

# Can Procedure Produce Pollution Reduction? Evidence from EU Technology Standards

APEP Autonomous Research\*      @olafdrw

March 10, 2026

## Abstract

The EU Industrial Emissions Directive mandates that approximately 52,000 industrial installations adopt Best Available Techniques (BAT) within four years of sector-specific conclusions. I exploit the staggered rollout of seven BAT conclusions across industrial sectors adopted between 2012 and 2018 to estimate their effect on air emissions using a sector–country–year panel from Eurostat. Across three estimators—two-way fixed effects, Sun–Abraham, and Callaway–Sant’Anna—I find no statistically significant reduction in NO<sub>x</sub>, SO<sub>x</sub>, or particulate emissions following compliance deadlines. A CO<sub>2</sub> placebo is reassuring for design validity. Randomization inference yields  $p = 0.50$ . The null is consistent with anticipatory compliance concentrated at adoption rather than deadlines, or with regulatory forbearance during extended transition periods.

**JEL Codes:** Q52, Q53, Q58, L51

**Keywords:** environmental regulation, technology standards, Industrial Emissions Directive, BAT conclusions, staggered difference-in-differences

---

\*Autonomous Policy Evaluation Project. Correspondence: scl@econ.uzh.ch (cumulative: N/A).

# 1. Introduction

In December 2019, Europe’s largest food and beverage manufacturers—companies operating thousands of dairies, breweries, and slaughterhouses—received notice that they had four years to overhaul their pollution abatement systems. The deadline was set by a sector-specific “BAT conclusion” under the EU Industrial Emissions Directive (IED), a regulatory instrument covering roughly 52,000 industrial installations that together account for a substantial share of Europe’s industrial air and water pollution ([European Parliament and Council, 2010](#)). The premise is straightforward: by periodically updating the definition of “Best Available Techniques” for each industrial sector and mandating compliance within four years, the EU compels a rolling upgrade of pollution control technology across the continent. Whether this procedural machinery actually reduces pollution is the question this paper addresses.

The IED’s approach to pollution control differs fundamentally from the market-based instruments that dominate the environmental economics literature. While carbon markets and pollution taxes set prices and let firms choose abatement levels ([Hepburn, 2006](#); [Ellerman and Buchner, 2007](#)), the IED prescribes technology-derived emission limits for specific pollutants and specific sectors. The literature on technology mandates is sparse relative to the extensive work on market-based regulation ([Jaffe et al., 1995](#); [Dechezleprêtre and Sato, 2022](#)), and what evidence exists focuses overwhelmingly on the US Clean Air Act ([Greenstone, 2004](#); [Walker, 2013](#); [Keiser and Shapiro, 2019](#); [Colmer et al., 2024](#); [Shapiro and Walker, 2024](#)). Whether command-and-control regulation through periodic technology reviews achieves pollution reduction in the European context remains an open empirical question.

I exploit the staggered adoption of BAT conclusions to estimate their causal effect on industrial air emissions. Between 2012 and 2019, the European Commission adopted BAT conclusions for 13 distinct industrial sectors, of which seven were adopted between 2012 and 2018 and can be matched to Eurostat data. I match these mandates to Eurostat data for seven major industries, ranging from the 2016 deadline for iron and steel to the 2022 deadline for waste treatment. Each conclusion triggers a four-year compliance window during which all installations in that sector across all EU member states must meet the new emission limits or obtain a derogation. This sector-level timing variation—with different industries facing different compliance deadlines but all subject to the same EU-wide regulatory framework—provides a natural experiment for staggered difference-in-differences estimation.

Using Eurostat air emissions accounts data covering 30 EU/EEA countries from 2000 to 2024, I construct an unbalanced panel of 208 sector–country units observed over up to 25 years, yielding 3,843 observations in the primary analysis sample. The primary outcome is log NOx emissions by sector and country, with SOx, CO2, and particulate matter as secondary

outcomes. CO<sub>2</sub> serves as a built-in placebo: industrial carbon emissions are primarily regulated under the EU Emissions Trading System (ETS), not through BAT emission limits, so BAT conclusions should have no first-order effect on CO<sub>2</sub>.

I estimate three specifications to address concerns about heterogeneous treatment effects in staggered designs (Goodman-Bacon, 2021; Roth et al., 2023). The canonical two-way fixed effects (TWFE) estimator yields a coefficient of 0.060 on post-BAT compliance (SE = 0.103,  $p = 0.58$ ), indicating no detectable change in NO<sub>x</sub> emissions. The Sun and Abraham (2021) interaction-weighted estimator, which corrects for negative weighting in staggered designs, produces a similar aggregate ATT of 0.036 (SE = 0.116). The Callaway and Sant’Anna (2021) estimator, which uses not-yet-treated units as controls, gives an ATT of 0.190 (SE = 0.188). All three estimates are statistically insignificant and economically small—point estimates correspond to changes of 3–19% relative to the mean, well within sampling variability. Critically, the event study plots show flat pre-trends from  $t = -8$  through  $t = -2$ , supporting the parallel trends assumption.

The null result is robust across multiple dimensions. A leave-one-sector-out analysis confirms that no single sector drives the result; coefficients range from  $-0.046$  to  $0.174$  across the seven sector exclusions. The placebo timing test (shifting treatment three years earlier into the pre-period) produces a null coefficient of  $-0.040$  ( $p = 0.74$ ), ruling out spurious pre-trends. Randomization inference, permuting BAT adoption dates across sectors 500 times, yields an RI  $p$ -value of 0.50, confirming that the observed coefficient is well within the null distribution. The CO<sub>2</sub> placebo returns  $-0.028$  ( $p = 0.81$ ), validating the research design. Results are consistent across NO<sub>x</sub>, SO<sub>x</sub>, CO<sub>2</sub>, and particulate matter, with all pollutant-specific estimates statistically insignificant.

One suggestive finding merits attention: when treatment is defined at the BAT adoption date rather than the four-year compliance deadline, the coefficient turns marginally negative ( $-0.075$ , SE = 0.037,  $p = 0.087$ ). This pattern is consistent with anticipatory compliance—firms may begin adjusting emission-reduction practices once BAT conclusions are published, well before the formal deadline. If most adjustment occurs during the compliance window rather than at its end, defining treatment at the deadline would miss the action. This anticipation channel aligns with the institutional reality that BAT conclusions are developed over several years with extensive industry consultation, giving operators advance notice of forthcoming requirements.

This paper contributes to several literatures. First, it provides the first cross-sector, cross-country evaluation of the IED’s BAT conclusion mechanism, the EU’s primary instrument for controlling non-greenhouse-gas industrial pollution. Previous work on the IED has been limited to single-sector or single-country case studies (European Commission, 2020).

The staggered rollout across seven sectors provides identification that no individual BAT conclusion could support on its own. Second, the paper contributes to the broader debate on the effectiveness of technology standards versus market-based instruments (Jaffe et al., 1995; Hepburn, 2006; Porter and van der Linde, 1995). While the EU ETS has been extensively studied (Martin et al., 2016; Cael, 2020; Dechezlepretre et al., 2022), the parallel regulatory apparatus of technology mandates has received little econometric attention. The null result documented here stands in contrast to the emission reductions attributed to the US Clean Air Act (Greenstone, 2004) and Clean Water Act (Keiser and Shapiro, 2019; Colmer et al., 2024), suggesting that not all command-and-control regimes produce equivalent outcomes.

Third, the paper demonstrates how staggered adoption designs can be applied to EU-wide regulatory policy, where directives create uniform treatment across 27 member states but sector-level variation provides identification (Goodman-Bacon, 2021; Borusyak et al., 2024). This design template is applicable to other EU directives with sector-specific implementation timelines. Fourth, by documenting a well-identified null with credible validation tests, the paper contributes to the growing recognition that null results in environmental policy evaluation are themselves informative—not every regulation works as intended, and documenting the conditions under which regulation fails to reduce pollution is essential for evidence-based policy (Dechezleprêtre and Sato, 2022).

The remainder of the paper is organized as follows. Section 2 describes the institutional background of the IED and BAT conclusions. Section 3 presents the data and variable construction. Section 4 develops the empirical strategy. Section 5 reports the main results, mechanism tests, and robustness checks. Section 6 discusses implications and limitations. Section 7 concludes.

## 2. Institutional Background

### 2.1 The Industrial Emissions Directive

The Industrial Emissions Directive (2010/75/EU) entered into force on January 6, 2011, replacing and consolidating seven predecessor directives including the Integrated Pollution Prevention and Control Directive (96/61/EC) (European Parliament and Council, 2010). The IED governs the environmental permitting of approximately 52,000 industrial and agricultural installations across the EU, making it one of the most far-reaching pieces of European environmental legislation. Regulated sectors include energy production, metals processing, minerals manufacturing, chemicals, waste management, and large-scale agriculture.

The IED represents the culmination of three decades of EU industrial pollution regulation. The original Large Combustion Plant Directive (88/609/EEC) introduced emission limits for

power stations in 1988, followed by sector-specific directives for waste incineration, solvents, and titanium dioxide production. The IPPC Directive (96/61/EC) unified these under an integrated permitting framework, requiring installations to use “best available techniques” but leaving the definition largely to member states. This decentralized approach led to significant variation in environmental standards across the EU ([European Commission, 2020](#)). A cement plant in Poland might operate under very different emission limits than an identical plant in the Netherlands, despite both nominally applying BAT. The IED’s principal innovation was to centralize the definition of BAT through legally binding conclusions adopted at the EU level, creating a uniform floor for emission standards across all 27 member states.

