Energy Institute WP 355

Benefits, Costs, and Distributional Outcomes in an Environmental Market

Olivier Deschenes and Paige Weber

December 2025

Benefits, Costs, and Distributional Outcomes in an Environmental Market*

Olivier Deschenes Paige Weber

December 10, 2025

Abstract

We study the benefits, abatement costs, and distributional outcomes of the EPA's NO_x Budget Program, a seasonal cap-and-trade market for NO_x emissions. Using a new empirical approach, we recover source-specific marginal abatement cost curves and combine them with source-specific air pollution damages to quantify market outcomes. We find that abatement costs under the market are roughly one-sixth those of an abatement-equivalent non-market policy. Despite these large differences in aggregate costs, the market and the non-market policy yield similar air quality benefits, both in overall magnitude and across demographic groups. We show that these results reflect the weak correlation between source-specific marginal damages and abatement behavior.

Keywords: Environmental markets, distributional outcomes, air pollution, electricity sector regulation
JEL classification codes: Q52, L94, Q58

*Deschenes: Department of Economics, University of California, Santa Barbara, and NBER, email: olivier@econ.ucsb.edu. Weber: Energy and Resources Group, University of California, Berkeley, email: paigeweber@berkeley.edu. We thank numerous seminar participants and colleagues for their helpful comments, especially Jonathan Colmer, Meredith Fowlie, and Erin Mansur. We have not received financial compensation for this project nor do we have financial relationships related to this research. The analysis relies entirely on publicly-available data and all data and code will be posted upon completion of the project.

1 Introduction

In the last decades several notable market-based policies to address environmental externalities have been implemented.[1] Environmental markets achieve cost-effective emissions reductions by providing flexibility on how much regulated sources abate. While this flexibility lowers the cost of the regulation compared to alternative prescriptive policies that achieve the same emissions reduction, it also generates ex-ante uncertainty in the spatial distribution of abatement. In contrast to uniform pollutants like greenhouse gases, for which spatial variation in abatement is inconsequential, this variation is critical for determining the benefits from non-uniform pollutants such as air pollution.[2]

The benefits of an environmental market for a non-uniform pollutant depend on the following factors: one, the spatial distribution of regulated sources; two, the abatement induced by the market at each regulated source; and, three the marginal damages of pollution released by each regulated source. To date, no study has empirically analyzed how emissions markets impact the distribution of non-uniform pollutants by separately examining the contribution of each of these three determinants of distributional impacts.[3]

This paper develops a methodology to estimate source-specific abatement under a cap-and-trade system, and quantifies how a market-based regulation changes the spatial distribution and marginal damages of air pollution, accounting for source-specific atmospheric transport. We implement this approach using data on emissions from all fossil-fueled elec-

[1] For example, California's Cap-and-Trade Program launched in 2012, the Regional Greenhouse Gas Initiative in 2009, the European Union's Emissions Trading Scheme in 2005, the U.S. Acid Rain Program in 1995 (Phase I), and the Regional Clean Air Incentives Market in Southern California in 1994.

[2] For a uniform pollutant, where marginal damages are equal regardless of the location of the source of emissions, to regulate pollution more efficiently than non-market approaches. However, this efficiency property does not extend to non-uniform pollutants, where marginal damages vary across sources. Muller and Mendelsohn (2009) and Doyle et al. (2014) discuss how trading ratios can be used to improve efficiency in markets for non-uniform pollutants.

[3] For example, Hernandez-Cortes and Meng (2023) include transport and receptor locations but not source-specific abatement. Fowlie et al. (2012) using a matching based approach to estimate average treatment effect and heterogeneous effect by income and minority status, though do not include a pollution transport model. Grainger and Ruangmas (2018) uses an air pollution transport model with the same identification approach as Fowlie et al. (2012). Shapiro and Walker (2021) focuses locations and prices of permit trades and local characteristics but does not estimate source-specific marginal damages.

tricity generating units (units or sources) in the United States and focus on the NO_x Budget Program (NBP)—a seasonal cap-and-trade market for nitrogen oxides (NO_x), a precursor of ozone and fine particulate matter. The NBP market operated from 2003 to 2008 and capped emissions from 1,769 fossil-fueled electricity generating units located across the Midwest and Eastern United States during the ozone season (May through September) each year. In contrast to previous studies, our empirical strategy leverages observed source-level emissions data together with institutional variation in regulatory status to identify heterogeneous abatement responses to the cap-and-trade market.[4] In turn, this allows us to estimate the source-specific and aggregate costs of the market-based policy, and then compare these costs and benefits to counterfactual non-market policy that achieves the same aggregate amount of emissions reduction.

