects of spatially fixed pollution sources, including aggregate regional emissions (Deryugina et al., 2019) and highways (Anderson, 2020), on a range of outcomes. I adopt this approach in the industrial-zone setting by estimating the average effect on cell-level PM_{2.5} of downwind exposure to nearby zones, thereby recovering the magnitude of pollution attributable to industrial-zone activity.

To do so, I construct a cell-zone-month downwind measure from ERA5-Land reanalysis wind fields (Copernicus Climate Change Service (C3S), 2019; Muñoz-Sabater et al., 2021), using the 10-meter eastward and northward wind components at 0.1 degree resolution. For each calendar day I retain a single snapshot at 06:00 UTC, corresponding to 13:00 local time in Vietnam, to capture wind conditions during the daytime boundary layer mixing regime when industrial emissions and plume dispersion are most active. For each cell-zone pair within 25 km, I compare the direction from the zone to the cell with the direction in which the wind is blowing at the zone on each day. A day is classified as downwind if the two directions are within 45 degrees of each other, and the monthly downwind share for a cell-zone pair is the share of days in the month so classified. Figure 4 illustrates the classification for a single zone.¹⁵ A cell's exposure measure aggregates this classification across all operational zones within 25 km, weighting each zone's monthly downwind

¹⁵ Although wind direction is given at 0.1° resolution (one value per ERA5 cell), the analysis is conducted on a 0.05° cell grid, so the four 0.05° cells inside a single 0.1° wind cell can receive different classifications when the wedge edge from the zone passes through them. Appendix Figure A6 illustrates this construction at the boundary between the two resolutions.

share by inverse distance. Cells within 5 km of the nearest zone are excluded from the estimation sample because direct emissions dominate exposure regardless of wind direction. This source area, as documented by the Type-2 thermal anomaly results in Section 4, captures emissions from both within-zone manufacturing and proximate economic activity induced by zone establishment.

Figure 4: Wind-based Classification of Cell Exposure to an Industrial Zone



Notes: The figure illustrates the wind-based classification of grid cells around an industrial zone, using a representative zone in the Mekong River Delta as a concrete example. The zone centroid is shown at the center of each panel. The outer circle marks the 25 km sample radius around the zone. The dashed inner circle at 5 km marks the source area, which is excluded from the analysis sample because direct emissions dominate exposure regardless of wind direction. Cells in the 5-to-25 km transport ring are shaded red if the bearing from the zone to the cell falls within $\pm 45^\circ$ of the wind direction at the zone (downwind) and grey otherwise. Cells in the 0-to-5 km source area are shown with diagonal hatching to indicate exclusion. The shaded red wedge depicts this 90° sector at the analysis boundary. Red dots show MODIS Collection 6.1 Type-2 thermal anomaly detections within 25 km of the zone during the illustrated season. Panel A shows the winter monsoon (December to February), when the northeast monsoon carries emissions toward the southwest. Panel B shows the summer monsoon (June to August), when the monsoon reverses. The same cells are classified as downwind in one season and upwind in the other, which is the identifying variation exploited by the wind-based exposure measure. The empirical measure uses daily ERA5-Land wind fields aggregated to the monthly cell-zone downwind share. *Source:* Fire detection data from NASA FIRMS (2021), wind data from Muñoz-Sabater et al. (2021), and industrial zone locations from the Ministry of Planning and Investment database.

5.1 Empirical Strategy

To examine the effect of industrial zones on PM_{2.5}, I estimate the following equation:

$$\ln \text{PM}_{2.5_{ipmt}} = \beta E_{imt} + \alpha_i + \mu_{mt} + \varepsilon_{ipmt} \quad (4)$$

where $\ln \text{PM2.5}_{imt}$ is the log of monthly mean PM2.5 concentration in cell i during month m of year t , and

$$E_{imt} = \sum_{j \in \mathcal{J}_i} \mathbb{I}(t \geq T_j) \cdot \text{distance}_{ij}^{-1} \cdot D_{ijmt} \quad (5)$$

is the cell's downwind exposure measure. The set $\mathcal{J}_i$ comprises industrial zones within 25 km of cell i . For each zone j , T_j is the year prior to its recorded operational status, defined as in equation (3) to absorb site-preparation activity that precedes the official operational date. The variable distance_{ij} is the distance in kilometers from cell i to zone j , and D_{ijmt} is the monthly downwind share for the cell-zone pair defined in Figure 4. The terms α_i and μ_{mt} denote cell and year-month fixed effects, respectively.

The key difference from the fires specification is that the PM2.5 analysis uses monthly data. Year-month fixed effects therefore absorb seasonal pollution patterns and aggregate shocks common to all cells. Cell fixed effects absorb, among other time-invariant characteristics, each cell's long-run prevailing wind climatology. Identification comes from within-cell deviations in monthly wind direction rather than from persistent differences in wind exposure across cells. Standard errors are clustered at the province level p .

The coefficient β captures the percentage change in PM2.5 associated with a one-unit increase in downwind exposure, identified from within-cell variation in E_{imt} that remains after absorbing the fixed effects. This variation has two sources: the opening of new zones within 25 km of a cell, which raises the cell's exposure once a zone becomes operational; and the seasonal reversal of monsoon winds, which alternates the position of a given cell between the downwind and upwind sectors of nearby zones across months (Figure 4). The coefficient therefore captures the reduced-form effect of industrial zone establishment on downwind air quality, operating through all emission channels that the policy activates within the source area defined in Section 4. Emissions from supporting economic activity induced by zone establishment are part of the causal effect of the policy. The event-study evidence below supports this interpretation: the downwind PM2.5 effect is absent before any nearby zone becomes operational and emerges only afterward (Figure 5, Panel A), indicating that the source area's emissions were triggered by zone establishment rather than pre-existing.

Identification assumptions. For β to be interpreted as the causal effect of downwind exposure on PM2.5, the wind-direction component of E_{imt} must be exogenous to unobserved determinants of PM2.5, conditional on cell and year-month fixed effects.¹⁶ After absorbing cell and year-month fixed effects, the residual variation in D_{ijmt} reflects short-run weather rather than climatological patterns or regional seasonal cycles. This residual variation is plausibly orthogonal to the determinants of zone siting and operational timing, which follow a multi-year central-government approval process initiated years before zones become operational. Consistent with this, Appendix

¹⁶The exposure measure has the structure of a shift-share regressor, where zone operational status $\mathbb{I}(t \geq T_j)$ provides the shifts and the inverse-distance-weighted monthly downwind share $\text{distance}_{ij}^{-1} \cdot D_{ijmt}$ provides the shares. Recent work on shift-share identification establishes that consistent estimation requires exogeneity of only one component (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022). Here the share component carries the identification.

Table A4 reports that across districts within the same province and distance band in the analysis sample, those with higher average downwind exposure do not systematically differ on pre-treatment characteristics (urban share, ethnic minority share, secondary education share, and farm labor share from the 2002 VHLSS), cell elevation, the cell’s share of its province’s 2000 population, or pre-treatment mean log PM2.5. As a further check, I estimate equation (4) replacing log PM2.5 with the cell’s share of provincial population as the outcome. The coefficient on downwind exposure is small and statistically insignificant ($\hat{\beta} = 0.0001$, $p = 0.34$), indicating that the wind-based measure does not predict demographic changes in the transport ring. I present additional validation tests alongside the main results in the next subsection.

5.2 Results

Table 4 reports the pooled estimates. Column (2) is the preferred specification, estimated on cells in the 5–25 km transport ring around the nearest zone with cell fixed effects, year-month fixed effects, and time-varying weather controls. A one-standard-deviation increase in inverse-distance-weighted downwind exposure raises monthly PM2.5 by 2.7%, equivalent to $0.66 \mu\text{g}/\text{m}^3$ at the unweighted sample mean of $24.4 \mu\text{g}/\text{m}^3$. The standardizing unit, 0.078 km^{-1} , corresponds to a cell fully downwind of a single operational zone at approximately 13 km for the entire month.

The timing of the effect matches zone establishment. Panel A of Figure 5 reports the effect by event time: a positive effect appears once a zone becomes operational and persists across all post-treatment windows, estimated more precisely at longer horizons (blue markers). Two placebo tests show that the effect appears only under the design’s own variation: the actual timing of operations and the actual location of sites.

First, when the exposure measure’s wind and proximity geometry is applied before any nearby zone operates, the placebo coefficient is 0.008 (s.e. 0.006): no pre-treatment window shows a significant effect, and the year immediately before operation is estimated most precisely and sits near zero (gray markers). Second, when every zone is relocated to a randomly drawn site in its own province (at least 25 km from the province’s actual zones) and the operational years are reshuffled among them, the effect is equally absent: across 500 such permutations, the placebo coefficients center tightly on zero (mean 0.0002, standard deviation 0.0034).^{17,18} The baseline estimate of 0.027 lies approximately eight standard deviations above their mean, outside the support of the distribution entirely (Panel B). The notes to Figure 5 detail the construction of both tests.

¹⁷The exclusion buffer is enforced only against same-province real zones. Fake zones in dense industrial corridors near province boundaries may fall within 25 km of real zones in neighboring provinces, in which case placebo exposure absorbs some signal from genuine downwind exposure. This residual overlap biases the placebo distribution slightly toward the baseline estimate, making the test more conservative than an idealized null.

¹⁸Several small, densely industrialized provinces cannot accommodate all of their zones under this within-province constraint: Bac Ninh, with 10 real zones in approximately 820 km^2 , is the clearest case, but Ha Nam, Thai Binh, Hai Duong, and Ho Chi Minh City also exhibit insufficient candidate area outside the 25 km exclusion buffer of their own zones. For these provinces, the algorithm falls back to drawing from candidate districts in the same broader region (e.g., the Red River Delta for the northern provinces; the Southeast region for Ho Chi Minh City). Across the 500 permutations, on average 81 of 287 zones (28%) are placed through this regional fallback, with the remaining 206 zones (72%) placed strictly within province.

Table 4: Wind-based Estimates of Industrial Zone Exposure on PM2.5

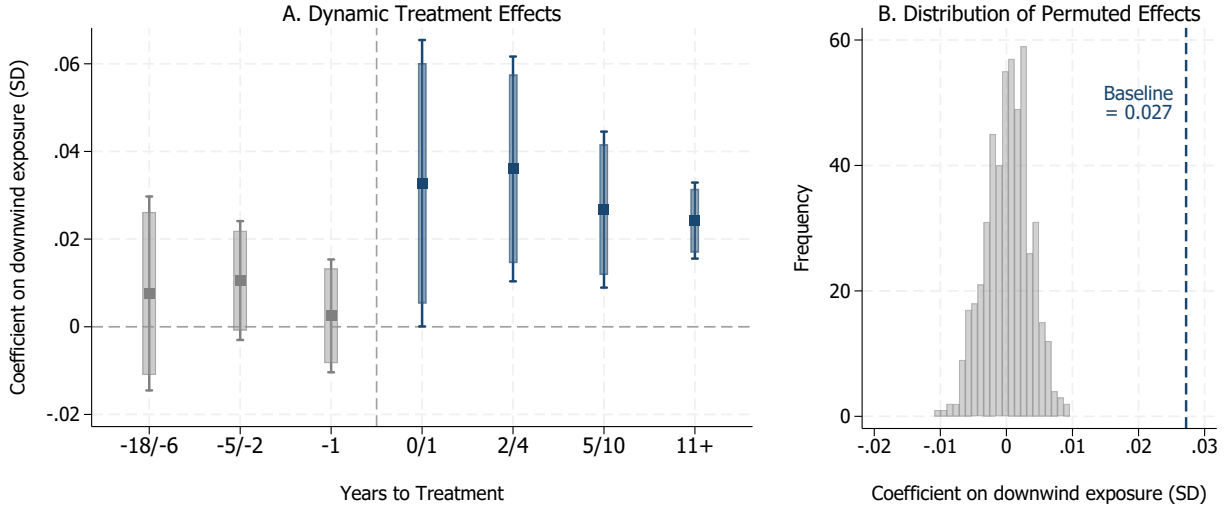
	Dependent Variable is $\ln(\text{PM}_{2.5})$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Downwind Exposure (SD)	0.025 (0.005)	0.027 (0.005)	0.025 (0.005)	0.029 (0.006)	0.021 (0.006)	0.032 (0.007)	0.024 (0.004)	0.030 (0.006)
Implied effect ($\mu\text{g}/\text{m}^3$)	0.61	0.66	0.61	0.71	0.52	0.77	0.59	0.73
Sample SD of exposure	0.078	0.078	0.078	1.152	0.151	0.097	0.054	0.100
Unstandardized coefficient	0.322	0.348	0.322		0.139	0.331	0.445	0.300
Observations	793212	793212	793212	793212	909036	1317156	793212	793212
Adjusted R-squared	0.722	0.739	0.770	0.740	0.742	0.726	0.739	0.739
Cell FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Weather Controls		Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year-Month FE $\times$ Baseline Controls			Yes					
Sample	5-25km	5-25km	5-25km	5-25km	0-25km	5-40km	5-25km	5-25km
Wind Aggregate	IDW	IDW	IDW	EW	IDW	IDW	IDW	IDW
Wind Angle $\pm^\circ$	45	45	45	45	45	45	30	60

Notes: The table reports estimates from equation (4). The dependent variable is the log of monthly mean PM_{2.5} concentration ($\mu\text{g}/\text{m}^3$) in a 0.05 degree cell. The regressor is the cell's monthly downwind exposure to operational industrial zones, defined in equation (5) and scaled by its standard deviation within each column's estimation sample, so that coefficients give the effect of a one-standard-deviation increase in exposure. The downwind share for a cell-zone pair is the share of days in the month on which the wind at the zone, from ERA5-Land 06:00 UTC 10-meter wind fields (Copernicus Climate Change Service (C3S), 2019; Muñoz-Sabater et al., 2021), blew within $\pm 45^\circ$ of the bearing from the zone to the cell; columns (7) and (8) use $\pm 30^\circ$ and $\pm 60^\circ$. Column (4) replaces the inverse-distance weights with equal weights (EW). Zones are treated as operational from one year before their recorded operational date, as in equation (3). The row "Implied effect ($\mu\text{g}/\text{m}^3$)" is the coefficient times the sample-mean PM_{2.5}, "Sample SD of exposure" the standard deviation of the unstandardized exposure, and "Unstandardized coefficient" the coefficient on the raw exposure measure, which is comparable across the inverse-distance-weighted columns, whose measure shares the same units (km^{-1}), and is omitted for column (4), whose measure has different units. All columns except column (1) include cell fixed effects, year-by-month fixed effects, and weather controls (a quadratic in monthly mean 2-meter temperature and a linear term in monthly total precipitation, both from ERA5-Land); column (1) omits the weather controls. Column (3) adds interactions of the year-by-month fixed effects with the four baseline district characteristics listed in the notes to Table 3. Columns (1)–(4), (7), and (8) use cells in the 5–25 km transport ring around the nearest zone; column (5) adds the 0–5 km source ring, where direct emissions dominate exposure regardless of wind direction; column (6) extends the buffer to 40 km. Province-clustered standard errors in parentheses.

The baseline estimate summarizes the average response of PM_{2.5} to a one-standard-deviation shift in downwind exposure, not the burden this implies for the people living in the transport ring: cells differ in how much exposure they receive, and population is unevenly distributed across them. I therefore compute the implied level effect of zone exposure for each cell-month, $\text{PM}_{2.5}^{\text{imt}} \cdot [1 - \exp(-\hat{\beta} \cdot E_{\text{imt}} / \sigma_E)]$, and average it over the 5–25 km ring weighted by year-2000 cell population. The result is an increment of $0.70 \mu\text{g}/\text{m}^3$ per month, 14% of the World Health Organization's 2021 annual PM_{2.5} guideline of $5 \mu\text{g}/\text{m}^3$. It measures the average monthly PM_{2.5} increment from zone exposure borne by a typical resident of the ring.