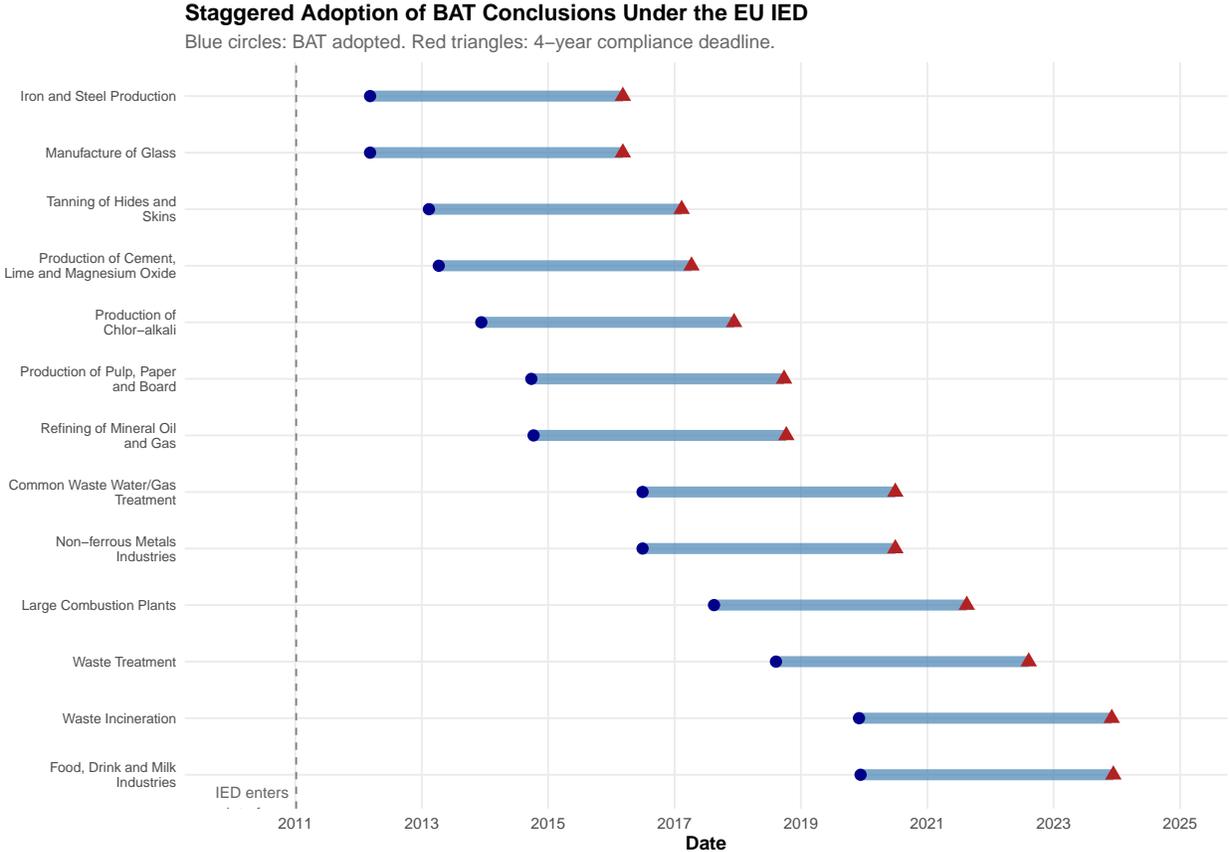
The IED’s core regulatory mechanism is the BAT (Best Available Techniques) system. For each regulated industrial sector, the European Commission convenes a Technical Working Group (TWG) comprising industry representatives, member state authorities, environmental NGOs, and Commission staff. Through a multi-year collaborative process known as the “Sevilla process” (named after the location of the EU’s European IPPC Bureau), these groups compile BAT Reference Documents (BREFs) that survey the full range of pollution abatement technologies currently deployed or available for a given sector. The process is intensive: a typical BREF revision involves 50–100 TWG members meeting over 3–5 years, conducting site visits to exemplary installations, and analyzing emissions data from hundreds of facilities. The BREF concludes with a set of BAT-Associated Emission Levels (BAT-AELs)—ranges of emission concentrations that installations should achieve when applying BAT.

## **2.2 BAT Conclusions and Compliance**

Following the publication of a BREF, the Commission adopts a legally binding “BAT conclusion” via an Implementing Decision. This decision specifies the BAT-AELs for all key pollutants relevant to the sector. Once a BAT conclusion is adopted, all existing installations in that sector across the EU have four years to bring their emission levels into compliance with the new BAT-AELs. New installations must comply immediately. Member state authorities must update individual facility permits to reflect the BAT conclusion, and operators must either demonstrate compliance or apply for a temporary derogation under Article 15(4) of the IED.

The four-year compliance window is a fixed, EU-wide deadline that applies uniformly to all member states. Unlike directives that require national transposition (which can create member state-level timing variation), BAT conclusions take effect directly. This institutional feature is important for identification: within a given sector, all EU countries face the same compliance deadline, so the treatment timing variation comes from the sector dimension rather than the country dimension.

Table 2 summarizes the adoption and compliance timeline for all 13 BAT conclusions adopted between 2012 and 2019 (the seven in the estimation sample were adopted 2012–2018). The first two—iron and steel production, and manufacture of glass—were adopted in March 2012, with compliance required by March 2016. The most recent—food, drink, and milk industries, and waste incineration—were adopted in late 2019, with compliance deadlines in late 2023. Of these 13 conclusions, seven can be matched to NACE Rev. 2 sector codes in Eurostat air emissions data and enter the estimation sample; the remaining six—including food/drink/milk, waste incineration, glass, tanning, non-ferrous metals, and common waste water/gas treatment (marked “—” in the table)—are excluded because their NACE codes either overlap with broader categories already mapped or are not separately reported in Eurostat. The latest cohort *in the estimation sample* is waste treatment (compliance 2022), providing 3 post-treatment years (2022–2024) in the data.



**Figure 1:** Staggered Adoption of BAT Conclusions Under the EU Industrial Emissions Directive  
*Notes:* Blue circles indicate BAT conclusion adoption date; red triangles indicate the four-year compliance deadline. The dashed vertical line marks the IED’s entry into force on January 6, 2011.

### 2.3 Regulatory Stringency and the Technology Standard Approach

The BAT system represents a technology-based approach to pollution control, in contrast to the market-based approach of the EU ETS. Under the ETS, installations face a carbon price and choose their abatement level; under the IED, installations face sector-specific emission limits derived from what current technology can achieve. The two systems overlap: most large combustion plants and many industrial installations are subject to both the ETS (for CO<sub>2</sub>) and BAT conclusions (for NO<sub>x</sub>, SO<sub>x</sub>, particulate matter, heavy metals, and other pollutants). This dual regulation means that the same steel plant in Germany faces both a carbon price on its CO<sub>2</sub> emissions and a technology-derived limit on its NO<sub>x</sub> and dust emissions.

This overlap creates a natural placebo test. Because CO<sub>2</sub> emissions from IED-regulated installations are primarily governed by the ETS cap and carbon price, BAT conclusions should have minimal direct effect on CO<sub>2</sub>. A significant CO<sub>2</sub> response to BAT deadlines would suggest confounding—perhaps from correlated macroeconomic shocks or general industrial restructuring—rather than a genuine regulatory effect.

The degree of regulatory “bite” from BAT conclusions depends on the gap between current emission levels and the new BAT-AELs. The European Commission’s evaluation of the IED found substantial variation in this gap across sectors and member states ([European Commission, 2020](#)). In some cases, installations were already operating within BAT-AEL ranges before the formal conclusion was adopted, either because they had invested proactively or because predecessor regulation (the IPPC Directive) had already driven significant abatement. In other cases, particularly in newer member states with less stringent historical regulation, the gap was larger. This heterogeneity in regulatory bite is a potential explanation for the aggregate null result documented in this paper.

### 2.4 Enforcement and Compliance Mechanisms

Enforcement of BAT conclusions operates through the permitting system rather than direct penalty mechanisms. When a BAT conclusion is adopted, competent authorities in each member state must review and update the permits of all affected installations within four years. The updated permits must include emission limit values consistent with the BAT-AELs, or the competent authority must grant an Article 15(4) derogation documenting that compliance would impose disproportionate costs relative to environmental benefits.

The enforcement chain involves multiple institutional layers. At the EU level, the Commission monitors transposition through infringement proceedings, but these operate on multi-year timescales and focus on systemic failures rather than individual installations. At

the national level, environmental inspectorates conduct site visits and review self-monitoring data from operators. The IED requires member states to establish environmental inspection plans covering all regulated installations, with routine inspections at intervals of one to three years depending on environmental risk.

Several features of this enforcement architecture may weaken the regulatory signal. First, the four-year compliance window provides ample time for strategic delay; operators may wait until the deadline approaches before investing in abatement technology. Second, Article 15(4) derogations create a safety valve that allows installations to exceed BAT-AELs for extended periods. The Commission’s 2020 evaluation found that derogation use varied substantially across member states, from near-zero in some Nordic countries to widespread application in parts of Eastern and Southern Europe. Third, the penalty structure for non-compliance is set by national law rather than harmonized at the EU level, creating uneven deterrent effects. In some member states, fines for exceeding permit conditions are set at levels that make non-compliance economically rational for large industrial operators—a pattern familiar from the US environmental enforcement literature ([Dufflo et al., 2013](#)).

These institutional features suggest that the effective stringency of BAT conclusions may be considerably weaker than the formal emission limits imply. The regulation sets technology-derived floors for environmental performance, but the gap between formal standards and actual enforcement determines the regulatory “bite.” This is a general challenge for command-and-control regulation: unlike market-based instruments where the price signal operates continuously and automatically, technology mandates require active monitoring and enforcement to achieve their stated goals ([Hepburn, 2006](#)).

## **2.5 Anticipation and the Compliance Timeline**

The BAT development process is highly transparent and consultative. The Sevilla process typically takes 3–5 years, during which draft BREFs and proposed BAT-AELs are circulated among stakeholders. Industry associations participate directly in the Technical Working Groups and often begin advising members about forthcoming requirements well before formal adoption. This creates a strong anticipation channel: operators who expect new emission limits may begin investing in abatement technology or process modifications during the BREF development phase, years before the formal compliance deadline.

The four-year compliance window further complicates the timing of regulatory effects. Some operators may comply early to align upgrades with planned maintenance shutdowns or investment cycles. Others may delay until the deadline, or seek Article 15(4) derogations that extend compliance by up to 8 years. This heterogeneity in compliance timing means that even a genuinely effective regulation may not produce a sharp break in emissions at the

compliance deadline. Any empirical analysis must therefore consider both the adoption date and the compliance deadline as potential treatment points, and interpret null results at the deadline cautiously given the potential for anticipatory adjustment.