We leverage the spatial, seasonal, and annual variation in treatment status across regulated sources to estimate a model representing counterfactual NO_x emissions in a counterfactual state without the environmental market. These correspond to the NO_x emissions from regulated sources during the ozone season months absent the cap imposed by the NBP market. Our model for counterfactual emissions includes unit and time fixed effects as well as predictors of demand for emissions (such as weather and market load), and we show it performs notably well in cross-validation exercises. In addition, we show that our estimated counterfactual emissions track the average monthly emissions of unregulated sources exceptionally well, mirroring the intricate seasonal patterns of emissions over time. The difference between observed emissions and predicted counterfactual emissions generates source-specific estimates of market-driven abatement. We use these estimates to construct source-specific marginal abatement cost (MAC) curves, which we apply to estimate the total abatement cost of the regulation. Importantly, these source-specific MAC curves allow us to simulate counterfactual abatement costs for alternative emission-control policies using only observed

[4] Previous research has considered heterogeneous responses across regulated sources by pre-specifying the dimensions of heterogeneity, for example, see Fowlie et al. (2012). Shapiro and Walker (2025) allow for heterogeneity in abatement costs across markets but not across sources within a market. Greenstone et al. (2025) use rich emissions permit bidding data to infer heterogeneity in abatement costs across sources.

emissions data from regulated sources.[5]

Our empirical analysis starts with implementing a standard triple-difference estimator in Section 3.1 that matches the three levels of identifying variation generated by the NBP market (inside-outside the NBP market region, before-after the start of the cap-and-trade policy, and during-outside months of the summer, May–September, ozone season). Using monthly emissions data from a sample of 3,722 electricity generating units in the continental US for 2000–2008, we find that the NBP market led to a 28% reduction in monthly NO_x emissions, corresponding to an abatement of about 40 tons of NO_x per month on average, and consistent with the results in Deschenes et al. (2017). We also find that the NBP market did not change emissions of other pollutants such as sulfur dioxide and carbon dioxide, and that the reduction in NO_x emissions is almost entirely attributable to abatement from coal units as opposed to gas and oil units. A standard "event-study" analysis finds no evidence of pre-NBP differences in relative trends in NO_x emissions.

Next, we implement our approach to estimate source-level impacts of the NBP market on NO_x emissions. The analysis reveals considerable treatment effect heterogeneity, which we interpret as abatement behavior induced by the market. While the majority of regulated sources reduce emissions, some sources increase emissions, highlighting the uneven distribution of abatement across sources. Coal units exhibit the largest changes in emissions, with an average abatement of –112 tons per month during the ozone season months of the market period and a standard deviation of 229 tons across units. For coal units the change in emission ranges from -1,038 tons per month to +165 tons per month during the ozone season months of the market period.[6]

We use our source-level treatment effect estimates in a predictive regression to identify

[5] Greenstone et al. (2025) construct source-level MAC curves and estimate compliance costs in counterfactual policies by making using of rich permit bidding data. Such data has not been available in other settings and the previous literature on environmental markets has otherwise not been able to estimate source-level MAC curves.

[6] The average abatement associated with the market is smaller for gas and oil units, at -0.26 and -1.80 tons per month, respectively. Heterogeneity in abatement across gas and oil units is significantly larger than across coal units.

key sources of heterogeneity across fuel types. Among coal units, those adopting selective catalytic reduction or combustion modification technologies show larger emission reductions in response to the NBP market. Coal units in regulated electricity markets, that is, markets not subject to restructuring, reduced NO_x emissions by 39 tons per ozone-season month, a 45% decline relative to non-treated units. This is consistent with the findings in Fowlie (2010), who shows that units in deregulated markets were less likely to adopt capital-intensive abatement technologies due to weaker incentives to invest in emissions control. Older and larger units also tend to abate more.