The increment is large relative to the deterioration in air quality over the study period and uneven across cells. Population-weighted PM_{2.5} in the transport ring rose from $22.0 \mu\text{g}/\text{m}^3$ in 2000–2002 to $27.9 \mu\text{g}/\text{m}^3$ in 2017–2019, an increase of $5.9 \mu\text{g}/\text{m}^3$ that took the average resident

Figure 5: Industrial Zones and PM2.5: Dynamic Effects and Permutation Test



Notes: Panel A reports pre-treatment placebo coefficients (gray markers, left of the dashed line) and post-treatment dynamic effects (blue markers, right). The placebo exposure omits the operational-status indicator from equation (5), so that it reflects only the directional and proximity geometry of zones, and is interacted with three pre-treatment window indicators (≤ -6 , -5 to -2 , and -1 years before the first nearby zone becomes operational) in one regression on the pre-treatment sample. Post-treatment coefficients come from separate regressions, each combining all pre-treatment cell-months of treated cells with the post-treatment cell-months in a given event-time window; never-treated cells are included in all regressions. Event time is years since the first zone within 25 km became operational. Panel B plots the distribution of placebo coefficients from a permutation test. In each of 500 permutations, every zone is relocated to a randomly drawn point within a district in the same province, at least 25 km from every actual zone in that province (if no such point is found after 40 attempts, the zone is placed in a randomly drawn district of the same broader region), and operational years are reshuffled across placebo zones within each province. For each permutation, E_{imt} is recomputed from equation (5), scaled by its own within-sample standard deviation, and equation (4) is re-estimated. The vertical dashed line marks the baseline estimate. All specifications are those of column (2) of Table 4. Navy markers are point estimates, bars 90% and caps 95% confidence intervals from province-clustered standard errors. Source: PM2.5 data from Shen et al. (2024); wind data from ERA5-Land (Muñoz-Sabater et al., 2021); industrial zone locations from the Ministry of Planning and Investment.

from below to above Vietnam's national ambient standard of $25 \mu\text{g}/\text{m}^3$ (QCVN 05:2023/BTNMT). The zone-attributable increment accounts for approximately 12% of this rise. Across cells, the effect is considerably larger for the heavily exposed: at the 95th percentile of cell-month exposure (2.1 standard deviations), the implied increment is $1.39 \mu\text{g}/\text{m}^3$, 28% of the WHO annual guideline, and at the 99th percentile (5.1 standard deviations) it reaches $3.34 \mu\text{g}/\text{m}^3$, or 67% of the guideline. The most exposed cells are those located close to and directly downwind of multiple operational zones simultaneously.

Robustness. The estimate does not depend on how downwind exposure is constructed or on the sample of cells over which it is estimated (columns (3) to (8) of Table 4). Allowing districts with different 2002 characteristics to follow their own time paths leaves it unchanged. An equal-weighted sum of monthly downwind shares in place of the inverse-distance-weighted measure, a buffer extended to 40 km, and angular tolerances of $\pm 30^\circ$ or $\pm 60^\circ$ give standardized coefficients close to the baseline. Including the 0–5 km source ring attenuates the per-unit effect, as expected where

direct emissions dominate exposure regardless of wind direction. Dropping one province at a time leaves every estimate positive and statistically significant (Appendix Figure A7). Under more conservative fixed effects, region-by-year-month and province-by-year-month, the coefficients are smaller but positive and statistically significant. These fixed effects compare a cell with neighbors that share its surroundings and thus exclude the pollution transported from other zones. A decomposition by distance shows that the effect of a cell's own exposure is similar under all three sets of fixed effects, and the smaller composite coefficients therefore reflect the excluded transport rather than a weaker effect (Appendix C).

5.3 Heterogeneity in the Pollution Effect

Table 5: Industrial Zones and PM2.5: Dynamic Effects, Overall and by Establishment Cohort

	Dependent Variable is $\ln(\text{PM}_{2.5})$			
	Event-time Window			
	Pooled (1)	0–4 (2)	5–10 (3)	11+ (4)
Panel A: All Zones				
Downwind Exposure (SD)	0.027 (0.005)	0.040 (0.012)	0.027 (0.009)	0.024 (0.004)
Implied effect ($\mu\text{g}/\text{m}^3$)	0.65	0.96	0.65	0.61
Mean PM2.5 ($\mu\text{g}/\text{m}^3$)		24.0	24.5	25.2
Observations	793212	268080	346176	420660
Adjusted R-squared	0.739	0.749	0.747	0.739
Panel B: By Establishment Cohort				
Pre-2003 cohort	0.023 (0.004) [322164]	0.013 (0.019) [24300]	0.017 (0.007) [86028]	0.021 (0.004) [211836]
2003–2009 cohort	0.047 (0.013) [321480]	0.045 (0.012) [106140]	0.042 (0.016) [127368]	0.049 (0.013) [87972]
Post-2009 cohort	0.017 (0.021) [28716]	-0.007 (0.018) [16788]	0.052 (0.031) [11928]	

Notes: All columns use the specification of column (2) of Table 4, equation (4) with cell and year-month fixed effects and weather controls on cells in the 5–25 km transport ring, estimated by event-time window, where event time is years since the first zone within 25 km became operational. Panel A pools all zones. Column (1) reproduces column (2) of Table 4. Columns (2)–(4) restrict the sample to treated cells and combine their pre-treatment observations with the post-treatment observations in the indicated window. The implied effect is the coefficient times the mean PM2.5 among treated cells in the window, reported in the next row. Panel B estimates the same specifications separately by establishment cohort (the year the first zone within 25 km became operational: before 2003, 2003 to 2009, and after 2009). Column (1) of Panel B pools all post-treatment windows of the cohort, and brackets report the treated cell-months in each window. Mean PM2.5 over the post-treatment cell-months is $25.5 \mu\text{g}/\text{m}^3$ for the pre-2003 cohort and 24.0 for the 2003–2009 cohort, and the implied effects per standard deviation of exposure, 0.57 and $1.14 \mu\text{g}/\text{m}^3$, differ by the same factor as the coefficients. The pre-2003 cohort has no pre-treatment observations in the panel, and its estimates are identified from monthly variation in downwind exposure after treatment. The post-2009 cohort is not observed beyond event year 10. Exposure is standardized on the full sample in every cell. Standard errors clustered at the province level are in parentheses.

The pooled estimate in Table 4 combines zones observed at different ages and established in different periods. Panel A of Table 5 summarizes the event-time profile of Figure 5 in three windows of years since the first zone within 25 km became operational. Because the later windows are identified from earlier cohorts, a change across them could reflect either the aging of zones or differences across cohorts. Panel B therefore estimates the same windows separately by establishment cohort (the year of the first zone within 25 km), grouped into the three regulatory periods of Section 2.

The effect per unit of exposure shows no sign of weakening as zones age. The point estimates in Panel A fall from 0.040 in the first five years to 0.024 beyond year 10, but the decline is not statistically significant.¹⁹ Panel B shows that it reflects cohort composition rather than changes over time within a cohort. Within each cohort, the estimates do not differ across the windows, while across cohorts the effect of the 2003–2009 cohort, 0.047 per standard deviation of exposure, is about twice the pre-2003 cohort’s 0.023 and larger in every window. The pooled 11+ window in Panel A is identified mostly from the pre-2003 cohort, which is why it is the lowest. For the post-2009 cohort, Panel B reports only two windows and both are imprecise. Its zones are the most recent, observed for at most ten years after they became operational, and most are still filling in: as of 2023, their median occupancy rate is 30% and 66% of them remain under construction (Appendix Table A5). Its effect therefore cannot yet be assessed.²⁰

6 DISCUSSION

In this section, I discuss potential explanations for the pollution effects documented above, both for the rise in combustion at the source and for the differences in the PM_{2.5} effect across cohorts. I then translate the pollution into a health burden and compare it with the income gains from zones.

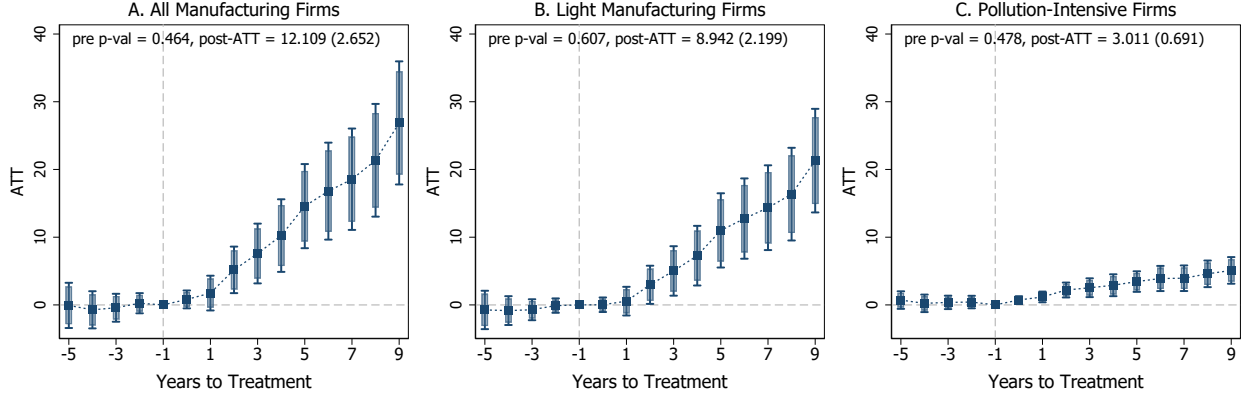
6.1 Potential Mechanisms

To organize the discussion, I follow the trade-and-environment literature, which distinguishes three channels through which emissions change. These are the volume of activity (scale), the industrial mix of that activity (composition), and emissions per unit of output within industries (technique) (Grossman & Krueger, 1995; Antweiler et al., 2001; Shapiro & Walker, 2018).

¹⁹The p-values are from Wald tests on a stacked version of the Table 5 regressions, in which the pre-treatment observations enter each window’s stack and standard errors are clustered at the province level: $p = 0.13$ for the first against the last window in Panel A, 0.74 and 0.75 for equality across windows within the pre-2003 and 2003–2009 cohorts, and 0.03 for the difference between the two cohorts.

²⁰One might be concerned about the right truncation of the post-2009 cohort: its zones are observed only in their first years of operation, and many of the zones approved in this period were not yet operating or filled by the end of the sample. If the effect built up over time, an average over those years would understate it, and I therefore draw no conclusion from this cohort’s estimates. The truncation does not affect the comparison between the two older cohorts, both of which are observed beyond ten years of operation and within each of which the estimates do not differ across the windows.

Figure 6: Industrial Zone Establishment and Manufacturing Firm Counts



Notes: This figure reports event-study estimates of industrial zone establishment on the number of all manufacturing firms (Panel A), light manufacturing firms (Panel B) and pollution-intensive manufacturing firms (Panel C) in districts within 5 km of a zone. The estimates are from $y_{dt} = \sum_{k=-5}^9 \beta_k \mathbb{I}\{e_{dt} = k\} + \alpha_d + \mu_t + \varepsilon_{dt}$, where the outcome y_{dt} is the firm count in district d in year t , over all firms in the survey, and α_d and μ_t are district and year fixed effects. The survey enumerates every state and foreign-invested firm and every private firm with 10 or more workers in each round and covers smaller private firms on a sample basis in some rounds, and estimates restricted to firms with 10 or more workers are similar. Event time $e_{dt} = t - T_d$ is defined relative to the year in which the first industrial zone within 5 km of district d became operational, with $k = -1$ as the reference period. Pollution-intensive and light manufacturing are classified as in Table 1. Estimation uses the heterogeneity-robust estimator of de Chaisemartin and d’Haultfoeuille (2024), with districts never within 5 km of a zone as controls. Because the estimator compares districts around the opening of their first nearby zone with never-treated districts, identification comes from zones becoming operational during the sample period, most of which belong to the 2003–2009 cohort. The pre-trend p-value reports a joint test of the null that all pre-treatment coefficients equal zero. Post-ATT is the average treatment effect across all post-treatment periods, with the standard error in parentheses. Province-clustered standard errors; outer bars 95% and inner bars 90% confidence intervals. Data are from the Vietnam Enterprise Survey for firm counts and the Ministry of Planning and Investment for industrial zone locations.

The rise in combustion at the source. The rise in thermal anomalies after zone establishment documented in Section 4 is consistent with the scale channel, which operates here through the entry of firms (the extensive margin). In districts within 5 km of a zone, the number of manufacturing firms grows steadily over the zone’s first decade of operation, by about 12 per district on average (Figure 6, Panel A), and manufacturing revenue rises by roughly 30% ($p = 0.01$, Table 6, Panel A). Although most of the increase is in light manufacturing (Figure 6, Panel B), about 3 of the 12 are pollution-intensive (Panel C), including about 2 in hot-process industries. Some of the firms are also large: the number with 100 or more workers rises by about 2.6 per district, roughly one in five (Table 6, Panel A). Within an industry, larger firms are more likely to burn fossil fuel and burn more of it per unit of output, and the gradient is present within light manufacturing (Appendix Figure A8). The predominance of light manufacturing therefore does not necessarily make this activity cleaner, and these large and high-output plants are likely the baseline foundation of a zone’s environmental footprint.

The differences across cohorts. To understand why the 2003–2009 cohort’s effect on PM2.5 is about twice the pre-2003 cohort’s and larger in every window (Table 5, Panel B), I examine the types of firms that operate inside each cohort’s zones. The post-2009 cohort is left out because

Table 6: Manufacturing Entry near Zones and the Tenant Firms of Each Cohort

Panel A: Effect of Zone Establishment on Manufacturing Activity, All Zones

	Number of firms with 100+ employees (1)	Number of hot-process firms (2)	ln(Manufacturing revenue) (3)
Post-ATT	2.6 (0.4)	1.8 (0.5)	0.263 (0.104)
pre-trend p-value	0.114	0.356	0.509
Observations	7665	7665	6964

Panel B: Manufacturing Firms inside Zones, by Establishment Cohort

	Median employment (1)	Share of firms with 100+ employees (2)	Share of firms with 500+ employees (3)	Share pollution- intensive (4)	Share hot-process (5)
<i>Pre-2003 cohort</i>					
All tenants	82	0.458	0.139	0.190	0.087
Established before 2003	158	0.604	0.214	0.197	0.097
Established 2003–2009	85	0.464	0.134	0.210	0.095
Established after 2009	48	0.323	0.079	0.172	0.088
<i>2003–2009 cohort</i>					
All tenants	86	0.466	0.205	0.251	0.207
Established 2003–2009	102	0.520	0.167	0.265	0.216
Established after 2009	60	0.381	0.143	0.262	0.226

Notes: Panel A reports the average post-treatment effect of zone establishment on district totals in districts within 5 km of a zone, estimated as in Figure 6. Revenue is gross turnover in 2010 prices. Because the survey is a repeated cross section, the estimates are net changes. Panel B describes manufacturing firms that report operating inside an industrial zone (2011 and 2013 to 2016). The survey records a firm's district but not its zone, so firms are assigned to the cohort of the first zone within 25 km of their district, as in Panel B of Table 5. The rows by year of establishment use the 2013 round, the only round with both the zone question and the year of establishment for every firm, and therefore describe the tenants operating in 2013 by the year they were established rather than the entrants of each period. The 24 tenants of the 2003–2009 zones established before 2003 and the post-2009 cohort, whose zones were still filling in (Appendix Table A5), are not reported. Pollution-intensive industries are defined as in Table 1, and hot-process industries are non-metallic minerals, basic metals, and coke and refined petroleum. Province-clustered standard errors in parentheses. Source: Vietnam Enterprise Survey and the Ministry of Planning and Investment zone database.

its zones were still filling in at the end of the sample period. Since the survey does not record when a firm moved into a zone, I treat tenants established in the period in which a cohort's zones opened as the zones' first tenants and those established afterwards as later tenants.²¹ As seen in Panel B of Table 6, the two cohorts differ mainly in industrial composition. Among first tenants, hot-process industries account for 0.22 in the 2003–2009 zones against 0.10 in the pre-2003 zones, and the broader group of pollution-intensive industries for about a quarter against a fifth. The two shares are similar among later tenants in each cohort, suggesting that the mix of each cohort's zones changed little as they filled (see also Panel A of Appendix Table D1 for national trends). Differences in size, on the other hand, are less consistent. Among first tenants, a larger share

²¹The year of establishment is not necessarily the year a firm moved into a zone: an enterprise that existed before a zone opened can move into it from within its own district or open a plant there. In the 2013 round, 7% of the pre-2003 cohort's tenants and 33% of the 2003–2009 cohort's were established before the first zone within 5 km of their district.

employs 500 or more workers in the pre-2003 zones, whereas among all tenants the share is one in five in the 2003–2009 zones against one in seven in the pre-2003 zones.

The per-exposure effect also varies with the industrial mix around a zone within a cohort. For the 2003–2009 cohort, it is about twice as large in districts whose manufacturing base was pollution-intensive before the zones opened as in districts with a lighter base: 0.080 against 0.038 per standard deviation of exposure, positive at every horizon (Table 7).

Table 7: Industrial Zones and PM2.5: Dynamic Effects by Baseline Pollution Intensity

2003–2009 cohort	Dependent Variable is $\ln(\text{PM}_{2.5})$			
	Event-time Window			
	Pooled (1)	0–4 (2)	5–10 (3)	11+ (4)
Downwind Exposure (SD), Low	0.038 (0.013)	0.031 (0.018)	0.025 (0.017)	0.043 (0.012)
Downwind Exposure (SD), High	0.080 (0.015)	0.077 (0.022)	0.092 (0.018)	0.077 (0.016)
Difference, High – Low	0.042 (0.015)	0.046 (0.030)	0.067 (0.018)	0.033 (0.016)
	[303480]	[99900]	[119880]	[83700]

Notes: All columns use the specification of column (2) of Table 4, estimated by event-time window on the cells of the 2003–2009 establishment cohort, defined as in Panel B of Table 5, with sample construction following that table. Cells are classified by whether their district’s baseline (2001–2003) pollution-intensive manufacturing share is above or below the national median. The pollution-intensive classification is the VES energy-based measure (Appendix B). The comparison is restricted to the 2003–2009 cohort because its zones enter after the baseline window. The pre-2003 cohort is omitted because its zones predate the window, and the split does not measure the base that preceded them. The post-2009 cohort is omitted because its short exposure history and small cell counts make the split by baseline share too imprecise to be informative. Downwind exposure is standardized on the common estimation sample and interacted with the high/low indicator; the two coefficients are therefore directly comparable. The rows “Difference, High – Low” report the gap between the two coefficients with its standard error, and the bracketed figures are the treated cell-months in each window. Province-clustered standard errors in parentheses.