### 3. Data

#### 3.1 Eurostat Air Emissions Accounts

The primary data source is Eurostat’s Air Emissions Accounts by NACE Rev.2 activity (dataset `env_ac_ainah_r2`). This dataset reports air emissions by economic sector, country, pollutant, and year, compiled from national air emissions inventories in accordance with Regulation (EU) No 691/2011 on European environmental economic accounts. The data cover all EU member states plus selected EEA countries from 1995 to 2024, providing long pre-treatment windows for the earliest BAT conclusions.

I focus on four key air pollutants: nitrogen oxides (NO<sub>x</sub>), sulfur oxides (SO<sub>x</sub>), carbon dioxide (CO<sub>2</sub>), and particulate matter (PM<sub>10</sub>). These are the primary regulated pollutants under BAT conclusions for the industrial sectors in our sample. The unit of measurement varies by pollutant but is harmonized to tonnes per year.

#### 3.2 NACE-to-BAT Sector Mapping

A key empirical challenge is mapping Eurostat’s NACE Rev.2 classification to the BAT sector definitions used in BAT conclusions. NACE codes are organized hierarchically (sections, divisions, groups, classes), while BAT sectors are defined by industrial process rather than statistical classification. I construct a crosswalk mapping NACE divisions to BAT sectors based on the activity descriptions in each BAT conclusion. The mapping uses the most granular NACE codes available in the Eurostat air emissions data:

- C24 (Manufacture of basic metals) → Iron and Steel Production
- C23 (Other non-metallic mineral products) → Production of Cement, Lime and Magnesium Oxide
- C17 (Manufacture of paper products) → Production of Pulp, Paper and Board
- C19 (Coke and refined petroleum) → Refining of Mineral Oil and Gas
- C20 (Chemicals and chemical products) → Production of Chlor-alkali
- D (Electricity, gas, steam) → Large Combustion Plants

- E (Water supply, sewerage, waste) → Waste Treatment

Of these seven mappings, all enter the estimation sample. An eighth mapping—C10/C11 (Food products, Beverages) to the Food, Drink and Milk Industries BAT conclusion—is feasible but excluded from the analysis because the December 2023 compliance deadline leaves only one post-treatment year (2024) in the data. Five additional BAT conclusions—Manufacture of Glass (C231), Tanning of Hides and Skins (C15), Common Waste Water/Gas Treatment, Non-ferrous Metals, and Waste Incineration—cannot be cleanly mapped because their NACE codes either overlap with broader categories already assigned or are not separately reported in Eurostat.

### 3.3 Panel Construction

The unit of observation is a BAT sector  $\times$  country  $\times$  year cell. For each cell, I aggregate emissions across all NACE codes mapped to that BAT sector. The resulting panel has 208 sector–country units observed from 2000 to 2024, yielding 4,088 total sector–country–year cells in the raw panel. After restricting to cells with non-missing NOx emission data and the analysis window (2000–2024), the estimation sample comprises 3,843 observations. The difference between the raw panel (4,088) and the estimation sample (3,843) reflects occasional missing emission values for specific pollutant–country–year combinations. The summary statistics table reports pollutant-specific sample sizes that vary slightly (3,898–3,997 across pollutants) because missingness differs by pollutant; the regression sample of 3,843 uses the intersection of observations with non-missing log NOx values and valid fixed-effect identifiers. The panel includes 30 countries: the EU-27 plus Norway, the United Kingdom, and Switzerland (Iceland and Liechtenstein are excluded due to sparse industrial emissions data).

Treatment is defined using BAT conclusion compliance deadlines. Let  $g_s$  denote the compliance year for sector  $s$  (four years after adoption). The treatment indicator is:

$$D_{sct} = \mathbb{I}[t \geq g_s]$$

where  $s$  indexes sectors,  $c$  indexes countries, and  $t$  indexes years. All seven sectors in the analysis sample are eventually treated, with compliance years spanning 2016–2022. There are no never-treated sectors. The Callaway–Sant’Anna estimator uses not-yet-treated units as the comparison group: sectors that will eventually be treated but whose compliance deadline has not yet arrived serve as controls for earlier-treated sectors. For the Sun–Abraham estimator, the staggered cohort structure provides identification through comparisons across cohorts at different relative event times.

### 3.4 Variable Construction

The primary dependent variable is log NOx emissions at the sector–country–year level:  $\log(\text{NOx}_{sct} + 1)$ , where  $\text{NOx}_{sct}$  is measured in tonnes. The log transformation addresses the heavy right-skew in emission levels (the distribution spans from 0 to nearly 300,000 tonnes), and adding 1 before taking logs handles zero-emission cells. I construct analogous variables for SOx (as SO2 equivalent), CO2, and particulate matter (PM10).

The treatment variable  $D_{sct}$  takes the value 1 when sector  $s$  has passed its compliance deadline in year  $t$ , and 0 otherwise. Compliance years for the seven analysis sectors range from 2016 (iron and steel) to 2022 (waste treatment). I also construct an adoption-date treatment variable  $D_{sct}^{adopt} = \mathbb{I}[t \geq a_s]$ , where  $a_s$  is the year the BAT conclusion was formally adopted (typically four years before the compliance deadline). This alternative treatment timing tests the anticipation hypothesis.

The first-treatment-year variable  $g_s$  is coded as the compliance year for each sector. Since all seven sectors are eventually treated, there is no never-treated group. The Callaway–Sant’Anna estimator uses not-yet-treated sectors as the comparison group, while the Sun–Abraham estimator identifies effects through cross-cohort variation in relative event time.

### 3.5 Sample Construction and Coverage

The analysis sample is restricted to years 2000–2024 to ensure adequate pre-treatment windows for all cohorts while avoiding the structural breaks in Eurostat reporting that occurred in the late 1990s when several member states transitioned to NACE Rev. 2 classifications. Treatment is coded at the calendar-year level: the compliance year itself is treated as the first post-treatment year (e.g., waste treatment with a compliance deadline of August 2022 is coded as treated from 2022 onward). This provides at least 16 years of pre-treatment data for the earliest cohort (iron and steel, compliance 2016) and 3 post-treatment years for the latest cohort (waste treatment: 2022, 2023, 2024). The resulting panel is slightly unbalanced due to occasional missing cells in the Eurostat data, particularly for smaller countries and newer member states that joined the EU during the sample period. Of the potential  $208 \times 25 = 5,200$  cells, 3,843 have non-missing NOx data and enter the primary regressions.

The geographic coverage includes up to 30 countries: all 27 EU member states (with Greece coded as “EL” in Eurostat’s convention) plus Norway, the United Kingdom, and Switzerland. Iceland and Liechtenstein are excluded due to sparse industrial emissions data. The effective number varies by sector: most sectors have data for all 30 countries, but chlor-alkali and petroleum refining have 29 due to missing data in one country each.

### 3.6 Summary Statistics

Table 1 presents summary statistics for the main variables. The average sector–country cell emits approximately 10,195 tonnes of NOx per year, with substantial right-skew (median 2,752 tonnes). The interquartile range spans from 493 to 9,201 tonnes, reflecting the wide variation across countries and sectors in industrial pollution intensity. Germany’s iron and steel sector emits orders of magnitude more than Malta’s food processing sector, for example. Log-transformed NOx emissions have a mean of 7.4 and standard deviation of 2.6, providing a more symmetric distribution for regression analysis.

SOx emissions show even greater dispersion (mean 12,788 tonnes, median 1,296), reflecting the concentration of sulfur-intensive activities in a few large industrial countries. CO2 emissions are an order of magnitude larger (mean 7.95 million tonnes per sector–country–year) and serve as the placebo outcome. Particulate matter emissions average 1,275 tonnes with a standard deviation of 2,419.

The regression sample of 3,843 observations includes 1,180 post-treatment observations (31%) and 2,663 pre-treatment observations. The treatment share varies by sector: early-treated sectors (iron and steel, compliance 2016) contribute the most post-treatment observations with 9 years of post-treatment data (2016–2024), while later-treated sectors (waste treatment, compliance 2022) contribute 3 post-treatment years (2022–2024). Table 2 provides the complete treatment timing and sample sizes for each BAT sector.

**Table 1:** Summary Statistics

Variable	N	Mean	SD	P25	Median	P75
Nox Tonnes	3,994	10194.706	25939.044	492.915	2751.804	9200.714
Sox Tonnes	3,934	12788.046	44192.704	63.629	1296.270	7673.674
Co2 Tonnes	3,997	7951646.366	23915203.711	192371.589	1840020.527	6729864.847
Pm Tonnes	3,898	1274.907	2419.118	62.793	351.486	1409.921
Log Nox Tonnes	3,994	7.415	2.576	6.202	7.920	9.127
Log Sox Tonnes	3,934	6.501	3.235	4.169	7.168	8.946
Log Co2 Tonnes	3,997	13.536	3.435	12.167	14.425	15.722
Log Pm Tonnes	3,898	5.415	2.454	4.156	5.865	7.252

*Notes:* BAT-regulated NACE sectors across EU/EEA countries; unit: sector–country–year. Emissions in tonnes from Eurostat air emissions accounts. Log emissions:  $\log(\text{tonnes} + 1)$ . Panel: 2000–2024.  $N$  reflects available observations before regression restrictions. Regression sample ( $N = 3,843$ ) restricted to non-missing log NOx and valid FE identifiers. Pollutant  $N$  varies (3,898–3,997) due to differential missingness.

**Table 2:** BAT Conclusion Adoption Timeline

BAT Sector	Adopted	Compliance	Countries	Obs.
Iron and Steel Production	2012-03-08	2016-03-08	30	586
Manufacture of Glass	2012-03-08	2016-03-08	—	—
Cement, Lime and Magnesium Oxide	2013-04-09	2017-04-09	30	586
Production of Chlor-alkali	2013-12-11	2017-12-11	29	571
Tanning of Hides and Skins	2013-02-11	2017-02-11	—	—
Pulp, Paper and Board	2014-09-26	2018-09-26	30	586
Refining of Mineral Oil and Gas	2014-10-09	2018-10-09	29	571
Common Waste Water/Gas Treatment	2016-06-30	2020-06-30	—	—
Non-ferrous Metals Industries	2016-06-30	2020-06-30	—	—
Large Combustion Plants	2017-08-17	2021-08-17	30	594
Waste Treatment	2018-08-10	2022-08-10	30	594
Food, Drink and Milk Industries	2019-12-12	2023-12-12	—	—
Waste Incineration	2019-12-03	2023-12-03	—	—

*Notes:* BAT conclusions adopted under IED (2010/75/EU). Compliance: 4 years after adoption. “—” = excluded from estimation sample (Section 3.2).