Section 3.3 shows how we leverage our source-level estimates of abatement to construct source-specific marginal abatement cost curves. From these MAC curves, we estimate that the total abatement cost of the NBP market was $413 million, around one sixth of the $2.5 billion benefits estimated by Deschenes et al. (2017).[7] We also use our estimates of source-specific MAC curves to construct the costs of a counterfactual non-market based regulation. This non-market counterfactual regulation achieves the same amount of aggregate ozone season NO_x abatement as the NBP, but instead requires each regulated coal unit to reduce its emissions by an equal share. We find that this 'uniform reduction policy' is around six times more costly as the NBP market for the same quantity of avoided NO_x emissions.

In the final part of the paper, Section 4, we demonstrate how to conduct an empirical analysis of the distribution of costs and benefits of a market-based environmental market with non-uniform pollutants. This requires comparing the distribution of marginal abatement costs across all sources to the distribution of marginal damages across all exposure end-points. We focus on the subset of regulatory benefits that can be quantified at the "receptor scale" in our setting, where source-specific marginal damages are observable. These include the benefits from avoided air pollution driven by lower ambient $PM_{2.5}$ concentrations. In the case of the NBP, we find that regulated sources' marginal damages from their contribution to $PM_{2.5}$ concentrations are not systematically related to their marginal abatement costs. This

[7] Deschenes et al. (2017) measure the benefits of the NBP using avoided pharmaceutical purchases and monetized avoided mortality attributable to the market.

result also holds when marginal damages are estimated separately by demographic group. Taken together, these findings suggest that the NBP market is neither inherently misaligned nor well suited to addressing demographic disparities in local air pollution exposure. As a result, we find little differences in the benefits, in both levels and distribution, achieved from the market and the simulated non-market policy.

We interpret these findings on the distribution of air quality benefits through the lens of a simple model of environmental markets where the quantity of emission abatement from a regulated pollution source depends on its marginal abatement cost curve. In this model, regulated sources with flatter MAC curves will reduce emissions more in response to the market regulation compared to sources with steeper MAC curves. Thus, individuals receiving air pollution (i.e., residing "downwind") from regulated sources with flatter MAC curves—smaller slope in absolute magnitude—will accrue more air quality benefits from an environmental market compared to individuals downwind from units with steeper—large slope in absolute magnitude—MAC curves. We demonstrate that the the joint distribution of community demographics and the slope of regulated sources MAC curves is sufficient to predict the impact of a market-based regulation on the distribution of air pollution. We then empirically estimate this joint distribution in our setting.

This study contributes and advances the literature on the distributional outcomes of environmental markets. Notably among this work, Fowlie (2010) and Grainger and Ruangmas (2018) study the distribution of benefits from an environmental market for air pollution in Southern California, RECLAIM. The former does not find evidence that the market is at odds with equity goals. The latter paper models air pollution dispersal, relaxing the assumption of uniform pollution dispersal, and finds larger air pollution reduction in higher income areas. They also find that Black communities benefit compared to White, while Hispanic communities lose. Shapiro and Walker (2021) study the movement of pollution based on locations, quantities, and prices of offset transactions in twelve prominent offset markets under the Clean Air Act. They find that the offset markets do not meaningfully increase or decrease demographic-based air pollution disparities. Hernandez-Cortes and Meng (2023) study the impact of the California carbon market, a uniform pollutant, on the distribution of local air pollution. Their results indicate that the GHG market does not exacerbate demographic-based pollution disparities and may contribute to reducing them.

There is now a substantial empirical literature studying the multifaceted effects of air pollution markets (see, e.g., Colmer et al. (2024), Dechezleprêtre et al. (2023), Fowlie et al. (2012), Martin et al. (2014), Meng and Thivierge (2025), Hernandez-Cortes and Meng (2023), Shapiro and Walker (2025)). We contribute to this literature in several important ways. First, we use a frontier econometric technique for treatment effect heterogeneity to estimate source-specific responses to the environmental market using observed emissions data. To the best of our knowledge, our paper is the first to use this approach to conduct such an analysis. Then, we interpret these responses within a simple model of environmental markets, which we use to translate estimates of source-specific abatement into source-specific marginal abatement cost curves, which we compare to source-specific damages. Shapiro and Walker (2025) is the most comparable paper in this regard: they use Clean Air Act offset transactions prices to construct mean market-level marginal abatement costs, which they compare to market-level damages. Our approach is distinct in that we estimate source-level MAC costs and we estimate the slope on the curve and not just a point. Source-specific marginal abatement cost curves in turn allow us to construct total market abatement costs and to simulate costs under alternative counterfactual policies. This enables us to compare the aggregate abatement costs and the distribution of ambient air pollution resulting from an environmental market with those under a "command-and-control" policy that constrains all sources to reduce emissions by the same proportion. Greenstone et al. (2025) is also highly comparable, where using rich emissions permit bidding data, they recover source-specific MAC curves. Our paper is unique in that we match source-specific MAC curves with source-specific marginal damages based on pollution transport and receptor populations, which is essential to study distribution outcomes of an environmental market.[8]