The setting described in Section 2 helps explain these patterns. Standards on operating firms were weakly enforced, and abatement after entry would thus do little to narrow the gap in fuel use between a heavier and a lighter mix. Consistent with this, for the pre-2009 cohorts, the per-exposure effect is no lower after the emission standards of 2009 and 2013 took effect than in the years before them ($p = 0.42$ for equality across the three periods).²² Firms near these zones also show no change in fuel use after the standards relative to firms in the same industries elsewhere (Appendix Table A6). Within each cohort, the mix and the per-exposure effect changed little as the zones filled. The results therefore suggest that it is the composition of the industries a zone admits that plays a decisive role in explaining the differences across cohorts.

It is worth noting that the evidence cannot separate the three channels or measure their contri-

²²The estimates come from equation (4) on the cells of the pre-2003 and 2003–2009 cohorts, pairing the pre-treatment months with the post-treatment months of one calendar period at a time: 0.025 (s.e. 0.005) through 2009, 0.021 (0.008) in 2010–2013, and 0.028 (0.005) from 2014 onward. The p-value is from a Wald test on a stacked version of the three regressions, as for the windows of Table 5.

butions to the rise in combustion at the source or to the gap between cohorts. Such an accounting would require the output and emissions of each zone’s tenants. The data instead record thermal anomalies within 5 km that cannot be assigned to plants, and the firm surveys give a firm’s district rather than its zone. For the gap between cohorts, scale therefore cannot be ruled out: a larger share of the tenants of the 2003–2009 zones employ 500 or more workers (Table 6, Panel B). Within a cohort, the relatively stable per-exposure effect does not imply that aggregate production scale and the entry of firms are negligible in generating a zone’s total emissions: the coefficient is a marginal effect per unit of exposure and does not vary with the amount of detected combustion at the source (Appendix E). Technique cannot be ruled out either. The energy data measure what firms burn but not how they burn it, and tenants that entered in different periods may operate equipment of different vintages. A technique gap between cohorts that is constant over time would produce the flat profiles within cohorts and the difference across them that Table 5 shows, and cleaner production that coincided with the growth of a zone’s activity would leave the per-exposure effect unchanged.

6.2 Health Burden

To provide a sense of the welfare magnitude, I translate the PM_{2.5} estimates into illustrative health costs using standard dose-response parameters from the epidemiological literature. The translation rests on three assumptions that go beyond the empirical design. First, the wind-based identification isolates only the directional component of zone-attributable pollution, the increment that varies with a cell’s position in the downwind plume. Pollution from zones that disperses uniformly or accumulates as elevated background concentration is absorbed by cell and year-month fixed effects, and the population-weighted increment of $0.70 \mu\text{g}/\text{m}^3$ is therefore a partial measure of the total burden. The health costs derived from it can be interpreted as a lower bound.

Second, the dose-response parameters used below are estimated from cross-sectional variation in chronic, long-run average exposure (Ebenstein et al., 2017), whereas the wind-based estimate reflects episodic, directional variation in plume exposure. Under a linear dose-response, the timing of exposure within the month does not matter. Under the concave exposure-response curves used in the Global Burden of Disease, an increment that arrives in spikes on a few downwind days causes less damage than the same increment spread evenly. The translation treats it as evenly spread and would then overstate the burden. The overstatement is unlikely to be large: under a logarithmic curvature, the shape of the Global Exposure Mortality Model (Burnett et al., 2018), even if the entire monthly increment arrived on a single day, the implied burden would fall by about a quarter.²³ Third, the monetary figures rest on a value of a statistical life transferred from

²³The Global Exposure Mortality Model specifies the log hazard ratio as proportional to $\ln(1 + z/\alpha)$ in the concentration z above a counterfactual (Burnett et al., 2018). The estimated population-weighted increment of $0.70 \mu\text{g}/\text{m}^3$ is a monthly average. Suppose it came entirely from a single day. The concentration on that day would then be $0.70 \times 30 = 21 \mu\text{g}/\text{m}^3$ above the sample-mean baseline of $24 \mu\text{g}/\text{m}^3$, and the other days of the month would be unchanged. Under a linear dose-response, the damage from that day is proportional to the change in concentration relative to the baseline, $21/24 = 0.88$. Under a logarithmic curve, it is proportional to the change in the log of concentration, $\ln(45/24) = 0.63$. The two coincide for small changes and diverge as the change grows. The concentrated day

the United States to Vietnam, scaled by relative income.

With these caveats, I apply the dose-response relationship from Ebenstein et al. (2017), in which each additional $10 \mu\text{g}/\text{m}^3$ of PM10 reduces life expectancy by 0.64 years, and convert it to PM2.5 terms with the Air Quality Life Index methodology, which yields 0.098 life-years per $\mu\text{g}/\text{m}^3$ of PM2.5.²⁴ At the point estimate, this implies a reduction of approximately 25 days of life expectancy per exposed individual ($0.70 \times 0.098 \times 365 \approx 25$). The 95% confidence interval on the PM2.5 coefficient gives a range of approximately 15 to 33 days. The burden is unevenly distributed: at the 95th percentile of cell-month exposure, the implied reduction is approximately 50 days, twice the population-weighted average.

Approximately 46 million people reside in the 5–25 km transport ring as of 2000, implying an aggregate burden of roughly 3.2 million person-years of life expectancy lost across the exposed population. To value this burden, I use the benefit-transfer value of a statistical life for Vietnam: \$342,000 in 2015 dollars under an income elasticity of 1.0, a common default for transfers to lower-income countries (Viscusi & Masterman, 2017), or about \$14,800 per life-year.²⁵ At this value, the aggregate burden amounts to an illustrative present value on the order of \$47 billion, and the 25 days per exposed person to about \$44 per year over a 40-year horizon. Under an income elasticity of 1.5, the reference case for low- and middle-income countries (Robinson et al., 2019), both figures fall by about four fifths, to roughly \$9 billion and \$8 per person per year.

For comparison, Pham (2026) estimates that household labor income per capita rises by about 7% in districts near zones, roughly \$30 to \$50 per person per year. With the three assumptions above, which understate the cost on one side and overstate it on the other, the health cost per exposed person ranges from roughly one fifth of the income gain to roughly its size.²⁶ A complete welfare evaluation would also require the fiscal costs of the program (foregone corporate income tax, duty-free imports of capital equipment, land-rent exemptions, and publicly built zone infrastructure), and these are not reported at any level of aggregation: Vietnam publishes no information on tax expenditures (Truong & Le, 2016), and provincial budgets do not separate zone-related spending. I therefore interpret the comparison as placing the environmental cost on the same order as the income gains rather than as a cost-benefit result.

would therefore cause $0.63/0.88 \approx 0.72$ of the damage that the linear translation assigns to the month. If the increment came instead from a quarter of the month's days, the ratio would be 0.95. A curve more concave than the logarithmic one would lower these ratios, and a milder one would raise them.

²⁴The conversion assumes a PM2.5-to-PM10 ratio of 0.65, so that $10 \mu\text{g}/\text{m}^3$ of PM2.5 corresponds to approximately $15.4 \mu\text{g}/\text{m}^3$ of PM10, implying $0.64 \times (10/6.5) \approx 0.98$ life-years per $10 \mu\text{g}/\text{m}^3$ of PM2.5. The assumed ratio is consistent with Vietnam's ambient monitoring network, which reports PM2.5-to-PM10 ratios of 0.65 to 0.75 across six reference-grade stations spanning the north-south gradient (Ministry of Natural Resources and Environment, 2021).

²⁵The value adjusts a U.S. base VSL of \$9.6 million by GNI per capita with an income elasticity of 1.0. It converts to a value of a statistical life-year of approximately \$14,800 at a 3% discount rate ($\text{VSLY} = \$342,000 / 23.1$, where 23.1 is the annuity factor over 40 years of remaining life expectancy). The 25 days per person, valued at this rate, amount to about \$1,000 in present value, or about \$44 per year over the same horizon. The figures are also sensitive to the base VSL and the discount rate.

²⁶The income gain converts the estimates of Pham (2026), in 2010 Vietnamese dong, to 2015 dollars. The gain accrues to households in districts within about 15 km of a zone, whereas the exposed population is that of the 5–25 km ring, and the comparison treats the two populations as interchangeable. The concavity bound above leaves this range essentially unchanged.

7 CONCLUSION

In this paper, I estimate the environmental cost of a place-based industrial policy in a setting where official measurement of industrial emissions is limited, and I examine the potential mechanisms behind it. I proceed in two steps that link the source of the pollution to the population exposed to it: satellite detections of industrial combustion indicate activity at the source, and wind-driven variation in downwind exposure identifies the zones' contribution to the ambient PM_{2.5} borne by the nearby population. In Vietnam, zone establishment more than doubles the frequency of detected combustion within 5 km. The rise is consistent with the scale of the activity that a zone attracts through the entry of firms, some of them far larger than the typical manufacturing firm. These large and high-output plants are likely the baseline foundation of a zone's environmental footprint. The transported pollution raises monthly ambient PM_{2.5} in the 5–25 km transport ring by a population-weighted 0.70 $\mu\text{g}/\text{m}^3$. Under standard dose-response estimates, the implied reduction in life expectancy is approximately 25 days per exposed individual, or 3.2 million person-years across the 46 million residents of the ring. Valued at the country's income level, the health cost per exposed person can be of the same order as the income gains that zones generate for nearby residents.

The PM_{2.5} effect per unit of exposure is about twice as large for zones established in 2003–2009 as for those established before 2003 and, within a cohort, remains relatively stable as the zones age. The 2003–2009 zones host a heavier mix of industries, and the mix of tenants changed little as the zones filled. In a setting where standards on operating firms are weakly enforced, these patterns suggest that the composition of the industries a zone admits plays a decisive role in explaining why the effect differs across cohorts. However, scale and technique cannot be ruled out: the same zones also have a higher share of very large plants among their tenants, and the data do not permit a decomposition of output scale or a comparison of production techniques within industries.

The estimates measure the local environmental footprint of zones, the pollution imposed on the populations around them, and not the net national effect of the policy. The two differ by what the activity inside zones would have emitted elsewhere had the zones not been built, and the sign of the difference is not determined in advance: a firm that moves to a zone lowers exposure at its origin and raises it at its destination, while one that expands on entering a zone emits more than before. The same problem arises for environmental regulation in China, where production shifted from regulated to unregulated plants of the same business group (Chen et al., 2025). Within the study area, the data show little sign of such displacement: combustion in the comparison bins does not change after a zone opens, pollution-intensive firm counts beyond 15 km fall by at most a tenth of the near-zone increase (Section 4, Appendix Figure A9), and few firms move into treated districts (Appendix Table A1). What these checks cannot detect is relocation from beyond 50 km, or a plant opened in a zone while another closes elsewhere, because the survey records enterprises rather than plants. Such relocation was official policy in the two largest cities, though the programs were directed at wastewater and residential nuisance rather than combustion and moved

few large sources.²⁷ For the populations near zones, the cost is the same whether the activity is new or relocated, but for a national accounting the part of the footprint that would otherwise have been emitted elsewhere in Vietnam remains unidentified.

For the evaluation of place-based industrial policy, environmental costs are a quantitatively meaningful component of its welfare impact, large enough to belong in cost-benefit assessments alongside the productivity, employment, and income effects on which the literature has focused. For the design of zone programs, the results highlight the importance of the industries a zone admits. In the United States, the decline in manufacturing emissions came mainly from cleaner production within industries (Levinson, 2009), attributed to environmental regulation enforced on operating firms (Shapiro & Walker, 2018). Where such enforcement is limited, as in Vietnam and other developing countries (Duflo et al., 2013; Greenstone & Hanna, 2014), the selection of industries at the approval stage is an instrument that remains available to policymakers. Screening projects by production type is one way to implement this selection, and Vietnam's list of production types with pollution risk (Decree 40/2019/ND-CP, Appendix IIa) already includes the hot-process industries that distinguish the 2003–2009 zones. Place-based programs elsewhere also shaped which industries entered. A tax incentive program in India denied its benefits to environmentally damaging industries, and forest cover in the region did not decline despite a large expansion of economic activity (Garg & Shenoy, 2021). In a new area in China, local governments raised entry barriers for polluting firms, and the regulation of these firms accounts for about half of the reduction in PM2.5 (Lin et al., 2026). In settings where enforcement on operating firms is limited, the environmental record of a zone program is largely decided when its projects are approved.

REFERENCES

- Abagna, M. A., Hornok, C., & Mulyukova, A. (2025). Place-based policies and household wealth in africa. *Journal of Development Economics*, 176, 103482.
- Akinci, G., & Crittle, J. (2008). *Special economic zone: Performance, lessons learned, and implication for zone development* (tech. rep.). The World Bank.
- Anderson, M. L. (2020). As the wind blows: The effects of long-term exposure to air pollution on mortality. *Journal of the European Economic Association*, 18(4), 1886–1927.
- Antweiler, W., Copeland, B. R., & Taylor, M. S. (2001). Is free trade good for the environment? *American Economic Review*, 91(4), 877–908.

²⁷Ho Chi Minh City listed about 1,400 polluting establishments in residential areas for relocation to industrial zones and outlying districts in 2002 (Decision 80/2002/QD-UB), the national plan for seriously polluting establishments set out 439 cases for 2003–2007 (Decision 64/2003/QD-TTg), and Hanoi listed 117 inner-city facilities for relocation by 2020 (Decision 130/QD-TTg of 2015). Of the 284 production establishments under the national plan, 28 were relocated and most of the remainder were required to build wastewater treatment. The Ho Chi Minh City list drew on the industries barred from residential areas, among them dyeing, tanning, electroplating, chemicals, and slaughtering. Implementation was slow in both cities, and the cement, steel, and tobacco plants on the lists had for the most part not moved a decade later.

- Athukorala, P.-c., & Nguyen, T. K. (2022). Vietnam's foreign direct investment and export performance. In J. D. London (Ed.), *Routledge handbook of contemporary Vietnam* (pp. 149–163). Taylor & Francis.
- Athukorala, P.-c., & Tran, T. Q. (2010). Foreign direct investment in industrial transition: The experience of Vietnam. In P.-c. Athukorala (Ed.), *The rise of asia: Trade and investment in global perspective* (pp. 229–251). Routledge.
- Borusyak, K., Hull, P., & Jaravel, X. (2022). Quasi-experimental shift-share research designs. *Review of economic studies*, 89(1), 181–213.
- Borusyak, K., Jaravel, X., & Spiess, J. (2024). Revisiting event-study designs: Robust and efficient estimation. *Review of Economic Studies*, 91(6), 3253–3285.
- Burke, M., Driscoll, A., Lobell, D. B., & Ermon, S. (2021). Using satellite imagery to understand and promote sustainable development. *Science*, 371(6535), eabe8628.
- Burnett, R., Chen, H., Szyszkowicz, M., Fann, N., Hubbell, B., Pope III, C. A., Apte, J. S., Brauer, M., Cohen, A., Weichenthal, S., et al. (2018). Global estimates of mortality associated with long-term exposure to outdoor fine particulate matter. *Proceedings of the National Academy of Sciences*, 115(38), 9592–9597.
- Busso, M., Gregory, J., & Kline, P. (2013). Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review*, 103(2), 897–947.
- Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.
- Center for International Earth Science Information Network (CIESIN), Columbia University. (2017). *Gridded population of the world, version 4 (GPWv4): Population density, revision 11*. Palisades, NY, NASA Socioeconomic Data; Applications Center (SEDAC). <https://doi.org/10.7927/H49C6VHW>
- Chen, J., & Roth, J. (2024). Logs with zeros? some problems and solutions. *Quarterly Journal of Economics*, 139(2), 891–936.
- Chen, Q., Chen, Z., Liu, Z., Suárez Serrato, J. C., & Xu, D. Y. (2025). Regulating conglomerates: Evidence from an energy conservation program in China. *American Economic Review*, 115(2), 408–447.
- Copernicus Climate Change Service (C3S). (2019). ERA5-Land hourly data from 1950 to present. <https://doi.org/10.24381/cds.e2161bac>
- Correia, S., Guimarães, P., & Zylkin, T. (2020). Fast poisson estimation with high-dimensional fixed effects. *Stata Journal*, 20(1), 95–115.
- de Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9), 2964–2996.
- de Chaisemartin, C., & d'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 1–45.