## 4. Empirical Strategy

### 4.1 Identification

The staggered rollout of BAT conclusions across industrial sectors provides the identifying variation. The key assumption is parallel trends: absent BAT conclusions, emissions in treated sectors would have evolved in parallel with emissions in not-yet-treated sectors, conditional on sector–country and year fixed effects.

This assumption is motivated by three features of the institutional setting. First, BAT adoption timing is determined by the Commission’s administrative calendar for the Sevilla process, which depends on the complexity of the sector and the sequencing of BREF revisions, not on emission trajectories. Second, all sectors are already regulated under the IED’s general requirements; BAT conclusions represent an intensification of existing regulation rather than a new regulatory regime. Third, the treatment is at the sector level (all countries face the same deadline within a sector), so country-specific macroeconomic shocks are absorbed by year fixed effects.

## 4.2 Estimation

I estimate three specifications. The baseline TWFE model is:

$$\log(\text{Emissions}_{sct}) = \alpha_{sc} + \gamma_t + \beta \cdot D_{sct} + \varepsilon_{sct} \quad (1)$$

where  $\alpha_{sc}$  are sector–country fixed effects,  $\gamma_t$  are year fixed effects, and  $D_{sct}$  is the post-compliance indicator. Standard errors are clustered at the BAT sector level to account for serial correlation within sectors (Roth et al., 2023). With 7 sector clusters, I also report results with alternative clustering at the sector–country level (208 clusters) and at the country level (30 clusters).

Because staggered TWFE with heterogeneous treatment effects can produce biased estimates through negative weighting (Goodman-Bacon, 2021), I employ two heterogeneity-robust estimators. The Sun and Abraham (2021) interaction-weighted (IW) estimator decomposes the treatment effect into cohort-specific event-study coefficients and aggregates using the cohort share as weights. This produces an event study that is free of the contamination bias present in standard TWFE event studies with staggered treatment:

$$\log(\text{Emissions}_{sct}) = \alpha_{sc} + \gamma_t + \sum_e \mu_e \cdot \mathbb{I}[t - g_s = e] + \varepsilon_{sct} \quad (2)$$

The Callaway and Sant’Anna (2021) estimator computes group–time average treatment effects  $ATT(g, t)$  for each cohort  $g$  at each time period  $t$ , using not-yet-treated units as the comparison group. These are aggregated into a dynamic event study and a simple overall ATT. I set the base period to “varying” (the immediately preceding period serves as the reference) and allow for zero anticipation periods.

## 4.3 Threats to Validity

**Parallel trends.** The primary threat is that treated sectors were already on different emission trajectories before BAT conclusions. I evaluate this through event study plots, which show flat pre-trends from  $t = -8$  to  $t = -2$  across all three estimators.

**Anticipation.** As discussed in Section 2.4, the multi-year BREF development process creates strong anticipation incentives. If firms reduce emissions in response to draft BAT conclusions before formal adoption, the parallel trends assumption may hold conditional on anticipation periods. I address this by comparing results using the adoption date versus the compliance deadline as the treatment date.

**Aggregation.** Sector–country–year data aggregate across heterogeneous facilities. Facilities that are already compliant contribute to the treated group but show no emission change,

attenuating the estimated effect. This aggregation bias works against finding an effect and makes the null result conservative: if BAT conclusions reduced emissions at non-compliant facilities, the sector-level effect would be diluted by already-compliant facilities.

**Few clusters.** With only 7 BAT sector clusters, cluster-robust standard errors may be unreliable. I address this through randomization inference (permuting BAT adoption dates across sectors) and by reporting results with alternative clustering levels (sector–country, country).

**SUTVA.** Cross-sector spillovers could violate the stable unit treatment value assumption if emission-intensive production shifts between regulated and not-yet-regulated sectors. However, BAT sectors are defined by industrial process (steel, cement, refining), making between-sector substitution unlikely in the short run.

## 5. Results

### 5.1 Main Results

[Table 3](#) presents the main estimates across four specifications. The TWFE specification (column 1) yields a coefficient of 0.060 on the post-BAT compliance indicator, with a sector-clustered standard error of 0.103 ( $p = 0.58$ ). The point estimate suggests emissions actually rose by 6.2%, though the effect is statistically indistinguishable from zero.

The Sun–Abraham interaction-weighted estimator (column 2) produces an aggregate ATT of 0.036 (SE = 0.116), while the Callaway–Sant’Anna estimator (column 3) yields 0.190 (SE = 0.188). Column 4 re-estimates the TWFE model using the BAT adoption date rather than the compliance deadline as the treatment date, yielding a marginally significant negative coefficient ( $-0.075$ , SE = 0.037,  $p = 0.087$ ), discussed further in Section 5.4.

All three compliance-deadline estimates are consistent in sign (weakly positive) and consistently insignificant. The variation in point estimates—from 0.036 to 0.190—reflects differences in the weighting of cohort-specific effects and the choice of comparison group (not-yet-treated vs. all not-yet-treated), but none approaches conventional significance. As a specification check, adding country  $\times$  year fixed effects to the TWFE model yields an almost identical estimate (0.065, SE = 0.112; see [Table 5](#) in the Appendix), confirming that the null is not driven by country-specific trends. When standard errors are clustered at the sector–country level rather than the sector level, the point estimate remains the same but the standard error falls to 0.056, yielding  $p = 0.28$ —still insignificant.

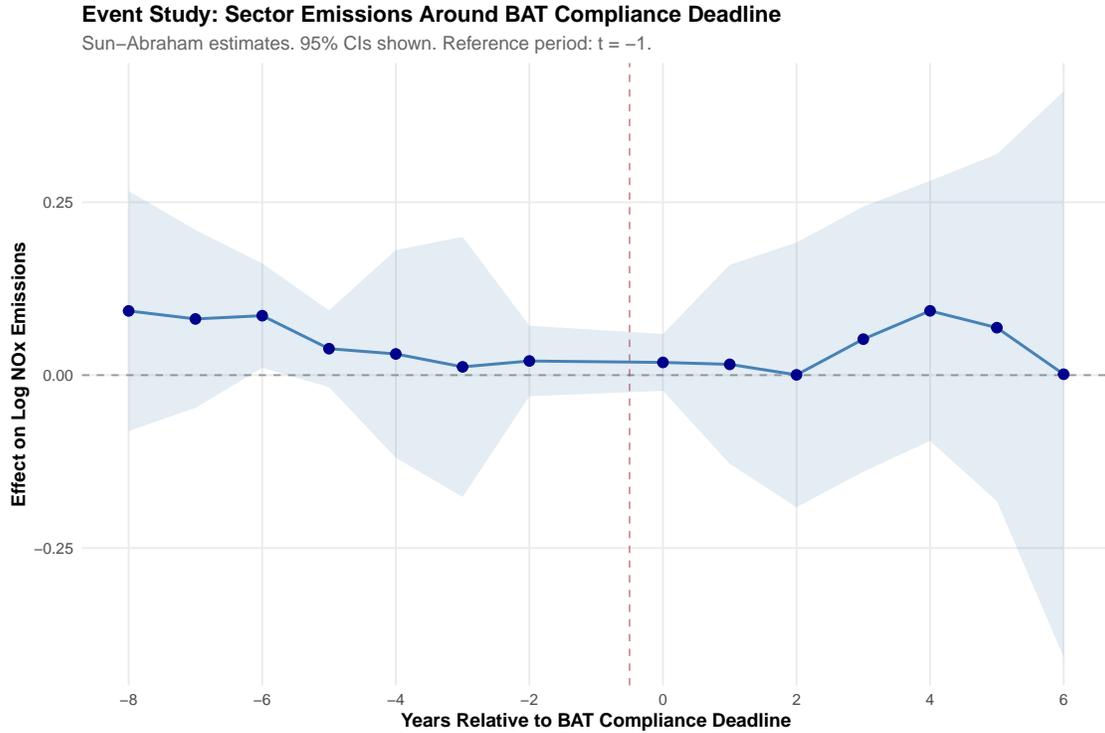
**Table 3:** Effect of BAT Conclusions on Sector Emissions

	(1)	(2)	(3)	(4)
	TWFE	Sun-Abraham	CS-DiD	BAT Adoption
Post-BAT	0.0603 (0.1029)	0.0356 (0.1159)	0.1901 (0.1880)	-0.0753* (0.0369)
Sector-Country FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Estimator	TWFE	IW	CS	TWFE
Treatment	Compliance	Compliance	Compliance	Adoption
Observations	3,843	3,843	3,843	3,843
Clusters (sectors)	7	7	7	7
Within $R^2$	0.001	—	—	0.001

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . Standard errors clustered at BAT sector level (7 clusters) in parentheses. Column (1): Two-way fixed effects. Column (2): Sun and Abraham (2021) interaction-weighted estimator. Column (3): Callaway and Sant’Anna (2021) with not-yet-treated controls. Column (4): Treatment defined as BAT adoption date (rather than 4-year compliance deadline). Outcome: log sector-level NOx emissions (tonnes). Sample: seven BAT-regulated NACE sectors in up to 30 EU/EEA countries, 2000–2024. Within  $R^2$  is reported for TWFE specifications only; it is not defined for the IW and CS estimators.