Finally, our estimates of source-specific abatement allow us to connect theoretical predictions about the distributional impacts of environmental markets to empirically estimated outcomes. Prior empirical research, as well as the theoretical and simulation-based work of Baumol and Oates (1988), Muller and Mendelsohn (2009), and Doyle et al. (2014), shows that the distributional consequences of environmental markets may vary substantially, depending on how abatement costs and marginal damages are distributed spatially. Our paper is the first to estimate both of the metrics empirically at the source-level, which allows us to demonstrate how the relationship between source-specific marginal abatement costs and damages informs the distribution of benefits under an environmental market.

2 Institutional background and data

2.1 The NO_x Budget Program

We study the NO_x Budget Program (NBP), an emission trading market for nitrogen oxides (NO_x) from industrial stationary sources. The NBP instituted a cap-and-trade market for seasonal NO_x emissions in 20 states and Washington, D.C., in the eastern U.S. from 2003 to 2008, with 8 states entering the market in 2003, 11 states entering in 2004, and one entering in 2007 (Missouri), which we exclude from our main estimation sample.[9] The NBP regulated fossil-fueled electricity generating units (units or sources) and industrial boilers, the two primary stationary sources of NO_x. As Appendix Figure 1 shows, in general all counties in a regulated state were covered by NBP. However not all counties in Alabama, Michigan, and Missouri were regulated by the NBP market. For our analysis, units located

[8] Greenstone et al. (2025) model benefits as accruing to the entire population of the regulated region without regard to pollution transport or the distribution of pollution across different population sub-groups.

[9] The Ozone Transport Commission Market, a smaller cap market regulating EGUs in Connecticut, Delaware, Maryland, Massachusetts, New Jersey, New York, Pennsylvania, Rhode Island, and Washington DC operated from 1999-2002. Compliance with the NBP cap begins in 2003 or 2004. We follow Deschenes et al. (2017) and ignore the OTC market which produced small declines in NO_x emissions. See Appendix Table 1 for details.

in regulated counties in Alabama and Michigan are assigned to the treatment group, while those in non-NBP counties are assigned to the control group.

We focus on emissions from fossil-fueled electricity generating units since they constitute 87% of total emissions from all regulated NBP units (EPA-430-R-07-009 September 2007), and high-frequency data on emissions from individual units are available. A key design feature of the NBP market is that the program's cap was binding only during May 1 to September 30 (henceforth, the ozone season) when NO_x's contribution to ozone formation is the highest. We exploit the seasonal nature of the market cap in the empirical analysis below.

In 2009 NBP was subsumed by the Clean Air Interstate Rule (CAIR) in the NBP regulated states, and expanded to additional states, while adding an annual emissions trading program for NO_x and SO_2 emissions. However CAIR was struck down by the D.C. Circuit Court, leading way to regulatory uncertainty surrounding its implementation and a dramatic fall in emissions allowance prices (U.S. Energy Information Administration, 2012; Schmalensee and Stavins, 2013). As a result, the market incentives embodied in CAIR have largely been considered non-binding. For this reason we focus our analysis on the NBP market alone in the period before CAIR was implemented.

2.2 Data Sources and Summary Statistics

Emissions, production, and electricity generating unit characteristics. We construct a monthly panel dataset of NO_x emissions (tons), SO_2 emissions (tons), and generation (MWh) at the electricity generating unit-level for the period 2000 to 2008 using the U.S. Environmental Protection Agency (U.S. EPA) Clean Air Markets Division continuous emissions monitoring systems data (henceforth, CAMD) (EPA, 2022a). We process the raw CAMD data and exclude units that report zero load across the entire NBP period. We also drop large outlier observations. We aggregate the hourly emission and generation data to the unit-month level for the empirical analysis.