- Deryugina, T., Heutel, G., Miller, N. H., Molitor, D., & Reif, J. (2019). The mortality and medical costs of air pollution: Evidence from changes in wind direction. *American Economic Review*, 109(12), 4178–4219.
- Duan, W., Jiang, M., & Qi, J. (2024). Poverty or pollution: The environmental cost of E-commerce for poverty reduction in China. *Journal of Environmental Economics and Management*, 127, 103029.
- Duflo, E., Greenstone, M., Pande, R., & Ryan, N. (2013). Truth-telling by third-party auditors and the response of polluting firms: Experimental evidence from India. *Quarterly Journal of Economics*, 128(4), 1499–1545.
- Ebenstein, A., Fan, M., Greenstone, M., He, G., & Zhou, M. (2017). New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai river policy. *Proceedings of the National Academy of Sciences*, 114(39), 10384–10389.
- Fan, M., Tang, Q., Wu, J., & Xiao, J. (2023). Unintended environmental consequences of place-based economic policies. *American Journal of Agricultural Economics*, 105(1), 291–315.
- Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., et al. (2007). The shuttle radar topography mission. *Reviews of geophysics*, 45(2).
- Garg, A., Kainou, K., & Pulles, T. (2006). Introduction. In S. Eggleston, L. Buendia, K. Miwa, T. Ngara, & K. Tanabe (Eds.), *2006 IPCC guidelines for national greenhouse gas inventories, volume 2: Energy*. Institute for Global Environmental Strategies.
- Garg, T., & Shenoy, A. (2021). The ecological impact of place-based economic policies. *American Journal of Agricultural Economics*, 103(4), 1239–1250.
- General Statistics Office. (2023). *Environmental statistics in Vietnam 2014–2021*. Statistical Publishing House. <https://www.nso.gov.vn/default/2024/01/so-lieu-thong-ke-moi-truong-vietnam-2014-2021/>
- Giglio, L., Schroeder, W., Hall, J. V., & Justice, C. O. (2021). *MODIS Collection 6 and Collection 6.1 Active Fire Product User's Guide, Version 1.0* (Technical Report). NASA. https://modis-fire.umd.edu/files/MODIS_C6_C6.1_Fire_User_Guide_1.0.pdf
- Giglio, L., Schroeder, W., & Justice, C. O. (2016). The collection 6 modis active fire detection algorithm and fire products. *Remote Sensing of Environment*, 178, 31–41.
- Goldsmith-Pinkham, P., Sorkin, I., & Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8), 2586–2624.
- Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2), 254–277.
- Greenstone, M., & Hanna, R. (2014). Environmental regulations, air and water pollution, and infant mortality in India. *American Economic Review*, 104(10), 3038–3072.
- Grossman, G. M., & Krueger, A. B. (1995). Economic growth and the environment. *Quarterly Journal of Economics*, 110(2), 353–377.

- Hallquist, M., Wenger, J. C., Baltensperger, U., Rudich, Y., Simpson, D., Claeys, M., Dommen, J., Donahue, N., George, C., Goldstein, A., et al. (2009). The formation, properties and impact of secondary organic aerosol: Current and emerging issues. *Atmospheric Chemistry and Physics*, 9(14), 5155–5236.
- Henderson, J. V., Storeygard, A., & Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2), 994–1028.
- Hess, S., Jaimovich, D., & Schündeln, M. (2021). Environmental effects of development programs: Experimental evidence from West African dryland forests. *Journal of Development Economics*, 153, 102737.
- Hollweg, C., Smith, T., & Taglioni, D. (2017). *Vietnam at a crossroads: Engaging in the next generation of global value chains*. World Bank Publications.
- Huang, R.-J., Zhang, Y., Bozzetti, C., Ho, K.-F., Cao, J.-J., Han, Y., Daellenbach, K. R., Slowik, J. G., Platt, S. M., Canonaco, F., et al. (2014). High secondary aerosol contribution to particulate pollution during haze events in China. *Nature*, 514(7521), 218–222.
- Jia, R., & Ku, H. (2019). Is china's pollution the culprit for the choking of south korea? evidence from the Asian dust. *Economic Journal*, 129(624), 3154–3188.
- Kline, P., & Moretti, E. (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *Quarterly Journal of Economics*, 129(1), 275–331.
- Levinson, A. (2009). Technology, international trade, and pollution from us manufacturing. *American Economic Review*, 99(5), 2177–2192.
- Lin, X., Liu, F., Zhang, P., & Zheng, S. (2026). Can we balance green and growth? the environmental effect of place-based policies. *Journal of Economic Behavior & Organization*, 246, 107548.
- Lu, F., Sun, W., & Wu, J. (2023). Special economic zones and human capital investment: 30 years of evidence from China. *American Economic Journal: Economic Policy*, 15(3), 35–64.
- Lu, Y., Wang, J., & Zhu, L. (2019). Place-based policies, creation, and agglomeration economies: Evidence from China's economic zone program. *American Economic Journal: Economic Policy*, 11(3), 325–360.
- Ma, C., Wu, J., Xiao, J., & Zhan, X. (2024). Environmental and regional economic development policy: Evidence from the Rise of Central China Program. *Journal of the Association of Environmental and Resource Economists*, 11(5), 1229–1273.
- McCaig, B., & Pavcnik, N. (2018). Export markets and labor allocation in a low-income country. *American Economic Review*, 108(7), 1899–1941.
- Ministry of Natural Resources and Environment. (2021). *National state of the environment report 2016–2020* (State of the Environment Report) (In Vietnamese. Chapter 9: Environmental Management). Ministry of Natural Resources and Environment, Socialist Republic of Vietnam. Hanoi. https://moit.gov.vn/upload/2005517/fck/files/20211108_Bao_cao_HTMT_2016-2020_F_a4980.pdf

- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J.-N. (2021). ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9), 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>
- NASA FIRMS. (2021). MODIS Collection 6.1 Active Fire Product (MCD14ML) [Accessed: 2026-03-15]. <https://firms.modaps.eosdis.nasa.gov/>
- Neumark, D., & Simpson, H. (2015). Place-based policies. In *Handbook of regional and urban economics* (pp. 1197–1287, Vol. 5). Elsevier.
- Nguyen, T. T. A., Luu, M. D., & Trinh, D. C. (2016). The evolution of Vietnamese industry. In J. Page, F. Tarp, J. Rand, A. Shimeles, C. Newman, & M. Söderbom (Eds.), *Manufacturing transformation: Comparative studies of industrial development in Africa and emerging Asia*. Oxford University Press.
- Pham, T. (2026). *Who benefits from place-based industrial policies? labor market adjustments and household welfare in Vietnam* (Working Paper).
- Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5), 2555–2591.
- Robinson, L. A., Hammitt, J. K., & O’Keeffe, L. (2019). Valuing mortality risk reductions in global benefit-cost analysis. *Journal of Benefit-Cost Analysis*, 10(S1), 15–50.
- Roth, J., Sant’Anna, P. H., Bilinski, A., & Poe, J. (2023). What’s trending in difference-in-differences? a synthesis of the recent econometrics literature. *Journal of Econometrics*, 235(2), 2218–2244.
- Shapiro, J. S., & Walker, R. (2018). Why is pollution from US manufacturing declining? the roles of environmental regulation, productivity, and trade. *American Economic Review*, 108(12), 3814–3854.
- Shen, S., Li, C., Van Donkelaar, A., Jacobs, N., Wang, C., & Martin, R. V. (2024). Enhancing global estimation of fine particulate matter concentrations by including geophysical a priori information in deep learning. *Acs Es&T Air*, 1(5), 332–345.
- Silva, J. S., & Tenreyro, S. (2006). The log of gravity. *Review of Economics and statistics*, 641–658.
- Tafese, T., Lay, J., & Tran, V. (2025). From fields to factories: Special economic zones, foreign direct investment, and labour markets in vietnam. *Journal of Development Economics*, 103467.
- Truong, B. T., & Le, Q. T. (2016). *Assessing Vietnam’s tax incentive policies* (tech. rep.). Oxfam in Vietnam, Hanoi. <https://vietnam.oxfam.org/latest/publications/report-assessing-vietnams-tax-incentive-policies>
- UNCTAD. (2019). World investment report 2019: Special economic zones. https://unctad.org/system/files/official-document/wir2019_en.pdf
- United Nations Department of Economic and Social Affairs. (2022). *2022 united nations energy statistics yearbook* [Table II: Selected Conversion Factors for Crude Oil and Oil Products]. United Nations.

- Viscusi, W. K., & Masterman, C. J. (2017). Income elasticities and global values of a statistical life. *Journal of Benefit-Cost Analysis*, 8(2), 226–250.
- Wang, G., Zhang, R., Gomez, M. E., Yang, L., Levy Zamora, M., Hu, M., Lin, Y., Peng, J., Guo, S., Meng, J., et al. (2016). Persistent sulfate formation from London Fog to Chinese haze. *Proceedings of the National Academy of Sciences*, 113(48), 13630–13635.
- Wang, J. (2013). The economic impact of special economic zones: Evidence from Chinese municipalities. *Journal of Development Economics*, 101, 133–147.
- Yang, L., Lin, Y., Wang, J., & Peng, F. (2024). Achieving air pollution control targets with technology-aided monitoring: Better enforcement or localized efforts? *American Economic Journal: Economic Policy*, 16(4), 280–315.

Supplementary Materials for Online Publication

The Environmental Cost of Place-Based Industrial Policy: Evidence from Industrial Zones in Vietnam

Trinh Pham

KDI School of Public Policy and Management

A APPENDIX TABLES AND FIGURES

A.1 Appendix Tables

Table A1: Origins of Manufacturing Firms Arriving in Districts near Industrial Zones

	Arrivals (count)	Movers (%)	From comparison districts within 50 km (%)	From another zone ring (%)	From another province (%)	Movers' share of employment (%)
<i>Panel A: All firms</i>						
Districts newly within 5 km of an operating zone, after establishment						
Pollution-intensive industries	2927	9.4	1.7	7.0	1.6	18.3
Light industries	15099	4.6	1.0	3.3	1.3	10.1
Districts never within 5 km of a zone						
Pollution-intensive industries	3777	6.6	1.4	4.5	1.1	17.3
Light industries	23793	6.9	2.0	4.2	1.4	16.4
<i>Panel B: Firms with 10 or more workers at arrival</i>						
Districts newly within 5 km of an operating zone, after establishment						
Pollution-intensive industries	1668	14.1	2.3	10.9	2.4	19.2
Light industries	6391	7.3	1.6	5.2	2.2	10.7
Districts never within 5 km of a zone						
Pollution-intensive industries	1905	9.4	2.3	6.0	1.8	18.4
Light industries	8425	10.8	3.4	6.3	2.7	17.6

Notes: The table follows manufacturing firms across districts in the Vietnam Enterprise Survey 2001–2018 using the enterprise tax code, whose first nine digits are constant across survey rounds. An arrival is the first year a firm is observed in a district. A mover is an arrival that was observed in a different district in an earlier year, and its origin is the district of that last earlier observation. Other arrivals are first observations in the survey. Districts newly within 5 km of an operating zone are those first within 5 km between 2003 and 2016, the cohorts of equation (1), and their arrivals are counted from the year of establishment onward. Movers “from comparison districts within 50 km” have an origin district that is never within 5 km of a zone and lies within 50 km of one. Movers “from another zone ring” have an origin district that was already within 5 km of an operating zone. Panel A includes all firms. Panel B restricts arrivals to firms with 10 or more workers, which the survey enumerates in every round together with all state and foreign-invested enterprises, whereas smaller private enterprises are covered on a sample basis for the full questionnaire in some rounds. Pollution-intensive industries follow the classification of Appendix B. The mover shares are lower bounds on relocation: a firm that closes in one district and re-registers under a new tax code in another is counted as a new arrival rather than a mover. Excluding enterprises ever listed with units in more than one district (0.3% of firms) lowers the shares in Panel A for districts newly within 5 km to 8.7 and 4.3%.

Table A2: Robustness Checks: Industrial Zones and Type-2 Thermal Anomalies (Days)

	Baseline	Donut - (5-10) km	Treatment 5 km, operation	Treatment 10 km
	(1)	(2)	(3)	(4)
Post-ATT (marginal effect)	0.422 (0.169)	0.498 (0.206)	0.279 (0.134)	0.314 (0.136)
pre-trend p-value	0.812	0.851	0.542	0.531
Observations	11362	8588	11362	9690
Cell FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Model	Poisson	Poisson	Poisson	Poisson

Notes: The table reports robustness checks for the Poisson estimates of column (4) of Table 3. Column (1) reproduces that specification: treatment is a cell being within 5 km of an operational industrial zone from the year before its recorded operational date, with cells never within 5 km of an operational zone during the study period as the comparison group. Column (2) excludes cells in the 5–10 km ring from the sample. Column (3) dates treatment from the year of recorded operational status rather than the year before. Column (4) defines treatment as a cell being within 10 km of an operational zone (from the year before operation). Post-ATT, marginal effects, and the pre-trend p-value are defined as in Table 3. Province-clustered standard errors in parentheses. Data are from the NASA FIRMS MODIS Collection 6.1 Active Fire Product, 2000–2020.

Table A3: Industrial Zones and Type-2 Thermal Anomalies (Days): Multiple Zones around a Cell

Panel A: Restricting the sample (event study)

	Baseline	Treated cells: one zone within 5 km	Comparison cells: no zone within 10 km	Comparison cells: no zone within 25 km
	(1)	(2)	(3)	(4)
Post-ATT (marginal effect)	0.422 (0.169)	0.352 (0.161)	0.498 (0.206)	0.400 (0.227)
pre-trend p-value	0.812	0.197	0.851	0.904
Observations	11362	10811	8588	4617
Treated cells	108	79	108	108
Comparison cells	490	490	344	135
Cell FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Model	Poisson	Poisson	Poisson	Poisson

Panel B: Exposure to all zones (pooled Poisson)

	(1)	(2)	(3)
Post $\times$ one operating zone within 5 km	0.765 (0.326)		
Post $\times$ two or more operating zones within 5 km	0.584 (0.483)		
Post (any operating zone within 5 km)		0.712 (0.325)	
Distance-weighted number of operating zones 5–25 km away (per s.d.)		0.067 (0.157)	
Distance-weighted number of operating zones within 25 km (per s.d.)			0.116 (0.122)
Observations	11362	11362	11362
Cell FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Model	Poisson	Poisson	Poisson

Notes: The table examines whether the event-study estimates of Table 3, column (4), depend on the number and proximity of the zones around a cell. Treatment is being within 5 km of any operating zone in the year before its recorded operational date. Panel A re-estimates the event study on restricted samples: column (2) keeps treated cells that never have more than one zone within 5 km, and columns (3) and (4) keep comparison cells that never have an operating zone within 10 or 25 km. Post-ATT marginal effects, pre-trend tests, and standard errors are computed as in Table 3. Panel B replaces the event-time indicators with a post-treatment indicator on the baseline sample and reports Poisson coefficients. Column (1) splits the indicator by the number of operating zones within 5 km of the cell in that year. Columns (2) and (3) use the number of operating zones in each 5 km band around the cell, weighted by the inverse of the band midpoint (2.5, 7.5, 12.5, 17.5, and 22.5 km) and standardized to unit variance, over 5–25 km in column (2), alongside the 5 km indicator, and over 0–25 km in column (3), in place of it. Province-clustered standard errors in parentheses.

Table A4: Directional Wind Exposure and Pre-Treatment Characteristics

	Year 2002				Year 2000		Pre-treatment
	Urban Share	Minority Share	Secondary Education Share	Farm Labor Share	Population Share	Elevation	ln(PM2.5)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Downwind Exposure (SD)	0.004 (0.034)	0.001 (0.008)	0.009 (0.007)	-0.018 (0.013)	0.000 (0.000)	10.521 (9.097)	-0.004 (0.014)
Observations	1071	1071	1071	1071	1071	1071	636
Adjusted R-squared	0.143	0.556	0.280	0.328	0.404	0.763	0.952
Province FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Distance Band FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The table reports OLS regressions of pre-treatment characteristics on the standardized downwind exposure variable E_{imt} defined in equation (5). The unit of observation is a district $\times$ distance band cell; outcomes measured at the cell level are averaged within district $\times$ distance band. The sample comprises district $\times$ distance band observations within 25 km of any industrial zone, excluding the innermost 0–5 km ring. In columns (1)–(4), the outcomes are the four district characteristics listed in the notes to Table 3, from the 2002 VHLSS. In Column (5), the outcome is the cell-level share of its province’s total population in 2000. In Column (6), the outcome is cell-level elevation in meters. In Column (7), the outcome is the cell-level mean of log PM2.5 across all panel years strictly before the cell’s first industrial zone within 25 km, with PM2.5 from Shen et al. (2024). All regressions include province and distance band fixed effects. Province-clustered standard errors in parentheses.

Table A5: Industrial Zone Characteristics by Establishment Cohort

	Pre-2003	2003–2009	Post-2009
Median planned area (ha)	197	249	262
Median leased area (ha)	123	98	35
Median occupancy rate	0.997	0.901	0.297
Share with occupancy below 0.25	0.011	0.140	0.448
Share under construction in 2023	0.011	0.124	0.664

Notes: The table reports characteristics of industrial zones by establishment cohort, for the 395 zones in the analysis sample with a recorded operation year: 92 established before 2003, 178 in 2003 to 2009, and 125 after 2009, where a zone's establishment year is its treatment year, the year before recorded operation, as throughout the paper. Areas are in hectares; leased area is the area occupied by tenant firms. Occupancy rates, leased areas, and construction status come from the 2023 vintage of the Ministry of Planning and Investment zone database and therefore reflect conditions three years after the end of the study period; occupancy among post-2009 zones during the study period is lower than reported here. Under construction denotes zones whose 2023 status is construction rather than operation. *Source:* Data from the Ministry of Planning and Investment zone database (2023).