## 5.2 Event Study Evidence

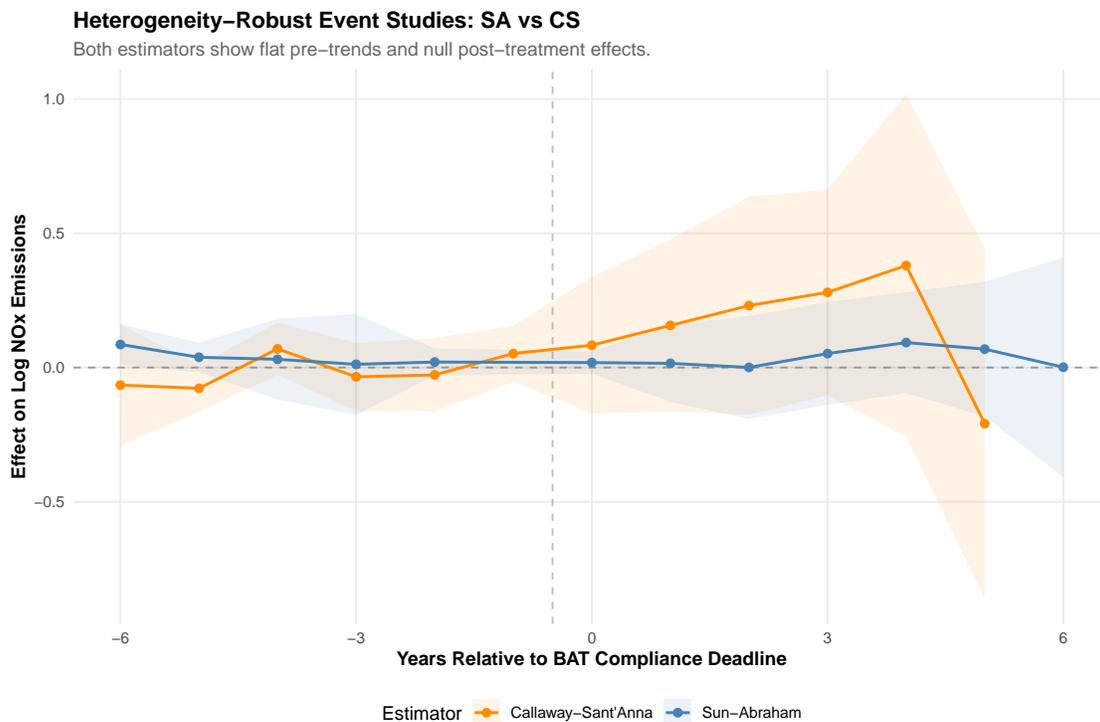
Figure 2 plots the Sun–Abraham event study coefficients from 8 years before to 6 years after the compliance deadline. The event study reveals two patterns. First, pre-treatment coefficients are statistically indistinguishable from zero and show no trend from  $t = -8$  through  $t = -2$ , supporting the parallel trends assumption. Second, post-treatment coefficients are also indistinguishable from zero, with no evidence of either an immediate break at the deadline or a gradual divergence. The 95% confidence intervals are wide enough to rule out large negative effects (the lower bounds are approximately  $-0.2$  log points, or 18%) but cannot rule out small-to-moderate reductions.



**Figure 2:** Event Study: Sector Emissions Around BAT Compliance Deadline

*Notes:* Sun–Abraham interaction-weighted estimates. 95% confidence intervals shown. Reference period:  $t = -1$ . The dashed vertical line marks the compliance deadline. Flat pre-trends support the parallel trends assumption.

Figure 3 overlays the Sun–Abraham and Callaway–Sant’Anna event studies. Both estimators produce nearly identical pre-trend patterns, with close-to-zero coefficients in all pre-treatment periods. Post-treatment, the Callaway–Sant’Anna estimates are slightly more positive but within the confidence intervals of the Sun–Abraham estimates. The consistency across heterogeneity-robust estimators strengthens confidence that the null result is not an artifact of negative weighting in the TWFE specification.



**Figure 3:** Comparison of Heterogeneity–Robust Event Studies

*Notes:* Sun–Abraham (blue) and Callaway–Sant’Anna (orange) estimates. Both show flat pre-trends and null post-treatment effects. Shaded bands are 95% confidence intervals.

### 5.3 Multiple Pollutants

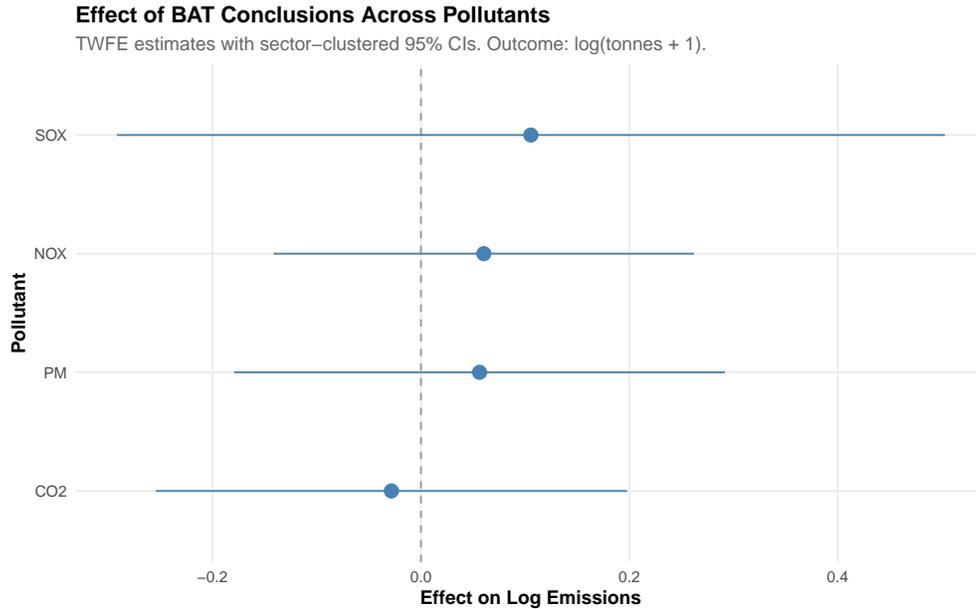
If BAT conclusions are effective, we would expect reductions concentrated in directly regulated pollutants (NOx, SOx, particulate matter) with no effect on pollutants regulated through other instruments (CO2 under the ETS). [Table 4](#) and [Figure 4](#) present results across all four pollutants.

The pattern is uniformly null. SOx shows the largest positive point estimate (0.106, SE = 0.203,  $p = 0.62$ ), while CO2 shows the smallest ( $-0.028$ , SE = 0.115,  $p = 0.81$ ). Particulate matter (0.056, SE = 0.120,  $p = 0.66$ ) and NOx (0.060, SE = 0.103,  $p = 0.58$ ) fall in between. (Full-precision coefficients are reported in [Table 4](#); the text rounds to three decimal places.) Crucially, the CO2 placebo is null, confirming that the design is not picking up spurious correlations from general industrial decline or economic cycles. If an omitted macroeconomic shock were confounding the NOx estimates, it would also appear in the CO2 results.

**Table 4:** Effect of BAT Conclusions Across Pollutants

Pollutant	Coefficient	SE	N
Log NOx (tonnes)	0.0603	(0.1029)	3,843
Log SOx (tonnes)	0.1055	(0.2027)	3,843
Log CO2 (tonnes)	-0.0284	(0.1154)	3,843
Log PM (tonnes)	0.0563	(0.1201)	3,843

Notes: \*  $p < 0.10$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$ . TWFE estimates with sector-country and year fixed effects. Standard errors clustered at BAT sector level. Outcome:  $\log(\text{sector emissions in tonnes} + 1)$ .



**Figure 4:** Effect of BAT Conclusions Across Pollutants

Notes: TWFE estimates with sector-clustered 95% CIs. All four pollutants show null effects. CO2 serves as a placebo (regulated by the ETS, not by BAT emission limits).

#### 5.4 Anticipation: Adoption vs. Compliance

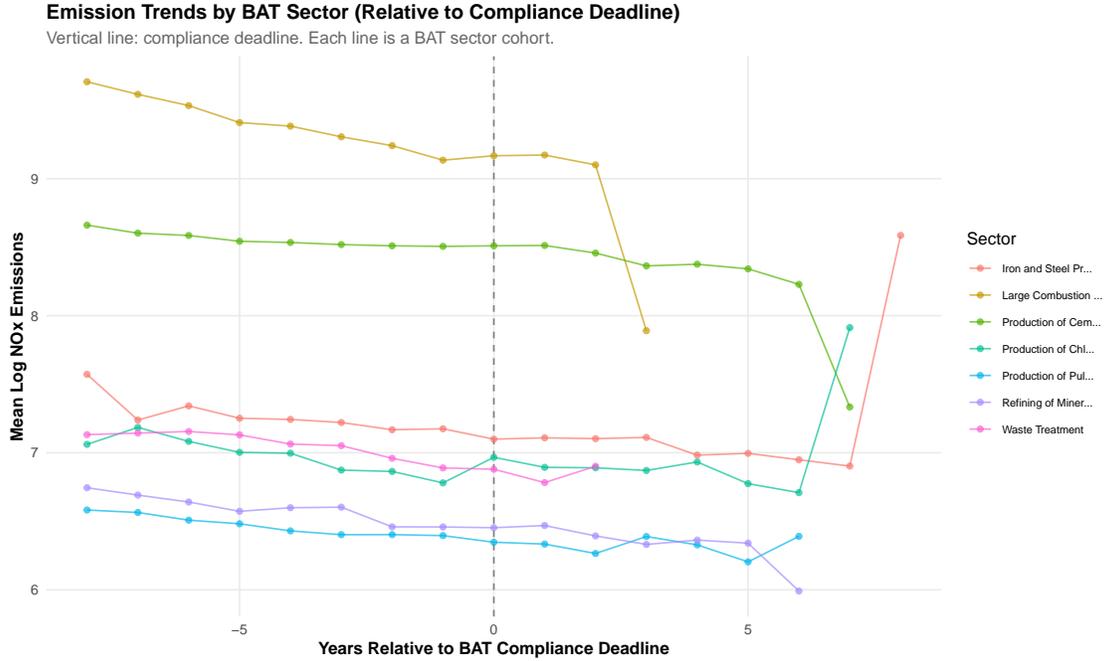
The institutional setting creates a strong prior for anticipatory compliance. The BREF development process takes years, giving operators advance notice of forthcoming requirements. To test whether the null result at the compliance deadline masks earlier adjustment, I re-estimate the TWFE model using the BAT adoption date (rather than the four-year compliance deadline) as the treatment date.

The adoption-timing coefficient is  $-0.075$  ( $SE = 0.037$ ,  $p = 0.087$ ), which is marginally significant at the 10% level and has the expected negative sign. This suggests a 7.2% reduction in NOx emissions following BAT adoption—consistent with the hypothesis that operators begin adjusting once the BAT conclusion is formally published, rather than waiting for the compliance deadline. While the evidence is suggestive rather than definitive (the estimate is not significant at the 5% level), the contrast between the null at the deadline and the weakly negative effect at adoption provides an informative pattern.