The CAMD data are merged to U.S. EPA facility attribute data by power plant identifiers ("ORIS" code) and generating unit identifiers to associate units with their adopted NO_x control technologies (EPA, 2022b). The NO_x control technologies are string-based descriptions, and eleven distinct technologies were identified with descriptive statistics on total counts by NO_x technology choice shown in Figure 4. In practice, units adopt between zero to five control technologies. We then merge the CAMD data using power plant and generating unit identifiers with data from the U.S. Electricity Information Administration's Form-860 (U.S. Energy Information Administration, 2022), which identifies the location of generating units and their attributes, such as primary fuel, year of commercial operation (which we use to derive the age of each unit) and capacity. Importantly the Form-860 data contains information on each unit's smokestack characteristics such as the smokestack's height, diameter, area, and velocity and temperature of emissions, which are required inputs for the air pollution transport model.

After processing the data, our sample includes 1,125 unique power plants and 3,722 unique electricity generating units that use coal, natural gas, or oil as primary fuel over the period 2000–2008. As explained below, the empirical analysis is done at the unit-level, which are split roughly equally between states in the NBP market region (48% of units) and states outside of the NBP region (52% of units). Table 1 reports summary statistics on units among NBP and non-NBP states in the baseline years prior to the implementation of NBP (2000–2002). In both market regions, coal units contribute the majority of ozone season NO_x emissions (93% and 71%, respectively. The NBP market region counts more coal units (495 vs. 344), but coal units outside of the NBP market have larger average capacity (380 vs. 310 MW) and release higher quantities of NO_x per month during ozone season months, with coal units outside the NBP market having higher monthly emissions.

We identify whether each unit-year is located in a restructured de-regulated electricity market versus a non-restructured regulated market using data from the Electricity Information Administration's (EIA) Form-860, which lists regulatory status for power plants by

Table 1: **Baseline Characteristics of Fossil-Fueled Electricity Generating Units**

	Coal (1)	Natural Gas (2)	Oil (3)
NBP States:			
Power plants	181	186	99
Units	495	656	318
Unit Size (MW)	310	101	146
Ozone Season NO_x Emissions (tons/month)	335	5	19
Percent of Ozone Season NO_x Emissions	93.3	1.4	5.3
Non-NBP States:			
Power plants	160	348	24
Units	344	906	52
Unit Size (MW)	380	178	217
Ozone Season NO_x Emissions (tons/month)	424	39	132
Percent of Ozone Season NO_x Emissions	71.3	6.6	22.2

Notes: This table shows baseline summary statistics for the electricity generating units in our sample in the years before NBP is implemented (2000–2002). The first two rows in each panel are counts, the next two rows show means, and the last rows show percentages.

year.[10]

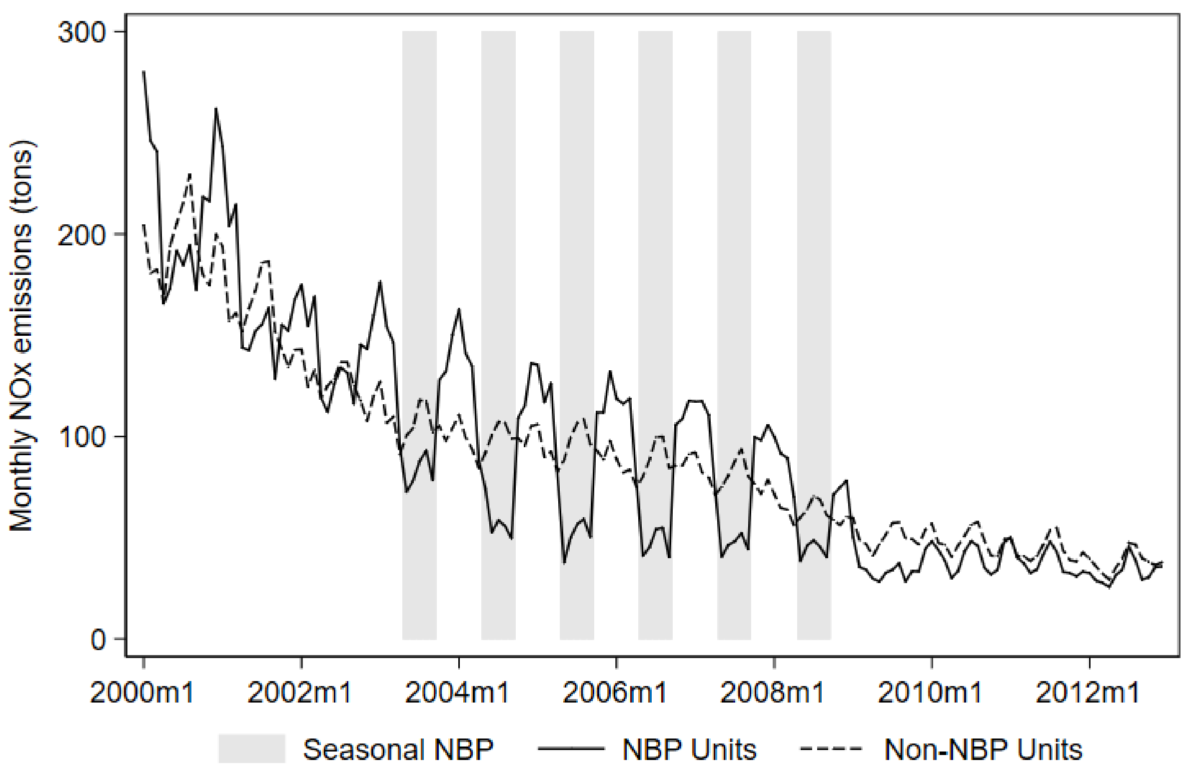
Figure 1 shows the average monthly NO_x emissions in tons at generating units by NBP status. We report these statistics over the period 2000–2012 in order to provide some context for the pre- and post-NBP periods.[11] There are several important patterns to report: first, there is a notable reduction in monthly NO_x emissions for units in the NBP market during the regulated ozone seasons during the NBP era (shown by the gray bar), compared to the months outside of the ozone season. This pattern is apparent across all units, and even sharper when focusing on coal units alone (see Appendix Figure 2). Second, emissions decreased in the