Table A6: Technique Margins of Firms near Pre-2009 Zones after the Emission Standards

	Pre-2003 cohort ×		2003–2009 cohort ×		Joint p-value
	$\mathbf{1}(t \geq 2011)$	$\mathbf{1}(t \geq 2014)$	$\mathbf{1}(t \geq 2011)$	$\mathbf{1}(t \geq 2014)$	$(t \geq 2014)$
Fossil fuel share of energy use	-0.006 (0.012)	-0.010 (0.008)	-0.005 (0.022)	-0.005 (0.015)	0.42
Any fossil fuel use	0.013 (0.028)	-0.024 (0.017)	-0.006 (0.019)	0.010 (0.021)	0.15
Fossil fuel TCE per unit of output	0.0006 (0.0006)	0.0002 (0.0004)	-0.0008 (0.0016)	-0.0000 (0.0005)	0.74
ln(total TCE per unit of output)	0.035 (0.066)	-0.023 (0.050)	-0.100 (0.093)	-0.022 (0.082)	0.87

Notes: Each row is one regression of the outcome on zone-cohort indicators, their interactions with indicators for the years from 2011 and from 2014 onward, and indicators for firm-size class (under 50, 50 to 99, 100 to 199, and 200 or more employees), with province and industry-by-year fixed effects and employment weights, on the VES energy rounds of 2009, 2011, and 2013–2015 (625885 firm-years for the first two outcomes and 611999 for the last two). Industry fixed effects are at the two-digit level of the Vietnamese industrial classification. The omitted category is firms in never-treated districts; the interactions for the post-2009 cohort are included but not reported. The first three outcomes are defined in the notes to Appendix Figure A8; ln(total TCE per unit of output) is the log of total energy consumption in TCE per million dong of gross turnover in 2010 prices, for firms reporting positive gross turnover. The years from 2011 are those in which the 2009 technical emission standards were in force, and the years from 2014 those after the 2013 ambient PM2.5 standard and the 2014 revision of the Law on Environmental Protection. The last column reports the p-value of the joint test that the two 2014-onward interactions for the pre-2009 cohorts are zero. Province-clustered standard errors in parentheses.

Table A7: Fuel-Use Margins of Firms near Zones, by Establishment Cohort

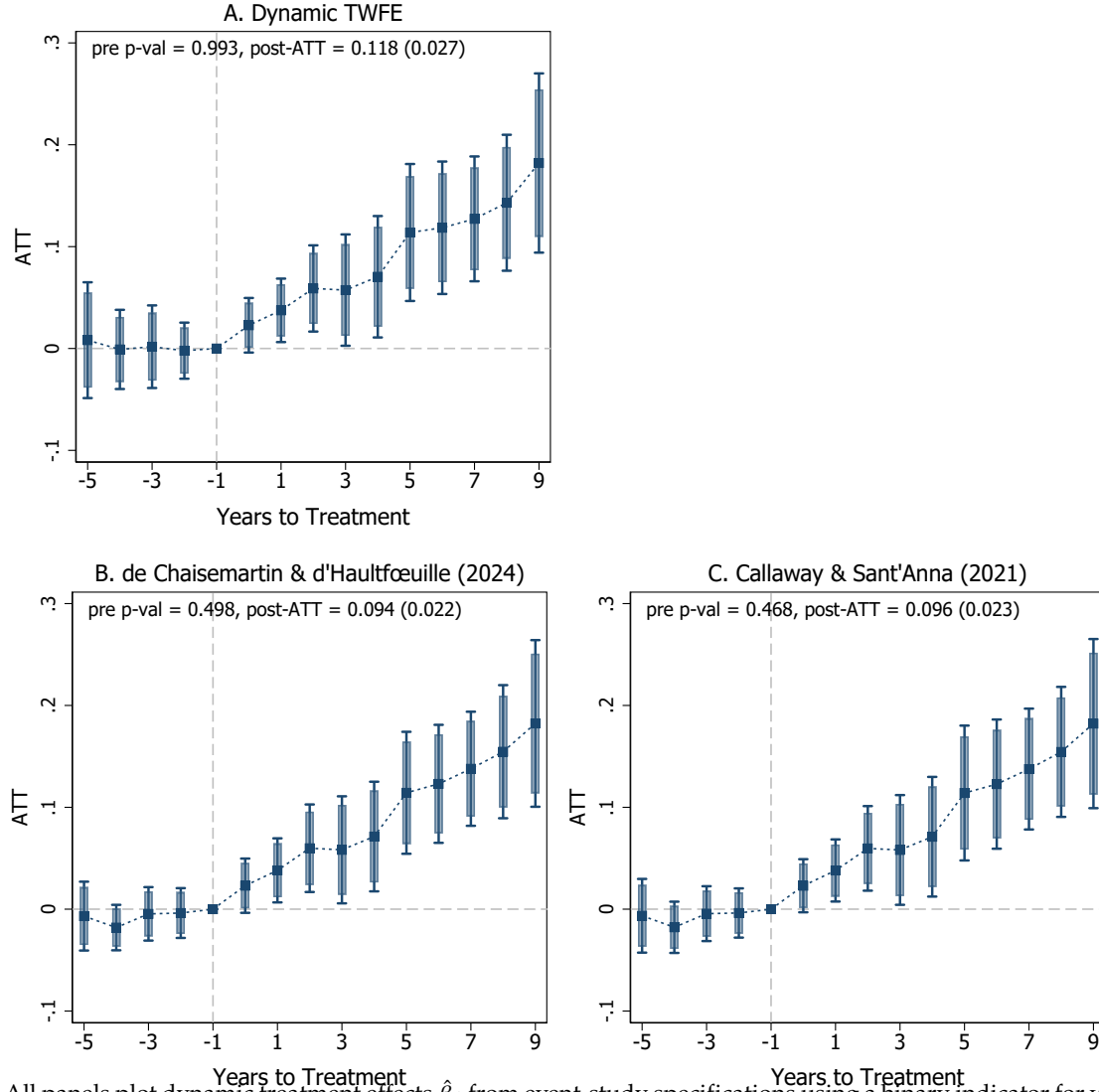
	Fossil fuel share (1)	Any fossil fuel use (2)	Fossil TCE per output (3)	ln(Total TCE per output) (4)
Pre-2003 zone cohort	0.006 (0.008)	0.013 (0.015)	-0.0012 (0.0005)	-0.013 (0.081)
2003–2009 zone cohort	0.018 (0.015)	0.055 (0.019)	0.0005 (0.0005)	0.067 (0.071)
Post-2009 zone cohort	0.008 (0.017)	0.048 (0.017)	0.0019 (0.0021)	0.238 (0.118)
Observations	625885	625885	611999	611999
Province FE	Yes	Yes	Yes	Yes
Industry $\times$ Year FE	Yes	Yes	Yes	Yes
Firm-size controls	Yes	Yes	Yes	Yes

Notes: The table reports the coefficients on the zone-cohort indicators from the regressions of Appendix Table A6 without the period interactions, which are also the regressions of Appendix Figure A8. Outcomes, fixed effects, firm-size indicators, weights, and sample are as in Appendix Table A6, with firms in never-treated districts as the omitted category. The employment-weighted never-treated means are 0.59 for any fossil fuel use and 0.0047 for fossil TCE per unit of output. Differences within two-digit industries may reflect product mix as well as technology. Province-clustered standard errors in parentheses.

A.2 Appendix Figures

Figure A1: Industrial Zones and Type-2 Thermal Anomalies: Linear Estimates

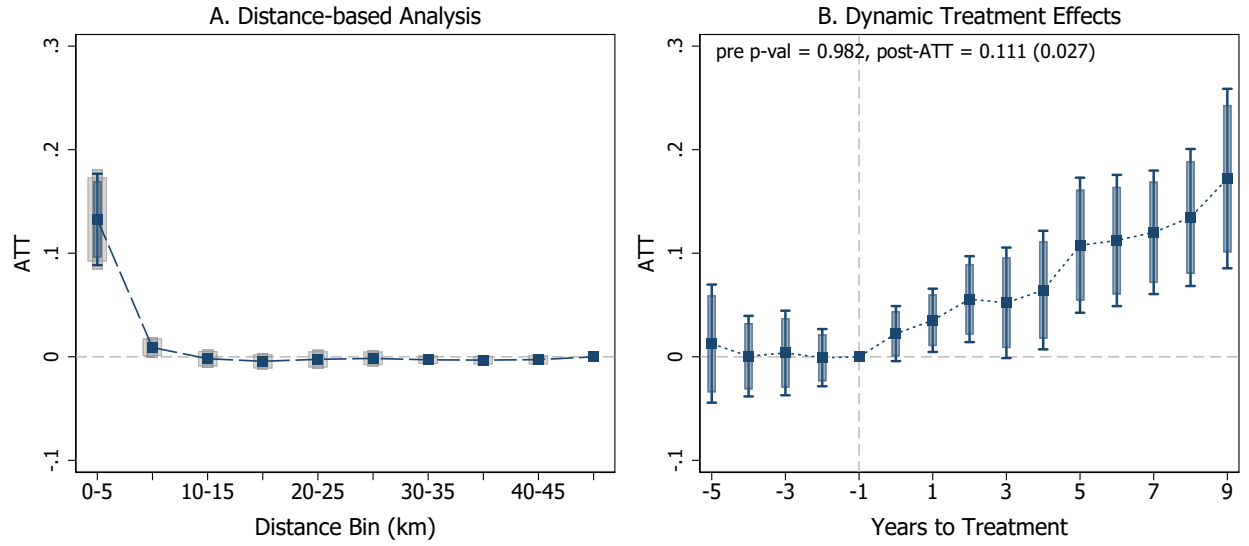
Robustness to heterogeneity-robust estimators



Notes: All panels plot dynamic treatment effects $\hat{\beta}_\ell$ from event-study specifications using a binary indicator for whether a cell is within 5 km of an operational industrial zone, with event time ℓ relative to first treatment. Outcome is $\ln(1 + \text{Type-2 detections})$ at the cell-year level. Panel A uses a two-way fixed-effects specification with event-time indicators pooled across cohorts and never-treated cells as the comparison group, with the post-ATT computed as a weighted average of post-treatment event-time coefficients, with weights proportional to the number of treated cells at each event-time. Panels B and C use the heterogeneity-robust estimator of de Chaisemartin and d'Haultfoeuille (2024) and Callaway and Sant'Anna (2021), which aggregate cohort-specific dynamic effects with non-negative weights. All specifications include cell fixed effects and year fixed effects, and endpoints $\ell = -5$ and $\ell = 9$ bin all event times outside the plotted range. Navy markers are point estimates, bars 90% and caps 95% confidence intervals from province-clustered standard errors. Source: Data are from the NASA FIRMS MODIS Collection 6.1 Active Fire Product, 2000–2020.

Figure A2: Industrial Zones and Type-2 Thermal Anomalies: Linear Estimates

Robustness to including time FE interacted with baseline characteristics



Notes: The left panel plots the weighted-average distance-bin coefficients $\bar{\delta}_b$ from equation (1), with 0–5 km through 40–45 km bins and 45–50 km as the reference. The right panel plots the dynamic treatment effects $\hat{\beta}_\ell$ using a binary indicator for whether a cell is within 5 km of an operational industrial zone, with event time ℓ relative to first treatment. Outcome is $\ln(1 + \text{Type-2 detections})$ at the cell-year level. Navy markers are point estimates, bars 90% and caps 95% confidence intervals from analytical province-clustered standard errors. Light shaded bands in the left panel are 90 and 95% confidence intervals from a stratified cluster bootstrap with 200 replications. *Source:* Data are from the NASA FIRMS MODIS Collection 6.1 Active Fire Product, 2000–2020.

Figure A3: Industrial Zones and Type-2 Thermal Anomalies: Poisson Event-Study

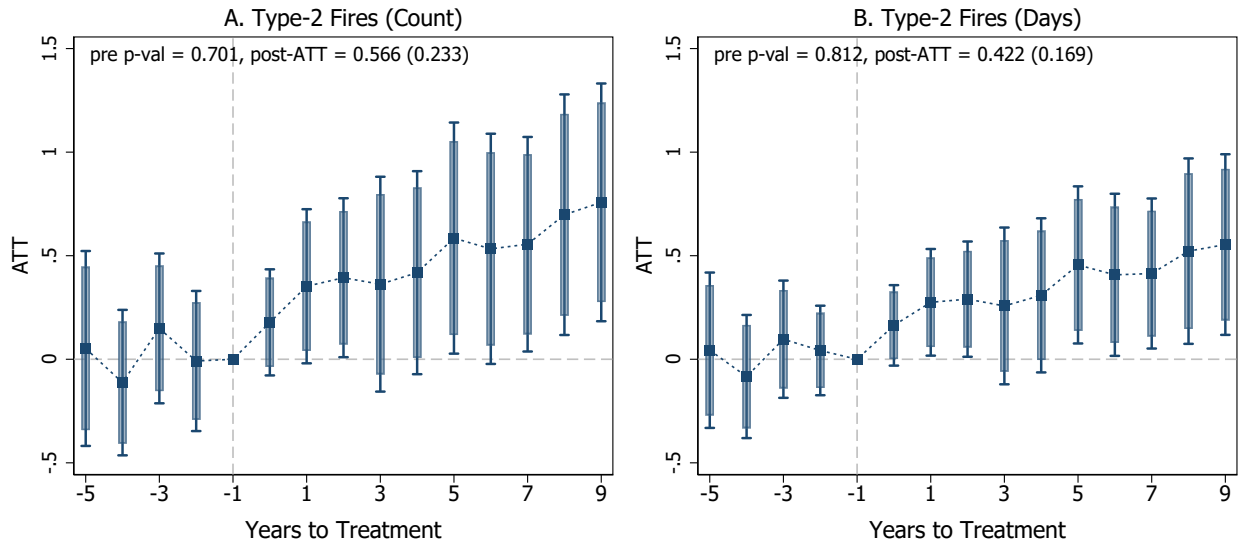
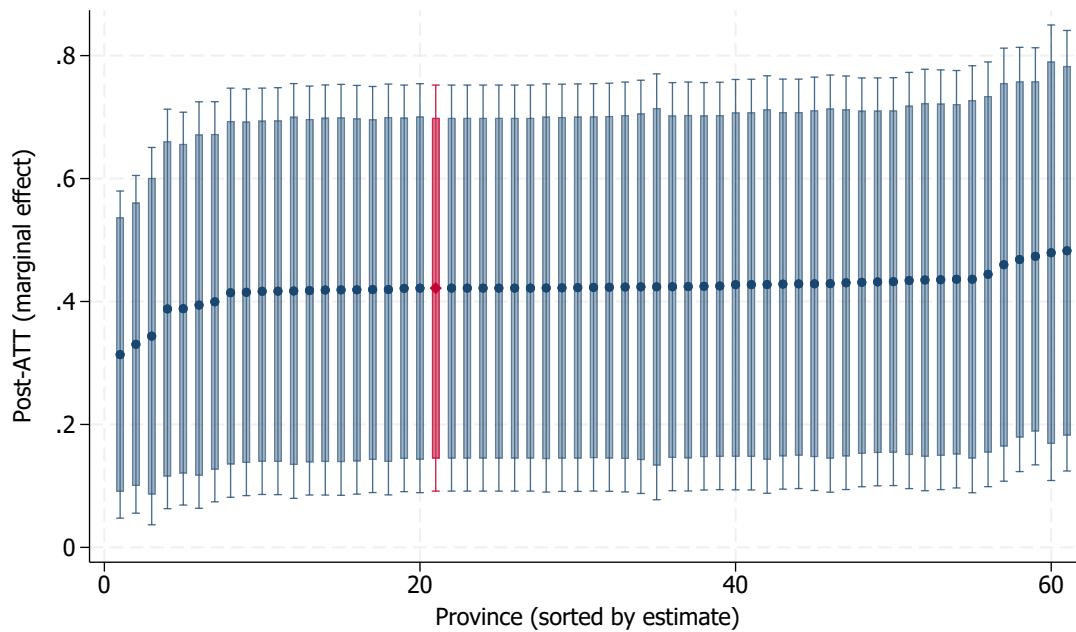


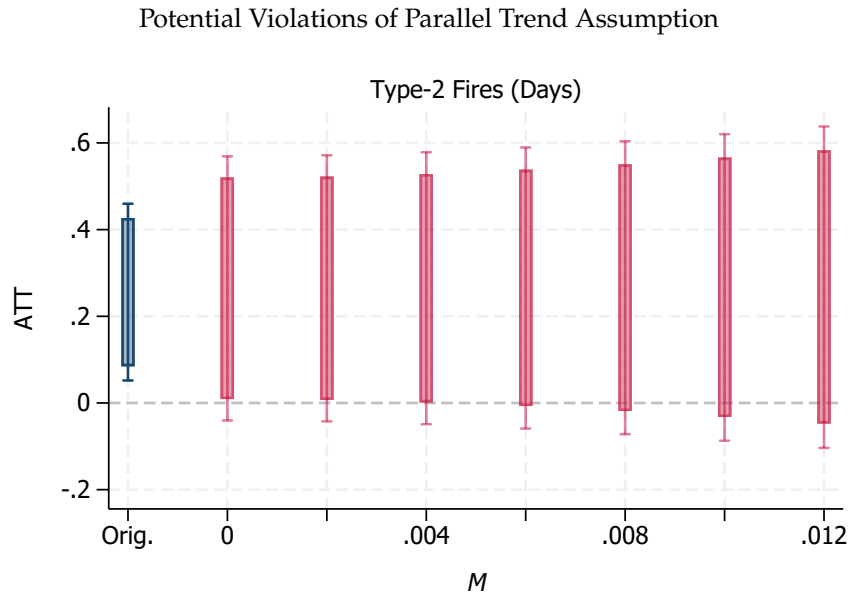
Figure A4: Industrial Zones and Type-2 Thermal Anomalies (Days)

Leave-One-Province-Out Distribution of Post-ATT Estimates



Notes: The figure reports leave-one-province-out estimates of the post-treatment marginal effect of industrial zones on Type-2 thermal anomaly days. Each point on the x-axis corresponds to a separate Poisson event-study regression in which one province is dropped from the sample, with provinces sorted along the x-axis by the size of the resulting estimate. Navy markers are point estimates, bars 90% and caps 95% confidence intervals. The cranberry marker shows the baseline estimate using the full sample (column 4 of Table 3). All specifications are those of column (4) of Table 3, with cell and year fixed effects and province-clustered standard errors. Post-ATT is defined as in Table 3. *Source:* Data are from the NASA FIRMS MODIS Collection 6.1 Active Fire Product, 2000–2020.