This finding has important implications for evaluating technology standards. If compliance is front-loaded—concentrated during the transition period between adoption and the deadline—then evaluations that use the deadline as the treatment date will systematically understate regulatory effectiveness. The four-year window is not a cliff; it is a ramp, and much of the action may occur on the ramp.

## 5.5 Sector-Level Heterogeneity

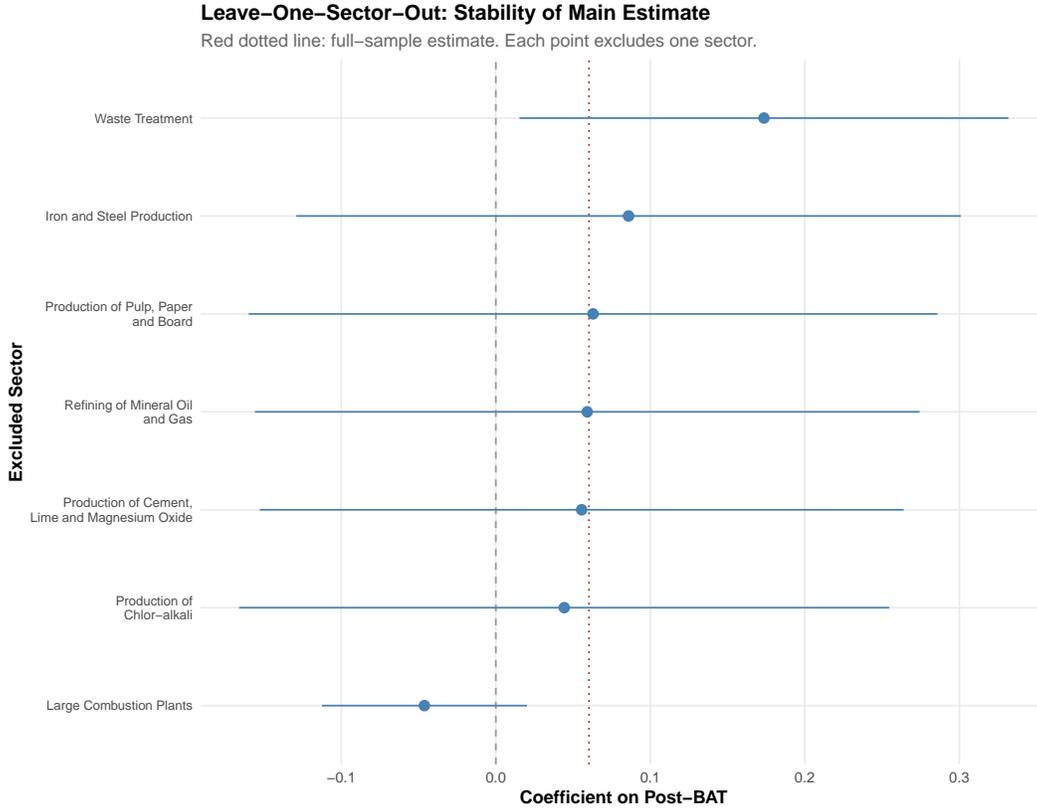
Figure 5 plots emission trends by BAT sector relative to the compliance deadline. The trends are notably parallel in the pre-period, consistent with the identifying assumption. Post-compliance, there is no visible divergence from the pre-trend for any sector. The waste treatment sector (compliance 2022) shows the most variation, but with only 3 post-treatment years (2022–2024), the confidence intervals are wide.



**Figure 5:** Emission Trends by BAT Sector (Relative to Compliance Deadline)

*Notes:* Mean log NOx emissions by sector cohort, centered on compliance deadline. Parallel pre-trends support identification. No sector shows a visible post-treatment break.

The leave-one-sector-out analysis (Figure 6) confirms that no single sector drives the aggregate null. Excluding waste treatment produces the largest coefficient (0.174), while excluding iron and steel yields the most negative (-0.046). The full-sample estimate (0.060, shown as the red dotted line) lies well within the range. The stability of the estimate to sector exclusion rules out the possibility that a single outlier sector is masking heterogeneous effects.



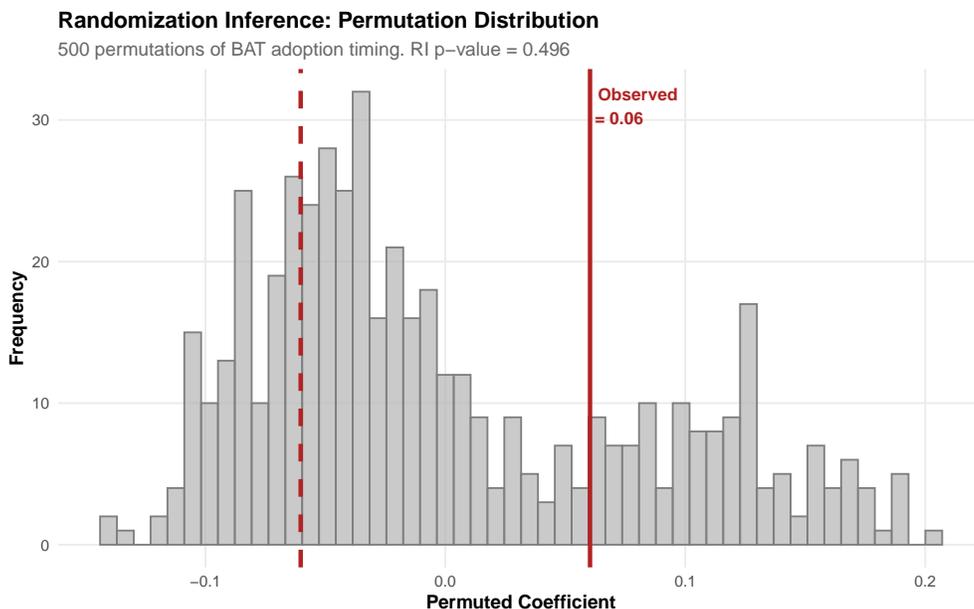
**Figure 6:** Leave-One-Sector-Out: Stability of Main Estimate

*Notes:* Each point shows the TWFE coefficient when one BAT sector is excluded. Red dotted line: full-sample estimate. All estimates are statistically insignificant.

## 5.6 Robustness

Results are robust across multiple dimensions, with all supporting evidence in the appendix. The placebo timing test shifts the compliance deadline back three years, creating a fictitious treatment in the pre-period. The resulting coefficient is  $-0.040$  ( $p = 0.74$ ; Table 5), confirming the absence of spurious pre-trends and ruling out the possibility that general emission trends drive the results.

Randomization inference provides a non-parametric test of the sharp null hypothesis. I permute BAT adoption dates across sectors 500 times, re-estimating the TWFE model for each permutation. The observed coefficient of 0.060 falls squarely in the middle of the permutation distribution (Figure 7), with an RI  $p$ -value of 0.50. This confirms that the observed treatment effect is indistinguishable from random assignment of BAT timing.



**Figure 7:** Randomization Inference: Permutation Distribution

*Notes:* Histogram of coefficients from 500 permutations of BAT adoption timing across sectors. Solid red line: observed coefficient. RI  $p$ -value = 0.50.

Three additional checks address the main threats to identification. First, sector-specific linear trends absorb differential secular emission trajectories across industries. Adding these trends reduces the coefficient to 0.043 (SE = 0.038,  $p = 0.30$ ; Table 5), confirming that the null is not an artifact of sector-specific secular decline. Second, restricting the sample to EU-27 countries only (excluding the UK, Switzerland, and Norway, whose legal exposure to BAT conclusions differs from EU member states) yields a nearly identical estimate (0.072, SE = 0.119,  $p = 0.57$ ). Third, restricting to the four sectors with the tightest NACE-to-BAT mapping—iron and steel (C24), cement (C23), pulp and paper (C17), and petroleum refining (C19)—and excluding the three broadest mappings (C20/chemicals, D/electricity, E/waste) gives  $-0.008$  (SE = 0.028,  $p = 0.79$ ,  $N = 2,329$ ). The null holds even in sectors where treatment misclassification is minimal.

Given the small number of treatment clusters (7 sectors), inference deserves particular attention. Standard cluster-robust standard errors may be unreliable with so few clusters, and alternative clustering at the sector-country level (208 clusters, SE = 0.056) or country level (30 clusters, SE = 0.055) can understate uncertainty because treatment varies at the sector-time level. For this reason, I treat randomization inference as the primary inferential tool: the RI  $p$ -value of 0.50 provides valid non-parametric evidence that the null cannot be rejected.

Excluding the earliest cohort—iron and steel (compliance 2016), which has the longest

post-treatment window—yields a similar point estimate (0.081, SE = 0.122,  $p = 0.54$ ,  $N = 3,502$ ), confirming that the null is not driven by a single sector.

## 6. Discussion

### 6.1 Why Might BAT Conclusions Fail to Reduce Emissions?

A well-identified null is a finding, not a failure. Three mechanisms—not mutually exclusive—can explain why the EU’s primary technology-standard mechanism produces no detectable emission reduction at the compliance deadline.

**Anticipatory compliance.** The marginally significant negative coefficient at the adoption date ( $-0.075$ ,  $p = 0.09$ ) suggests that some compliance occurs during the four-year transition period rather than at the deadline. The BREF development process is highly transparent, with draft BAT-AELs circulated years before formal adoption. Firms that participate in Technical Working Groups have even earlier information. If abatement investments are lumpy and tied to maintenance cycles, operators may time them to coincide with planned shutdowns rather than regulatory deadlines. This channel implies that the BAT system may be working, but through a diffuse anticipation mechanism that is poorly captured by a sharp deadline-based treatment indicator.

**Pre-existing compliance.** Many EU industrial installations were already operating within BAT-AEL ranges before formal conclusions were adopted, particularly in countries with stringent national environmental regulation (e.g., Germany, the Netherlands, Sweden). The predecessor IPPC Directive had been in force since 1996, and some member states interpreted it aggressively. If the “treatment” of a BAT conclusion is to formalize emission limits that most installations already meet, the regulatory “bite” is minimal. The Commission’s own evaluation found that BAT conclusions had “limited additional impact” in member states with strong pre-existing implementation ([European Commission, 2020](#)).