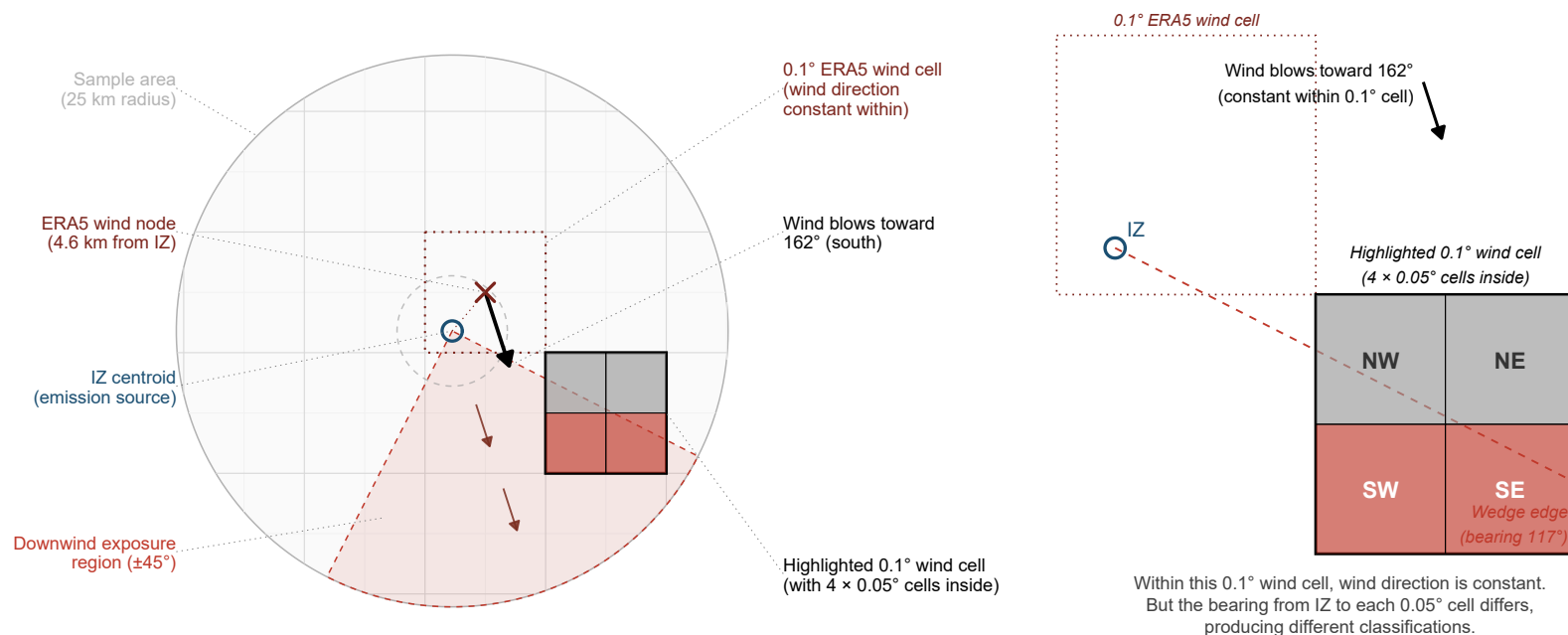
Figure A5: Industrial Zones and Type-2 Thermal Anomalies (Days)



Notes: The figure reports the sensitivity analysis of Rambachan and Roth (2023) for the post-treatment effect at event time $\ell = 2$ in the days-with-fire specification. The leftmost bar (“Orig.”) shows the original 90/95% confidence interval from the baseline Poisson event-study estimate. The remaining bars show 90/95% robust confidence intervals under the smoothness restriction $\Delta^{SD}(M)$, which bounds deviations from linearity to no more than M between consecutive periods. $M = 0$ allows for linear violations of parallel trends, and larger values of M allow for greater deviations from linearity. The post-treatment effect remains positive and statistically significant at the 10% level for M up to approximately 0.005. The baseline specification is the preferred Poisson estimate from column (5) of Table 3, with cell fixed effects and year fixed effects. Standard errors are clustered at the province level. *Source:* Data are from the NASA FIRMS MODIS Collection 6.1 Active Fire Product, 2000–2020.

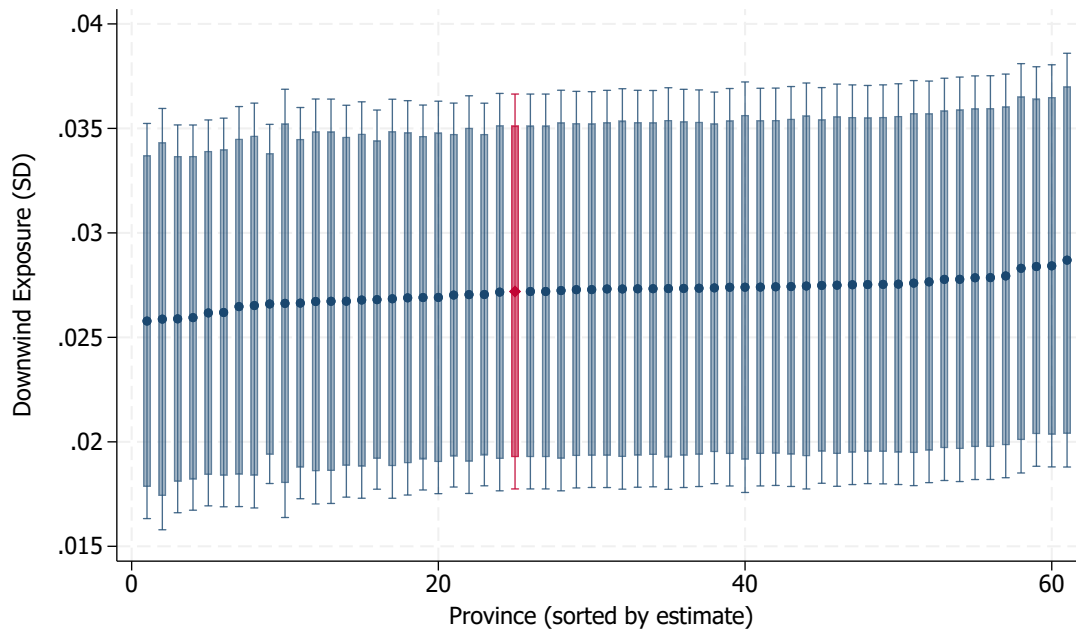
Figure A6: Wind-Based Downwind Classification: Schematic at 0.1° and 0.05° Resolutions

A. Overview at 25 km scale: 0.1° wind cells (red) and 0.05° analysis cells (inside highlighted cell) B. Wind direction is constant at 0.1°, classification varies at 0.05°



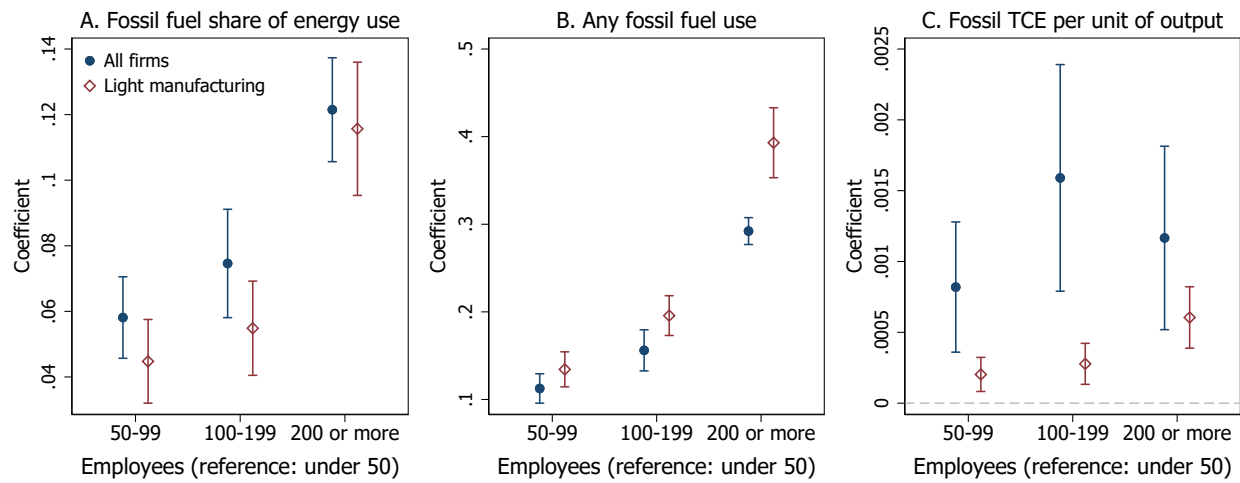
Notes: The figure illustrates the construction of downwind exposure when the ERA5-Land wind grid (0.1°) is coarser than the analysis grid (0.05°). Panel A shows the 25 km sample area around a representative industrial zone (IZ), whose centroid is the emission source. The ERA5 wind node is the nearest 0.1° grid node to the centroid, 4.6 km away in this schematic; the 0.1° wind cell containing both is outlined with a red dotted line, and the 0.1° and 0.05° grids are drawn faintly. The downwind region is the $\pm 45^\circ$ wedge anchored at the centroid and oriented in the direction the wind blows toward (162° here, approximately due south). The highlighted 0.1° wind cell in the southeast contains four 0.05° analysis cells, color-coded by downwind classification. Panel B zooms in on that cell, with the zone's own wind cell shown for context in the upper left. The wind direction is constant within a 0.1° cell, but the bearing from the centroid to each of the four analysis cells differs, and the wedge edge (bearing $117^\circ = 162 - 45$) therefore splits them into two downwind (SW, SE) and two upwind (NW, NE) cells. All geometric parameters (the offset between the centroid and the ERA5 node, the wind direction, and the position of the highlighted cell) are chosen for visual clarity rather than drawn from data.

Figure A7: Industrial Zones and PM2.5: Leave-One-Province-Out Estimates



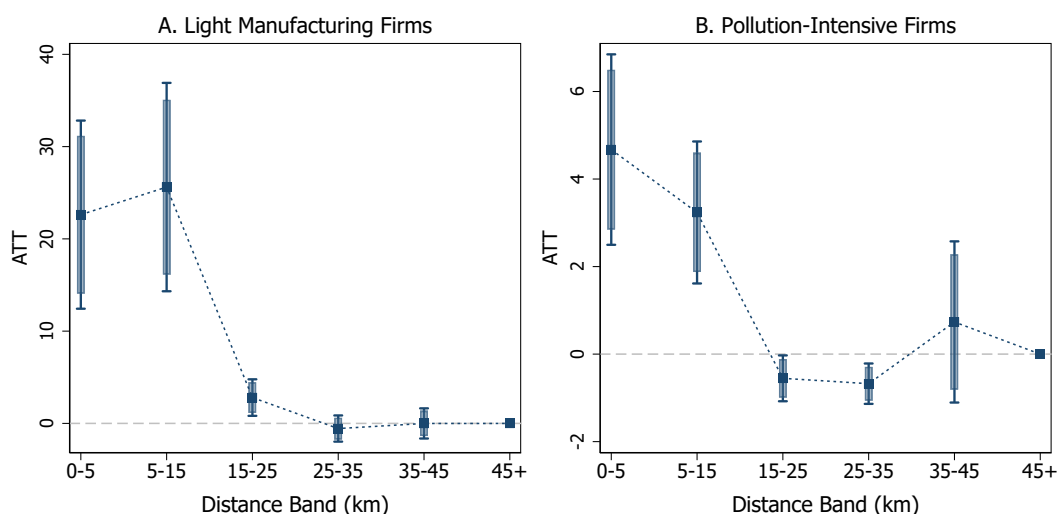
Notes: The figure reports leave-one-province-out estimates of the effect of downwind exposure to industrial zones on $\ln(\text{PM}_{2.5})$. Each point on the x-axis corresponds to a separate regression of equation (4) in which one province is dropped from the sample, with provinces sorted along the x-axis by the size of the resulting estimate. The exposure measure is re-standardized within each leave-one-out sample. Navy markers are point estimates, bars 90% and caps 95% confidence intervals. The cranberry marker shows the baseline estimate using the full sample (column 2 of Table 4). All specifications are those of column (2) of Table 4, with province-clustered standard errors. *Source:* Data are from Shen et al. (2024) for PM_{2.5}, ERA5-Land for wind (Muñoz-Sabater et al., 2021), and the Ministry of Planning and Investment for industrial zone locations.

Figure A8: Fossil Fuel Use by Firm Size within Industry, Year, and Province



Notes: Each panel plots coefficients on firm-size class indicators, relative to firms with fewer than 50 employees, from firm-level regressions of a fuel-use margin on size-class and zone-cohort indicators with province and industry-by-year fixed effects, weighted by firm employment. Bars are 95% confidence intervals from standard errors clustered at the province level. Circles use all firms in the energy module and diamonds manufacturing firms outside the pollution-intensive industries (Appendix B). The fossil fuel share is the ratio of energy from coal, briquettes, gasoline, kerosene, mazut, diesel, and LPG to total energy in tons of coal equivalent (TCE), converted using net calorific values from Garg et al. (2006) and volume-to-mass factors from United Nations Department of Economic and Social Affairs (2022). Any fossil fuel use is an indicator for positive fossil fuel consumption, defined for firms with positive energy consumption. Fossil TCE per unit of output is fossil fuel consumption in TCE per million dong of gross turnover in 2010 prices, winsorized at the top 2.5% within two-digit industry; this margin restricts the sample to firms reporting positive gross turnover, with revenue deflated with the World Bank CPI (2010 = 100). Industry fixed effects are at the two-digit level of the Vietnamese industrial classification. The zone-cohort coefficients from the same regressions are in Appendix Table A7. Firm-level data are from the Vietnam Enterprise Survey energy module for 2009, 2011, and 2013–2015.

Figure A9: Industrial Zone Establishment and Manufacturing Firm Counts by Distance Band



Notes: The figure estimates equation (1) with the district as the unit of analysis and the number of light (Panel A) or pollution-intensive (Panel B) manufacturing firms as the outcome, in levels. For each two-year cohort, treated districts are those first within 5 km of an operating zone in the cohort years, comparison districts are those never within 5 km of any zone, assigned to bands by the distance to their nearest zone, and the reference band is 45 km and beyond. Bands beyond 5 km are 10 km wide, because districts are larger than the cells of Section 4, so that each band contains at least 34 comparison districts. Coefficients are averaged across cohorts with weights proportional to the number of treated districts, as in equation (2), and standard errors are clustered at the province level. Classification, sample years (2001–2018), and data follow Figure 6. Outer bars 95% and inner bars 90% confidence intervals. Source: Data are from the Vietnam Enterprise Survey and the Ministry of Planning and Investment for industrial zone locations.

B POLLUTION-INTENSITY CLASSIFICATION FROM THE VIETNAM ENTERPRISE SURVEY

This appendix describes the construction of the pollution-intensity classification used in the paper and compares the resulting industry ranking with the U.S.-based classification of Shapiro and Walker (2018).