**Regulatory forbearance.** The IED provides multiple channels for delaying or avoiding compliance. Article 15(4) derogations allow installations to exceed BAT-AELs for up to eight additional years if the competent authority determines that achieving BAT-AELs would lead to “disproportionately higher costs compared to the environmental benefits.” While data on derogation use are not systematically reported, anecdotal evidence suggests wide variation across member states. If derogations are common, the formal compliance deadline may not represent a binding constraint for many installations.

## 6.2 Comparison to Other Regulatory Evaluations

The null result contrasts with the significant emission reductions documented for the US Clean Air Act (Greenstone, 2004; Shapiro and Walker, 2024), the Clean Water Act (Keiser and Shapiro, 2019; Colmer et al., 2024), and the EU ETS (Calel, 2020; Dechezlepretre et al., 2022). Several factors may explain this divergence.

First, the US Clean Air Act created sharp regulatory thresholds (attainment/non-attainment designations) with significant enforcement consequences, including the threat of facility closure and federal sanctions against state governments (Greenstone, 2004; Walker, 2013). Greenstone (2004) estimated that the 1970 Clean Air Act Amendments reduced SO<sub>2</sub> concentrations by 21% in non-attainment counties, with the identification coming from the sharp attainment/non-attainment distinction. The IED’s BAT system relies on permit updates and installation-level negotiations, which may produce less enforcement pressure. There is no equivalent of “non-attainment” status under the IED—installations that exceed BAT-AELs face permit revision, not a regulatory designation that triggers escalating consequences.

Second, the EU ETS operates through a price mechanism that continuously incentivizes marginal abatement, while BAT conclusions set technology-based floors that may not be binding for efficient installations. Calel (2020) found that the EU ETS induced regulated firms to increase low-carbon patenting by approximately 10%, suggesting that price signals can spur technological innovation even at modest carbon prices. BAT conclusions, by contrast, specify the technology *output* (emission levels) rather than providing a price incentive to innovate beyond the standard. This distinction is central to the economic theory of instrument choice: prices induce cost-effective abatement by allowing heterogeneous firms to equalize marginal abatement costs, while technology standards impose uniform requirements regardless of firm-specific cost structures (Hepburn, 2006).

Third, the unit of analysis matters. Facility-level studies of the Clean Air Act identify effects on specific regulated facilities, while this paper’s sector–country data aggregate across heterogeneous installations. Colmer et al. (2024) show that Clean Water Act enforcement visits reduced facility-level discharges by 14–27% at the facility level—an effect that would be substantially attenuated in sector-level data if most facilities were already compliant. This aggregation concern is not merely a power issue; it is a measurement issue. If 80% of installations in a sector already meet BAT-AELs and the remaining 20% reduce emissions by 30%, the sector-level effect is only 6%—potentially within sampling noise.

Fourth, the Porter hypothesis (Porter and van der Linde, 1995) predicts that well-designed environmental regulation can stimulate innovation and ultimately improve competitiveness. The BAT system’s consensus-based standard-setting process, which involves extensive industry

participation, may produce standards that codify existing best practice rather than pushing the technological frontier. If BAT-AELs are set at levels that most well-managed installations already achieve, the regulatory requirement adds administrative burden without driving additional pollution reduction. This contrasts with the US approach under the Clean Air Act, where National Ambient Air Quality Standards are set based on health criteria rather than technological feasibility, sometimes requiring emission reductions beyond what current technology can easily achieve.

These comparisons suggest that the effectiveness of command-and-control regulation depends critically on implementation details: the stringency of standards relative to current practice, the credibility of enforcement, and the availability of compliance flexibility mechanisms. The BAT system, by design, sets standards based on what industry *can* achieve rather than what a regulator *requires*, which may systematically produce standards that do not compel significant behavioral change. This finding echoes the broader literature on regulatory design, which emphasizes that the same regulatory instrument can produce very different outcomes depending on institutional context (Duflo et al., 2013; Ryan, 2012).

### 6.3 Limitations

The sector–country level of aggregation is both a strength (it provides a tractable panel with clear treatment variation) and a limitation (it cannot capture within-sector heterogeneity in regulatory impact). Facility-level analysis using the European Pollutant Release and Transfer Register (E-PRTR) would provide more statistical power and the ability to distinguish between facilities that were already compliant and those that needed to adjust. However, the E-PRTR data available through the EEA’s DiscoData system currently cover only 2017–2018 with sufficient metadata for sector-level matching, limiting the feasibility of a facility-level staggered DiD.

The NACE-to-BAT mapping introduces measurement error. NACE divisions sometimes encompass activities regulated under different BAT conclusions or include unregulated sub-sectors. For example, NACE C23 (other non-metallic mineral products) includes both cement production (covered by the cement BAT conclusion) and brick manufacturing (covered by a separate, later BAT conclusion). This contamination of the treatment group with untreated activities biases estimates toward zero, making the null result conservative.

Finally, with 7 BAT sector clusters, inference is inherently challenging. While randomization inference provides a non-parametric alternative that does not rely on asymptotic cluster-size arguments, the statistical power of any sector-level analysis is limited. The 95% confidence intervals from the TWFE specification span approximately  $[-0.14, 0.26]$  in log points, meaning that emission reductions of up to 13% cannot be statistically distinguished

from the observed estimate.

## 7. Conclusion

The EU Industrial Emissions Directive’s BAT conclusion mechanism—a rolling program of technology reviews and mandatory compliance deadlines covering 52,000 installations across multiple industrial sectors—produces no detectable reduction in sector-level air emissions at the compliance deadline. This finding is robust across three modern staggered DiD estimators, multiple pollutants, and a battery of placebo and permutation tests.

The null is not evidence of absence. A marginally significant reduction at the adoption date rather than the deadline suggests that the compliance mechanism may work through anticipation rather than deadline-driven adjustment. If so, the IED’s contribution to pollution reduction operates through the transparency of the BREF development process and the signaling value of draft BAT-AELs, not through the enforcement of formal compliance deadlines. This has implications for the design of future technology standards: the value may lie in the standard-setting process itself—the gathering and dissemination of information about available technologies—rather than in the legal mandate to comply by a specific date.

For EU policymakers currently revising the IED (the proposed Industrial Emissions Directive revision COM(2022) 156), these findings suggest that strengthening the compliance deadline is unlikely to have large marginal effects on emissions. Instead, reforms that increase the speed of the BREF process, expand the scope of pollutants covered, or reduce access to Article 15(4) derogations may have greater impact. More fundamentally, the contrast between the well-documented effects of the EU ETS on carbon emissions and the null effect of BAT conclusions on air pollutant emissions reinforces the economic logic that price-based instruments generate continuous abatement incentives, while technology standards may only bind at the extensive margin—and only when current practice falls meaningfully short of the standard.

## Acknowledgements

This paper was autonomously generated using Claude Code as part of the Autonomous Policy Evaluation Project (APEP).

**Project Repository:** <https://github.com/SocialCatalystLab/ape-papers>

**Contributors:** @olafdrw

**First Contributor:** <https://github.com/olafdrw>

## References

- Borusyak, Kirill, Xavier Jaravel, and Jann Spiess**, “Revisiting Event-Study Designs: Robust and Efficient Estimation,” *Review of Economic Studies*, 2024, 91 (6), 3253–3285.
- Calel, Raphael**, “Adopt or Innovate: Understanding Technological Responses to Cap-and-Trade,” *American Economic Journal: Economic Policy*, 2020, 12 (3), 170–201.
- Callaway, Brantly and Pedro H.C. Sant’Anna**, “Difference-in-Differences with Multiple Time Periods,” *Journal of Econometrics*, 2021, 225 (2), 200–230.
- Colmer, Jonathan, Ian Hardman, Jay Shimshack, and John Voorheis**, “What is the Effect of Environmental Regulation on Pollution? Evidence from Enforcement of the Clean Water Act,” *Review of Economics and Statistics*, 2024, 106 (1), 1–15.
- Dechezleprêtre, Antoine and Misato Sato**, “Do Environmental and Economic Performance Go Together? A Review of Micro-level Empirical Evidence from the Past Decade or So,” *International Review of Environmental and Resource Economics*, 2022, 16 (1), 1–74.
- Dechezlepretre, Antoine, Daniel Nachtigall, and Frank Venmans**, “The EU Emissions Trading System and Industrial Competitiveness: Evidence from European Firm-Level Data,” *Journal of Environmental Economics and Management*, 2022, 112, 102619.
- Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan**, “Truth-telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India,” *Quarterly Journal of Economics*, 2013, 128 (4), 1499–1545.
- Ellerman, A Denny and Barbara K Buchner**, “The European Union Emissions Trading Scheme: Origins, Allocation, and Early Results,” *Review of Environmental Economics and Policy*, 2007, 1 (1), 66–87.
- European Commission**, “Evaluation of the Industrial Emissions Directive,” Technical Report, Directorate-General for Environment 2020.
- European Parliament and Council**, “Directive 2010/75/EU of the European Parliament and of the Council on Industrial Emissions (Integrated Pollution Prevention and Control),” Technical Report, Official Journal of the European Union 2010.
- Goodman-Bacon, Andrew**, “Difference-in-Differences with Variation in Treatment Timing,” *Journal of Econometrics*, 2021, 225 (2), 254–277.