B.1 Construction

The classification is built from the energy module of the Vietnam Enterprise Survey, which reports physical quantities of nine energy types (electricity, coal, briquettes, gasoline, kerosene, mazut, diesel, LPG, and natural gas) consumed for production. I use the 2009, 2011, and 2013–2015 survey rounds (Section 3) and convert quantities to tons of coal equivalent (TCE) using default net calorific values from Garg et al. (2006) and volume-to-mass conversion factors from United Nations Department of Economic and Social Affairs (2022). Fossil energy is total TCE excluding electricity. The measure therefore captures on-site combustion, the process that generates the thermal anomalies and primary particulate emissions studied in the paper.

For each firm, pollution intensity is fossil TCE per unit of output, measured as gross turnover deflated to 2010 prices with the World Bank CPI series. I drop observations with nonpositive revenue and winsorize the firm-level intensity ratios at the top and bottom 2.5% within each industry, pooling across survey rounds. The industry-level intensity is the revenue-weighted mean of the winsorized firm ratios, equivalent to industry fossil TCE per unit of industry output. The estimation sample comprises approximately 141,000 manufacturing firm-year observations.

The classification unit is a set of 20 manufacturing categories that can be measured consistently over the full 2001–2018 firm panel, spanning the change in Vietnam’s industrial classification from VSIC 1993 to VSIC 2007.¹ Pollution-intensive industries are the top quartile of the intensity ranking, five of the 20 categories: non-metallic mineral products, paper and paper products, coke and refined petroleum products, basic metals, and textiles. The ranking is insensitive to construction choices: rank correlations between the baseline intensity and variants using 1% winsorization tails, unweighted means, unwinsorized ratios of sums, or intensity per worker range from 0.84 to 0.96.

B.2 Comparison with the U.S.-based classification

Table B1 reports the full ranking alongside the emissions intensity of Shapiro and Walker (2018), which is estimated from U.S. data and widely used in the literature, and therefore provides an external benchmark. The comparison shows two features. First, the four highest-ranked categories under the Vietnamese energy data, non-metallic minerals, paper, coke and refined petroleum, and basic metals, are the core of the U.S.-based classification. Second, the classifications differ below the top. The U.S.-based intensities are defined for broader sector groups: printing and

¹Recycling (VSIC 1993) and repair and installation of machinery (VSIC 2007) do not correspond to a consistent industry across the two classification vintages and are left unclassified.

recorded media is classified as pollution-intensive only because it belongs to the paper group, yet it ranks last, 20th of 20, in the Vietnamese energy data. Drugs and chemicals belongs to the chemicals group but ranks 8th, reflecting that chemical production pollutes primarily through process emissions and water effluent rather than fuel combustion, the margin relevant for the combustion-based outcomes studied here. Conversely, textiles ranks 5th in Vietnam, consistent with the coal-fired boilers used in dyeing and finishing, and with the inclusion of dyeing among the production types with pollution risk in Appendix IIa of Decree 40/2019/ND-CP.² Overall, the rank correlation between the two intensity measures is 0.38: agreement is concentrated at the polluting tail, the part of the ranking that determines the classification, and weaker in the middle of the distribution.

The two classifications differ more for individual firms than in aggregate shares. The Vietnamese classification marks 16.5% of manufacturing firm-year observations as pollution-intensive against 24.1% under the U.S.-based classification, and the two agree for 84% of observations; the pollution-intensive share of manufacturing employment is 13 to 14% under either. Appendix Figure B1 re-estimates the firm-count event studies of Figure 6 under the U.S.-based classification and under alternative cutoffs of the Vietnamese ranking, the top four categories (dropping textiles) and the top eight (adding food and beverages, other transport equipment, and drugs and chemicals, which fall just below the quartile cutoff).

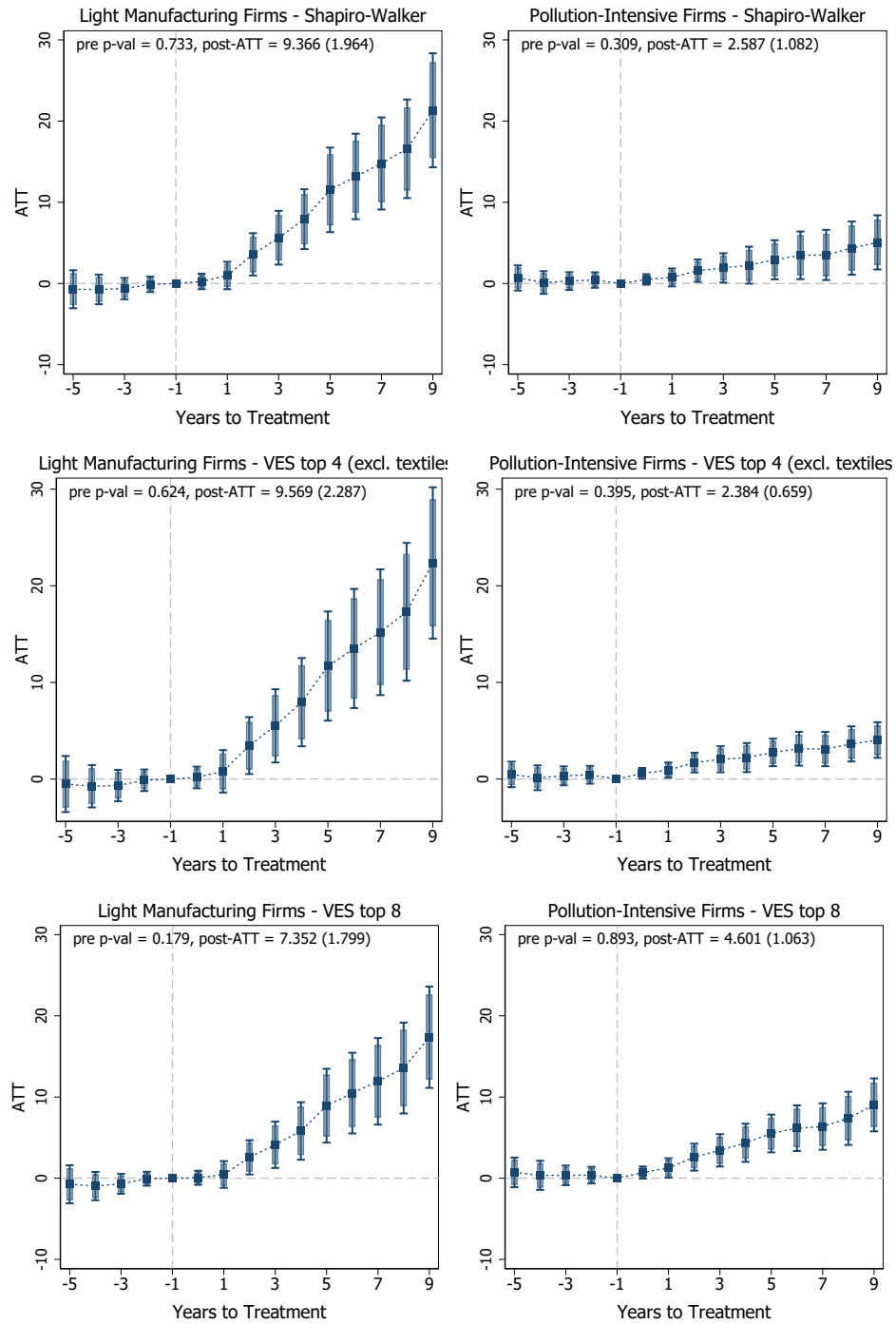
²At the more disaggregated two-digit VSIC 2007 level, available for the 2009–2015 energy rounds, beverages ranks fifth of 23 manufacturing industries, reflecting coal-fired boilers in breweries; beer and liquor production also appears in Appendix IIa of Decree 40/2019/ND-CP. In the panel-consistent categories, beverages is merged with the much larger food sector and the combined category ranks sixth.

Table B1: Pollution-Intensity Rankings: VES Energy-Based versus U.S.-Based

Manufacturing category	VES energy-based		Shapiro and Walker (2018)		Pollution-intensive	
	Intensity	Rank	Intensity	Rank	VES	SW
Non-metallic mineral products	0.0513	1	19.91	2.0	✓	✓
Paper and paper products	0.0071	2	14.61	3.5	✓	✓
Coke and refined petroleum products	0.0051	3	13.88	5.0	✓	✓
Metals	0.0037	4	36.57	1.0	✓	✓
Textiles	0.0027	5	1.44	15.0	✓	
Food and beverage	0.0022	6	2.60	10.5		
Other transport equipment	0.0021	7	1.26	17.0		
Drug and other chemical products	0.0021	8	13.42	6.0		✓
Rubber and plastic products	0.0010	9	3.13	8.0		
Wearing apparel	0.0009	10	1.44	15.0		
Fabricated metal products	0.0008	11	1.24	18.0		
Wood and wood products	0.0007	12	6.75	7.0		
Electronic, computer and optical products	0.0006	13	1.52	12.5		
Machinery and equipment nec	0.0006	14	1.00	20.0		
Leather and related products	0.0005	15	1.44	15.0		
Motor vehicles	0.0004	16	1.03	19.0		
Tobacco	0.0004	17	2.60	10.5		
Furniture and others	0.0003	18	3.06	9.0		
Electrical appliances	0.0003	19	1.52	12.5		
Recorded media	0.0002	20	14.61	3.5		✓

Notes: The table reports pollution intensity by manufacturing category for the 20 categories measured consistently over the 2001–2018 firm panel, sorted by the VES energy-based ranking. The VES intensity is fossil-fuel energy in tons of coal equivalent per million constant-2010 dong of gross output, computed from firm-level energy and revenue records in the 2009, 2011, and 2013–2015 VES rounds, winsorized at 2.5% tails within category and aggregated with revenue weights. The Shapiro and Walker (2018) intensity is emissions per unit of output estimated from U.S. data, matched to the Vietnamese manufacturing categories at the level of broader sector groups. The final two columns mark the pollution-intensive classification under each measure: VES = top quartile of the energy-based ranking (used throughout the paper); SW = the U.S.-based classification of Shapiro and Walker (2018). The Spearman rank correlation between the two intensity measures is 0.38. *Source:* Vietnam Enterprise Survey and Shapiro and Walker (2018).

Figure B1: Firm Counts under Alternative Pollution-Intensity Classifications



Notes: The figure re-estimates the firm-count event studies of Figure 6 under three alternative pollution-intensity classifications. The top row uses the U.S.-based classification of Shapiro and Walker (2018). The middle row uses the top four categories of the VES energy-based ranking (excluding textiles). The bottom row uses the top eight categories (adding food and beverages, other transport equipment, and drugs and chemicals). Specifications, estimator, and inference follow Figure 6. Source: Vietnam Enterprise Survey and Ministry of Planning and Investment.

C THE SPATIAL REACH OF ZONE EXPOSURE AND ALTERNATIVE FIXED EFFECTS

This appendix decomposes the PM2.5 estimate in Table 4, column (2), by the distance between a cell and the zones whose exposure affects it, and uses the decomposition to interpret estimates of equation (4) under alternative sets of fixed effects. Let d_{ij} denote the distance between the centroids of cells i and j , and define three distance bands around cell i : $\mathcal{N}_1(i) = \{j \neq i : d_{ij} \leq 25\}$, $\mathcal{N}_2(i) = \{j : 25 < d_{ij} \leq 50\}$, and $\mathcal{N}_3(i) = \{j : 50 < d_{ij} \leq 100\}$, with distances in km. The average downwind exposure of band k in month m of year t is

$$R_{imt}^{(k)} = \frac{1}{|\mathcal{N}_k(i)|} \sum_{j \in \mathcal{N}_k(i)} E_{jmt}, \quad k = 1, 2, 3 \quad (\text{C1})$$

where E_{jmt} is the downwind exposure of cell j from equation (5) and $|\mathcal{N}_k(i)|$ is the number of cells in the band. I divide each band average by the standard deviation that scales the own-cell measure, so that all coefficients below are per standard deviation of exposure. Adding the band averages to equation (4) gives

$$\ln \text{PM2.5}_{ipmt} = \kappa_0 E_{imt} + \sum_{k=1}^3 \kappa_k R_{imt}^{(k)} + \alpha_i + \mu_{mt} + \varepsilon_{ipmt} \quad (\text{C2})$$

where κ_k is the effect of average exposure in band k , conditional on the cell's own exposure and on exposure in the other bands.

Equation (4) omits the band averages. Its coefficient β therefore combines the own-cell effect with the band effects, weighted by the extent to which band exposure varies with own-cell exposure. Let δ_k denote the coefficient from a regression of $R_{imt}^{(k)}$ on E_{imt} with the same controls and fixed effects as equation (4). Because equations (4) and (C2) are estimated on the same sample with the same controls and fixed effects, the omitted variable bias formula (Angrist & Pischke, 2009, Section 3.2.2) holds exactly:

$$\beta = \kappa_0 + \sum_{k=1}^3 \kappa_k \delta_k \quad (\text{C3})$$

The term $\kappa_k \delta_k$ is the contribution to β of exposure in band k .

Panel A of Table C1, column (1), reports estimates of equation (C2) on the sample and with the fixed effects of Table 4, column (2). Conditional on its own exposure, a cell's PM2.5 increases with average exposure in the 0–25 km and 25–50 km bands, while the coefficient on the 50–100 km band is imprecisely estimated and average exposure in a 100–200 km band has no explanatory power when added as a fourth regressor. Under equation (C3), the estimate of 0.027 in Table 4 consists of 0.010 from the cell's own exposure, 0.012 from exposure within 25 km, 0.005 from exposure at 25–50 km, and less than 0.001 from exposure beyond 50 km. The exposure that affects a cell's PM2.5 is therefore located within about 50 km of the cell, consistent with the two distances used in the research design: the 25 km radius of the exposure measure and the 50 km outer bound of the

Table C1: PM2.5 Estimates under Alternative Fixed Effects: Decomposition by Distance Band

	Dependent Variable is $\ln(\text{PM}_{2.5})$		
	(1)	(2)	(3)
Panel A: Distance profile of exposure effects, equation (C2)			
Own cell	0.010 (0.004)	0.007 (0.002)	0.004 (0.001)
0–25 km band	0.031 (0.011)	0.024 (0.006)	0.030 (0.004)
25–50 km band	0.048 (0.015)	0.053 (0.018)	0.043 (0.008)
50–100 km band	0.023 (0.025)	0.015 (0.027)	0.060 (0.015)
Share of 25–50 km band variation absorbed	71%	77%	85%
Panel B: Estimate from equation (4)			
Downwind Exposure (SD)	0.027 (0.005)	0.017 (0.003)	0.009 (0.001)
Implied effect ($\mu\text{g}/\text{m}^3$)	0.66	0.42	0.22
Observations	793212	793212	793212
Adjusted R-squared	0.739	0.869	0.947
Cell FE	Yes	Yes	Yes
Year-Month FE	Yes		
Region $\times$ Year-Month FE		Yes	
Province $\times$ Year-Month FE			Yes
Weather Controls	Yes	Yes	Yes

Notes: The table reports estimates on the estimation sample of Table 4, column (2): cells in the 5–25 km transport ring, monthly, with the weather controls of that table and downwind exposure standardized within the sample. Column (1) includes cell and year-month fixed effects; columns (2) and (3) replace the year-month fixed effects with region-by-year-month and province-by-year-month fixed effects. Panel A reports equation (C2), which adds the average downwind exposure of the panel cells in three distance bands around each cell (equation (C1)), scaled by the same standard deviation as the own-cell measure. Band coefficients are not comparable across rows, because the bands contain different numbers of cells. The bottom row reports the share of the variation in 25–50 km band exposure absorbed by each set of fixed effects, computed as one minus the ratio of the within-transformed to the raw standard deviation. Panel B reports the coefficient on the cell's own downwind exposure from equation (4); the implied effect converts the coefficient to $\mu\text{g}/\text{m}^3$ at the sample-mean PM2.5 of 24.4 $\mu\text{g}/\text{m}^3$. Panels A and B are connected by the identity in equation (C3). Province-clustered standard errors in parentheses.

estimation panel. Alternative partitions of the bands, at 0–20/20–40/40–80 km and at 0–15/15–30/30–60 km, yield a similar pattern.

Columns (2) and (3) replace the year-month fixed effects with region-by-year-month and province-by-year-month fixed effects (7 regions and 58 provinces in the estimation sample). Panel B reports the corresponding estimates of equation (4): 0.027 per standard deviation of exposure with year-month fixed effects, 0.017 with region-by-year-month fixed effects, and 0.009 with province-by-year-month fixed effects. Equation (C3) separates two possible sources of the decline, a smaller effect of exposure at each distance (the κ_k) or a weaker relationship between band exposure and own-cell exposure (the δ_k), and Panel A indicates the latter. The band coefficients differ little across the columns. The effect of exposure at these distances is thus similar whether it is identified from comparisons across the whole country or among cells in the same province and month, which argues against a confounder operating at the region or province level as the source of the

band effects. What differs is the relationship between band exposure and own-cell exposure. Cells in the same province share most of their surroundings, and the tighter fixed effects absorb most of the variation in 25–50 km band exposure (the bottom row of Panel A). The estimate of δ_2 accordingly falls from 0.10 in column (1) to 0.04 in column (2) and -0.03 in column (3), and with the κ_k similar across columns the composite coefficient declines with it.

The coefficient in equation (4) measures the effect of zone exposure inclusive of pollution transported from surrounding zones whose exposure varies with the cell's own, which is the object relevant for the local externality the paper studies. Region-by-year-month and province-by-year-month fixed effects remove this transported component from the comparison, and their lower coefficients likely understate the full local effect rather than correct it. The effect of a one-standard-deviation increase in exposure lies between 0.017 and 0.027 log points, or 0.42 to 0.66 $\mu\text{g}/\text{m}^3$ at the sample mean. The province-by-year-month estimate of 0.009 sits below this range because its δ_2 is negative.

D COMPOSITION OF FIRMS INSIDE ZONES OVER TIME

Section 6.1 describes the tenants of each cohort's zones. This appendix asks whether the mix of firms inside zones changed over the years the data cover and whether the PM2.5 effect of exposure changed where it did. In its 2010 to 2016 rounds, the VES asks each firm whether it is located in a zone. Because the question covers all zone types and is answered by the firm itself, the reported in-zone firms span a broader set of zones than the empirical analysis. I use the 2011 and 2013–2016 rounds, in which the question was asked of nearly all firms, and keep manufacturing firms reporting a zone location. Pollution-intensive industries are defined as in the rest of the paper (Appendix B).

Panel A of Table D1 reports the national composition of in-zone manufacturing by year. The mix of firms inside zones did not shift toward light manufacturing over the period the data cover. The pollution-intensive share is 0.20 of firms in 2011 and 0.19 in 2016, and about 0.10 of employment in both years. The share changes little among firms with 100 or more workers and declines slightly among smaller firms. The industries whose processes run hot enough to register as satellite thermal anomalies (non-metallic minerals, basic metals, and coke and refined petroleum) account for about a tenth of in-zone manufacturing firms and 3 to 4% of their employment in every year (the last two rows of Panel A).

Panel B tests whether the PM2.5 effect of exposure declined where the in-zone mix did change. The VES records that a firm is in a zone but not which one, and the unit of analysis is therefore the district. The sample comprises the 94 districts hosting zones that were operational by 2009 and with at least five in-zone manufacturing firms in both 2011 and 2016, the first and last rounds in which the zone question was asked of nearly all firms. I measure the shift as the 2011–2016 change in the predicted fuel use per worker of a district's in-zone manufacturing firms, where predicted fuel use is employment multiplied by the fossil-fuel intensity per worker of the firm's two-digit industry (Appendix B). Unlike a pollution-intensive share, this measure gives every industry its own weight, and expressed per worker it reflects the composition of in-zone employment rather than its size. I sort districts into terciles of the change. Because the terciles differ in the level of the effect, the comparison is a difference-in-differences. For each tercile, I assign cells to the district hosting their nearest zone and, on the cells of the pre-2003 and 2003–2009 cohorts, regress log PM2.5 on standardized downwind exposure and its interaction with an indicator for the years from 2017 onward, with the fixed effects and weather controls of equation (4). The coefficient on exposure is the effect in the post-treatment months through 2011, before the interval over which the shift is measured, and the interaction is its change by the months from 2017 onward, after it. The months of 2012–2016 are excluded and pre-treatment months are always included.

In the first tercile, the per-exposure effect is 0.030 in the years through 2011 and 0.038 from 2017 onward, a change of 0.008 (s.e. 0.007). The change is the same in the third tercile and similar in the second (Panel B). The effect therefore did not decline in the districts whose in-zone mix moved most toward lighter industries. The terciles differ instead in the level of the effect, which is highest in the third, whose mix is the most fuel-intensive by 2016, and lowest in the second, whose

Table D1: Composition of Manufacturing Firms inside Zones, 2011–2016

Panel A: In-zone manufacturing firms, national					
	2011	2013	2014	2015	2016
Firms	5580	6671	7049	7712	8730
Employment (thousands)	1935	2413	2660	2914	3209
Pollution-intensive share of firms	0.200	0.202	0.190	0.191	0.190
Pollution-intensive share of employment	0.096	0.099	0.095	0.091	0.095
Share among firms with 100+ workers	0.164	0.176	0.168	0.163	0.168
Share among firms with <100 workers	0.231	0.223	0.210	0.214	0.206
Hot-process share of firms	0.098	0.101	0.092	0.089	0.091
Hot-process share of employment	0.042	0.042	0.039	0.035	0.034

Panel B: PM2.5 effect of exposure by change in the in-zone mix of the host district, 2011–2016			
	Districts sorted by the 2011–2016 change in predicted fuel use per worker of in-zone firms		
	(1) Decreased most	(2) Middle tercile	(3) Increased most
Mean change in log predicted fuel use per worker	-0.60	-0.07	0.57
Districts (zones)	32 (57)	31 (73)	31 (73)
Cells	586	325	515
Effect of exposure on $\ln(\text{PM}_{2.5})$, pre-2009 cohorts:			
Years 2011 and earlier	0.030 (0.013)	0.008 (0.006)	0.032 (0.007)
Years 2017 and later	0.038 (0.010)	0.020 (0.006)	0.040 (0.004)
Change, 2017 and later minus 2011 and earlier	0.008 (0.007)	0.012 (0.005)	0.008 (0.005)

Notes: Panel A reports manufacturing firms that report a zone location in the Vietnam Enterprise Survey, in the rounds in which the zone question was asked of nearly all firms (5580 firms in 2011 and 8730 in 2016). Pollution-intensive industries are the top quartile of the VES energy-based intensity ranking (Appendix B); hot-process industries are non-metallic minerals, basic metals, and coke and refined petroleum. Panel B sorts districts hosting zones operational by 2009, with at least five in-zone manufacturing firms in both 2011 and 2016 (94 districts hosting 203 zones), into terciles of the 2011–2016 change in the predicted fuel use per worker of their in-zone manufacturing firms, where predicted fuel use is employment multiplied by the fossil-fuel intensity per worker of the firm’s two-digit industry (Appendix B); the first row reports the mean change in log points. Each cell in the 5–25 km transport ring is assigned to the district hosting its nearest zone. For each tercile, the three coefficients come from one regression of $\ln(\text{PM}_{2.5})$ on standardized downwind exposure and its interaction with an indicator for the years from 2017 onward, with cell and year-month fixed effects and weather controls as in Panel B of Table 5, estimated on cells of the pre-2003 and 2003–2009 cohorts in the tercile, with pre-treatment cell-months always included and post-treatment cell-months restricted to the years through 2011 and from 2017 onward. The rows report the effect in the earlier period, the effect in the later period, and their difference. Province-clustered standard errors in parentheses.

zones host the largest concentrations of light assembly.

The shift in the first tercile came from light manufacturing growing beside the pollution-intensive plants rather than from those plants leaving: the number and employment of pollution-intensive in-zone firms rose over 2011–2016 while employment in other manufacturing more than doubled, and no district in the sample lost its pollution-intensive base.

Table D2: PM2.5 Effect of Exposure and Alternative Measures of the Shift in the In-Zone Mix

Measure of the shift ΔS_d	γ_1 (Exposure $\times \Delta S_d$)	γ_2 (Exposure $\times \mathbf{1}(t \geq 2017) \times \Delta S_d$)
Change in pollution-intensive share of employment	0.000 (0.005)	0.001 (0.005)
Log change in number of pollution-intensive firms	-0.003 (0.007)	-0.001 (0.004)
Log change in employment of pollution-intensive firms	-0.003 (0.007)	0.001 (0.004)
Log change in predicted fuel use	0.003 (0.007)	-0.008 (0.004)
Log change in predicted fuel use per worker	0.003 (0.005)	-0.003 (0.003)

Notes: Each pair of rows reports γ_1 and γ_2 from one estimate of $\ln PM_{imt} = \beta_1 E_{imt} + \beta_2 E_{imt} P_t + \gamma_1 E_{imt} \Delta S_d + \gamma_2 E_{imt} P_t \Delta S_d + \mathbf{X}'_{imt} \boldsymbol{\theta} + \alpha_i + \lambda_{mt} + \varepsilon_{imt}$, where E_{imt} is standardized downwind exposure, $P_t = \mathbf{1}(t \geq 2017)$, ΔS_d is the shift measure of the district hosting cell i 's nearest zone, standardized to unit variance across the 94 districts, and the controls and fixed effects are those of Panel B of Table D1. The effect of exposure is β_1 in the years through 2011 and $\beta_1 + \beta_2$ from 2017 onward; γ_1 measures how the early-period effect varies with the shift and γ_2 how the change in the effect between the two periods varies with it, so a positive γ_2 would indicate that the effect declined where the mix became lighter. The measures are the change in the share of in-zone manufacturing employment in pollution-intensive industries; the log change in the number of pollution-intensive in-zone firms; the log change in their employment; and the log changes in the predicted fuel use of all in-zone manufacturing firms, in total and per worker. The sample is that of Panel B of Table D1, pooling the three terciles and keeping the months through 2011 and from 2017 onward (1426 cells, 239568 cell-months). For the per-worker measure, the coefficient on exposure is 0.023 (0.006) and on its interaction with the later-period indicator 0.010 (0.004). Province-clustered standard errors in parentheses.

The conclusion does not depend on how the shift is measured (Table D2). Two caveats apply. The shifts observed over 2011–2016 were moderate and came from growth beside the heavy plants. The results therefore do not show what the exit of pollution-intensive firms would do, because it did not occur in these years. And the results are descriptive, since which firms enter and leave a zone is not exogenous.

E DOWNWIND EXPOSURE WEIGHTED BY OBSERVED COMBUSTION AT THE SOURCE

The exposure measure in equation (5) counts every operating zone within 25 km of a cell equally. This appendix examines how the PM2.5 effect of exposure relates to the thermal activity observed at the source, measured from the Type-2 thermal anomalies of Section 4.

Equation (5) is a sum of zone contributions,

$$E_{imt} = \sum_{j \in \mathcal{J}_i} c_{ijmt}, \quad c_{ijmt} = \mathbb{I}(t \geq T_j) \cdot \text{distance}_{ij}^{-1} \cdot D_{ijmt} \quad (\text{E1})$$

where zone j contributes once it is operating, with a weight that declines with its distance from cell i and rises with the share of the month the cell spends downwind of it. The weighted measure multiplies each contribution by the zone's observed combustion F_j ,

$$E_{imt}^F = \sum_{j \in \mathcal{J}_i} c_{ijmt} F_j \quad (\text{E2})$$

F_j is the zone's mean annual Type-2 detections over its operating years in 2002–2020, counted in the 0.05-degree cells within 5 or 10 km of the zone centroid, with each cell attributed to its nearest zone. Zones with no detections receive weight zero, and they are 40% of zones at 5 km and 29% at 10 km. Because F_j is fixed over time, month-to-month variation in E_{imt}^F comes from the downwind shares, as in E_{imt} , and setting $F_j \equiv 1$ returns E_{imt} itself. Alternative weights replace the detection count with fire radiative power or fire-days, or enter $\log(1 + F_j)$ in place of F_j .

The question is how a zone's effect varies with its observed combustion. Equation (4) assigns every zone the same effect per unit of contribution. Allowing each zone its own coefficient gives

$$\ln \text{PM2.5}_{imt} = \sum_{j \in \mathcal{J}_i} \beta_j c_{ijmt} + \mathbf{X}'_{imt} \boldsymbol{\theta} + \alpha_i + \mu_{mt} + \varepsilon_{imt}. \quad (\text{E3})$$

Estimating equation (4) with E_{imt}^F in place of E_{imt} imposes $\beta_j = \phi F_j$: a zone's effect is proportional to its observed combustion, and a zone without detections has none. The linear form $\beta_j = \beta + \phi F_j$ nests both cases, the proportional model at $\beta = 0$ and equation (4) at $\phi = 0$. Under the linear form, equation (E3) becomes

$$\ln \text{PM2.5}_{imt} = \beta E_{imt} + \phi E_{imt}^F + \mathbf{X}'_{imt} \boldsymbol{\theta} + \alpha_i + \mu_{mt} + \varepsilon_{imt} \quad (\text{E4})$$

where β is the effect of exposure to a zone with no detections and ϕ is the additional effect per unit of observed combustion.

Panel A of Table E1 reports two regressions for each weighting, on the sample of Table 4, column (2), with each measure standardized within it. The first uses the weighted measure as the only exposure regressor, which imposes $\beta = 0$, and gives a coefficient below the 0.027 for E_{imt} because it forces zones without detections to have no effect. The second is equation (E4), which

includes both measures: β equals the baseline estimate for every weighting and ϕ is indistinguishable from zero. Exposure to a zone with no detections raises PM2.5 by as much as exposure to any other zone, and additional observed combustion adds nothing.

Panel B sorts cells into terciles of their nearest zone's F_j and interacts E_{imt} with the tercile indicators. The effect does not rise across terciles, and zones with no detections raise downwind PM2.5 as much as zones in the top tercile, which record about four detections a year.

The per-exposure effect is therefore unrelated to the thermal activity observed at the source. This is the pattern the composition of zone tenants would predict. Thermal detections register the hot-process industries, which account for 9% of in-zone manufacturing firms and 3% of their employment (Appendix Table D1). The remaining firms burn fuel at temperatures below the detection threshold and contribute to ambient PM2.5 without contributing to detections. The thermal data record whether a zone hosts high-temperature sources, not how much it emits.

Table E1: Downwind Exposure Weighted by Observed Combustion at the Source

Panel A: Weighted exposure measures

Weight F_j on each zone's contribution	Weighted measure only	Both measures, equation (E4)	
	Coefficient (1)	β (2)	ϕ (3)
None (baseline exposure, Table 4, column 2)	0.027 (0.005)		
Mean annual detections within 5 km	0.014 (0.004)	0.026 (0.005)	0.002 (0.004)
Mean annual detections within 10 km	0.014 (0.005)	0.026 (0.005)	0.002 (0.004)
Mean annual radiative power within 5 km	0.013 (0.004)	0.026 (0.005)	0.002 (0.004)
Mean annual radiative power within 10 km	0.014 (0.004)	0.026 (0.005)	0.002 (0.004)
Mean annual fire-days within 5 km	0.015 (0.004)	0.026 (0.005)	0.002 (0.005)
log(1 + mean annual detections within 10 km)	0.022 (0.005)	0.026 (0.007)	0.001 (0.008)
log(1 + mean annual radiative power within 5 km)	0.023 (0.005)	0.026 (0.008)	0.001 (0.010)
Observations	793212		

Panel B: Effect of baseline exposure by tercile of the nearest zone's observed combustion

	Lowest (1)	Middle (2)	Highest (3)
Nearest zone's mean annual detections within 10 km	0.021 (0.005)	0.038 (0.009)	0.027 (0.007)
Tercile mean of the weight	0.0	0.3	3.9
Cells	1095	1112	1077
Nearest zone's mean annual radiative power within 5 km	0.025 (0.005)	0.045 (0.015)	0.026 (0.006)
Tercile mean of the weight	0.0	1.5	31.1
Cells	1648	557	1079

Notes: Panel A estimates equation (4) with the weighted measure of equation (E2) in place of E_{imt} (column (1)) and equation (E4) (columns (2) and (3)), on the sample of Table 4, column (2): cells in the 5–25 km transport ring, monthly, with cell and year-month fixed effects and weather controls. Each measure is standardized by its standard deviation in the sample, so coefficients are per standard deviation of the respective measure. Weights are defined in the text; zones with no detections receive weight zero. Panel B interacts E_{imt} with indicators for terciles of the nearest zone's weight, defined over the cells in the estimation sample, and reports the tercile mean of the weight and the number of cells; cells whose nearest zone is not in the fire file are excluded (748752 cell-months). Province-clustered standard errors in parentheses.

Appendix References

- Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.
- Callaway, B., & Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230.
- de Chaisemartin, C., & d'Haultfoeuille, X. (2024). Difference-in-differences estimators of intertemporal treatment effects. *Review of Economics and Statistics*, 1–45.
- Garg, A., Kainou, K., & Pulles, T. (2006). Introduction. In S. Eggleston, L. Buendia, K. Miwa, T. Ngara, & K. Tanabe (Eds.), *2006 IPCC guidelines for national greenhouse gas inventories, volume 2: Energy*. Institute for Global Environmental Strategies.
- Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J.-N. (2021). ERA5-Land: a state-of-the-art global reanalysis dataset for land applications. *Earth System Science Data*, 13(9), 4349–4383. <https://doi.org/10.5194/essd-13-4349-2021>
- Rambachan, A., & Roth, J. (2023). A more credible approach to parallel trends. *Review of Economic Studies*, 90(5), 2555–2591.
- Shapiro, J. S., & Walker, R. (2018). Why is pollution from US manufacturing declining? the roles of environmental regulation, productivity, and trade. *American Economic Review*, 108(12), 3814–3854.
- Shen, S., Li, C., Van Donkelaar, A., Jacobs, N., Wang, C., & Martin, R. V. (2024). Enhancing global estimation of fine particulate matter concentrations by including geophysical a priori information in deep learning. *Acs Es&T Air*, 1(5), 332–345.
- United Nations Department of Economic and Social Affairs. (2022). *2022 united nations energy statistics yearbook* [Table II: Selected Conversion Factors for Crude Oil and Oil Products]. United Nations.