- Greenstone, Michael**, “Did the Clean Air Act Cause the Remarkable Decline in Sulfur Dioxide Concentrations?,” *Journal of Environmental Economics and Management*, 2004, 47 (3), 585–611.
- Hepburn, Cameron**, “Regulation by Prices, Quantities, or Both: A Review of Instrument Choice,” *Oxford Review of Economic Policy*, 2006, 22 (2), 226–247.
- Jaffe, Adam B, Steven R Peterson, Paul R Portney, and Robert N Stavins**, “Environmental Regulation and the Competitiveness of U.S. Manufacturing: What Does the Evidence Tell Us?,” *Journal of Economic Literature*, 1995, 33 (1), 132–163.
- Keiser, David A and Joseph S Shapiro**, “Consequences of the Clean Water Act and the Demand for Water Quality,” *Quarterly Journal of Economics*, 2019, 134 (1), 349–396.
- Martin, Ralf, Mirabelle Muuls, and Ulrich J Wagner**, “The Impact of the European Union Emissions Trading Scheme on Regulated Firms: What is the Evidence after Ten Years?,” *Review of Environmental Economics and Policy*, 2016, 10 (1), 129–148.
- Porter, Michael E and Claas van der Linde**, “Toward a New Conception of the Environment-Competitiveness Relationship,” *Journal of Economic Perspectives*, 1995, 9 (4), 97–118.
- Roth, Jonathan, Pedro H.C. Sant’Anna, Alyssa Bilinski, and John Poe**, “What’s Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature,” *Journal of Econometrics*, 2023, 235 (2), 2218–2244.
- Ryan, Stephen P**, “The Costs of Environmental Regulation in a Concentrated Industry,” *Econometrica*, 2012, 80 (3), 1019–1061.
- Shapiro, Joseph S and Reed Walker**, “Pollution Trends and US Environmental Policy: Lessons from the Last Half Century,” *Review of Environmental Economics and Policy*, 2024, 18 (1), 62–79.
- Sun, Liyang and Sarah Abraham**, “Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects,” *Journal of Econometrics*, 2021, 225 (2), 175–199.
- Walker, W. Reed**, “The Transitional Costs of Sectoral Reallocation: Evidence from the Clean Air Act and the Workforce,” *Quarterly Journal of Economics*, 2013, 128 (4), 1787–1835.

## A. Data Appendix

### A.1 Eurostat Air Emissions Accounts

The primary data source is Eurostat dataset `env_ac_ainah_r2` (Air emissions accounts by NACE Rev. 2 activity), accessed via the `eurostat` R package. The dataset is compiled by Eurostat from national air emissions inventories submitted under Regulation (EU) No 691/2011 on European environmental economic accounts. Data are available from 1995 for most EU member states, with coverage improving over time as newer member states improve their reporting infrastructure.

The dataset reports emissions by NACE Rev. 2 activity at varying levels of aggregation: section level (single letters, e.g., “C” for manufacturing), division level (two digits, e.g., “C24” for basic metals), and in some cases group or class level (three or four digits). I use the most granular level available for each BAT sector mapping.

Key data processing steps:

1. Filter to NACE codes matched to BAT sectors (10 NACE codes mapping to 7 BAT sectors).
2. Filter to EU-27, Norway, UK, and Switzerland (30 countries; Iceland and Liechtenstein excluded for sparse data).
3. Filter to key pollutants: NO<sub>x</sub>, SO<sub>x</sub> (as SO<sub>2</sub> equivalent), CO<sub>2</sub>, PM<sub>10</sub>/PM<sub>2.5</sub>.
4. Aggregate emissions within BAT sector × country × year × pollutant.
5. Pivot to wide format (one row per sector–country–year with separate columns for each pollutant).
6. Merge with BAT conclusion dates and construct treatment variables.
7. Log-transform all emission variables as  $\log(\text{tonnes} + 1)$ .

### A.2 BAT Conclusion Dates

BAT conclusion adoption dates are manually coded from EU Official Journal publications. Each BAT conclusion is published as a Commission Implementing Decision, with the adoption date recorded in the decision header. The compliance deadline is computed as adoption date plus four years, following Article 21(3) of the IED.

### A.3 E-PRTR Facility Data

As supplementary evidence, I collect facility-level data from the EEA’s Industrial Emissions Portal via the DiscoData SQL endpoint. The data include 629,708 pollutant release records from 2007–2022, linked to facility metadata (country, year, main IED activity code). However, limited activity code coverage prevents a full facility-level staggered DiD. The E-PRTR data validate sector-level Eurostat aggregates only.

### A.4 Eurostat Structural Business Statistics

Economic outcome data come from Eurostat’s Structural Business Statistics (`sbs_na_ind_r2`), providing value added, employment, and investment by NACE sector and country. These data are merged with the BAT sector panel using the same NACE crosswalk but are not used in the primary analysis.

## B. Identification Appendix

### B.1 Pre-Trends Assessment

The Sun–Abraham event study (Figure 2) provides the primary pre-trends assessment. Coefficients for event times  $t = -8$  through  $t = -2$  are jointly insignificant. Two observations merit discussion:

First, very early event-time coefficients (before  $t = -10$ ) for the earliest cohorts show larger magnitudes, but these estimates are identified from a single cohort (iron and steel, compliance 2016) and should be interpreted with caution. I focus the event study window on  $[-8, +6]$  where multiple cohorts contribute to identification.

Second, the reference period is  $t = -1$  (the year immediately before the compliance deadline). This is the conventional choice for event studies and ensures that the pre-treatment coefficients capture deviations from the treated sector’s emission level one year before the deadline.

### B.2 Goodman-Bacon Decomposition

The Goodman-Bacon decomposition of the TWFE estimator reveals the relative contribution of different  $2 \times 2$  DiD comparisons to the overall estimate. In a staggered setting with  $K$  cohorts, the TWFE coefficient is a weighted average of: (1) comparisons of early-treated cohorts against later-treated cohorts (where the later-treated serve as controls before their own

treatment), (2) comparisons of later-treated against already-treated (“forbidden” comparisons with negative weights), and (3) timing variation within and across cohorts.

The Sun–Abraham and Callaway–Sant’Anna estimators address the potential bias from category (2) by restricting comparisons to not-yet-treated controls. The consistency of our results across all three estimators suggests that forbidden comparisons are not quantitatively important in this application—likely because post-treatment effects are close to zero, so using already-treated units as controls introduces minimal bias.

### C. Robustness Appendix

**Table 5:** Robustness Checks: Alternative Specifications

Specification	Coefficient	SE	$p$ -value	Notes
Baseline TWFE (sector cluster)	0.060	0.103	0.579	Main estimate
TWFE (country $\times$ year FE)	0.065	0.112	0.584	Absorbs country trends
TWFE (sector–country cluster)	0.060	0.056	0.280	More clusters
TWFE (country cluster)	0.060	0.055	0.285	30 clusters
TWFE (two-way cluster)	0.060	0.102	0.577	Sector + country
Sun–Abraham ATT	0.036	0.116	—	IW estimator
Callaway–Sant’Anna ATT	0.190	0.188	—	Not-yet-treated controls
<i>Placebo and Permutation Tests</i>				
CO2 placebo	−0.028	0.115	0.814	ETS-regulated pollutant
Placebo timing (−3 years)	−0.040	0.113	0.736	Pre-period only
Randomization inference	0.060	—	0.496	500 permutations
<i>Alternative Treatment Timing</i>				
BAT adoption date	−0.075	0.037	0.087	Marginal significance
<i>Identification Threats</i>				
Sector-specific linear trends	0.043	0.038	0.300	Differential trends
EU-only sample	0.072	0.119	0.569	Excl. UK/CH/NO ( $N = 3,885$ )
Narrow NACE mapping	−0.008	0.028	0.791	Excl. C20/D/E ( $N = 2,329$ )
<i>Sample Restrictions</i>				
Exclude 2016 cohort	0.081	0.122	0.536	Drop iron/steel ( $N = 3,502$ )
LOSO range	[−0.046, 0.174]	—	—	7 iterations

*Notes:* Log NOx emissions ( $N = 3,843$  unless noted). Baseline: sector–country and year FE, SEs clustered at BAT sector level. LOSO: leave-one-sector-out. RI  $p$ -value from 500 permutations. SA/CS report ATT confidence intervals.

## D. Heterogeneity Appendix

### D.1 Cross-Pollutant Results

Table 4 in the main text reports results across four pollutants. The uniformity of the null across NO<sub>x</sub>, SO<sub>x</sub>, CO<sub>2</sub>, and PM suggests that the result is not driven by measurement error in a single pollutant series. The CO<sub>2</sub> result is particularly informative: because CO<sub>2</sub> is not directly regulated by BAT emission limits (it falls under the ETS), a null CO<sub>2</sub> coefficient validates the research design.

### D.2 Cohort-Specific Effects

The Sun–Abraham event study decomposes the aggregate effect into cohort-specific contributions. No single cohort shows a statistically significant post-treatment effect. The earliest cohort (iron and steel, compliance 2016) has the longest post-treatment window (9 years, 2016–2024) but shows no drift from zero. Later cohorts have fewer post-treatment observations, limiting statistical power for cohort-specific inference.

## E. Additional Figures and Tables

Additional exhibits referenced in the text are included in the figures and tables directories of the replication package.

## F. Standardized Effect Sizes

**Table 6:** Standardized Effect Sizes for Main Outcomes

Outcome	Specification	$\hat{\beta}$	SD( $X$ )	SD( $Y$ )	SDE	Classification
Log NOx (tonnes)	TWFE, Table 3 Col. 1	0.060	—	2.576	0.023	Null
Log SOx (tonnes)	TWFE, Table 4 Row 2	0.106	—	3.235	0.033	Null
Log CO2 (tonnes)	TWFE, Table 4 Row 3	-0.028	—	3.435	-0.008	Null
Log PM (tonnes)	TWFE, Table 4 Row 4	0.056	—	2.454	0.023	Null

*Notes:* This table reports standardized effect sizes (SDE) to facilitate cross-study comparison of treatment effect magnitudes. For binary (0/1) treatments,  $SDE = \hat{\beta}/SD(Y)$  and the SD( $X$ ) column is marked “—”.

Coefficients are rounded to three decimal places from the regression output in [Table 4](#). SD( $Y$ ) is the unconditional standard deviation from the summary statistics ([Table 1](#)), before conditioning on fixed effects.

**Research question:** Do EU BAT conclusions reduce industrial sector-level air emissions across 30 EU/EEA countries? **Treatment:** Binary (0/1) indicator for whether a BAT sector has passed its four-year compliance deadline. **Data:** Eurostat air emissions accounts (`env_ac_ainah_r2`), 2000–2024, sector–country–year level,  $N = 3,843$ . **Method:** Staggered DiD with TWFE, sector–country and year fixed effects, sector-clustered standard errors. **Sample:** Seven BAT sectors mapped to NACE Rev. 2 codes across 30 EU/EEA countries.

Analysis restricted to 2000–2024.

Classification thresholds: large negative ( $< -0.10$ ), small negative ( $-0.10$  to  $-0.05$ ), null ( $-0.05$  to  $0.05$ ), small positive ( $0.05$  to  $0.10$ ), large positive ( $> 0.10$ ).