[10] Electricity market regulatory status is provided in the "Utility" tab of Form-860 for the years 2001, 2002, and 2003 and in the "Plant" tab of Form-860 for the years 2006, 2007, and 2008. Units within a power plant are assigned the regulatory status of their parent power plant. EIA Form-860 does not include regulatory status in the years 2004 and 2005—for these years, we carry forward the status of the unit as of the year 2003.

[11] See Holland et al. (2020) and Hernandez-Cortes et al. (2022) for an analysis of trends in power plant emissions over the 2010s.

years before the NBP, a trend which we will account for in the empirical analysis below. For example, average monthly emissions of NO_x dropped from 194 to 38 (non-NBP units) and 205 to 32 (NBP units) tons per month between 2000 and 2012. Finally, Appendix Figure 2 shows that NO_x emissions from natural gas and oil units are much smaller than from coal units and appear to only be weakly impacted by the regulation.

Figure 1: Average Monthly NO_x Emissions of Electricity Generating Units by NBP Regulatory Status



Notes: Figure 1 plots average monthly NO_x emissions by NBP and non-NBP regulatory status. The light gray areas denote the regulated months of the ozone season between 2003 and 2008. Units are assigned to NBP and non-NBP status based on their counties, as shown in Appendix Figure 1. Units in Missouri are excluded, as in the rest of the analysis.

3 Results

3.1 Impact of the NBP Market on Emissions

This section begins with the empirical methodology and results from estimating the effect of the NBP market on NO_x emissions using a triple-difference estimator. We then present an approach to predict counterfactual emissions from electricity generating units in absence of the market to derive unit-specific treatment effects. Finally, we use observable characteristics of each unit to identify the main predictors of treatment effect heterogeneity in the impact of NBP on NO_x emissions.

Triple-Difference Estimator. We use a triple-difference estimator (DDD) to estimate the effect of the NBP market on monthly unit-level NO_x and SO_2 emissions, pooling across all units, and also separately by primary fuel. This generalizes the approach in Deschenes et al. (2017), who estimate the effect of the NBP market on emissions using data at the county-season-year level. The estimating equation below relates unit-level emissions to a triple interaction term representing the treatment, D_{icmt}, which equals 1 for units located in counties where NBP was in effect, after the NBP began operating (2003 or 2004), and during the ozone season months, $m \in \{5, 6, 7, 8, 9\}$:

$$Y_{icmt} = \alpha + \beta D_{icmt} + \mu_i + \gamma_{ct} + \delta_{cm} + \eta_{mt} + \epsilon_{icmt}, \tag{1}$$

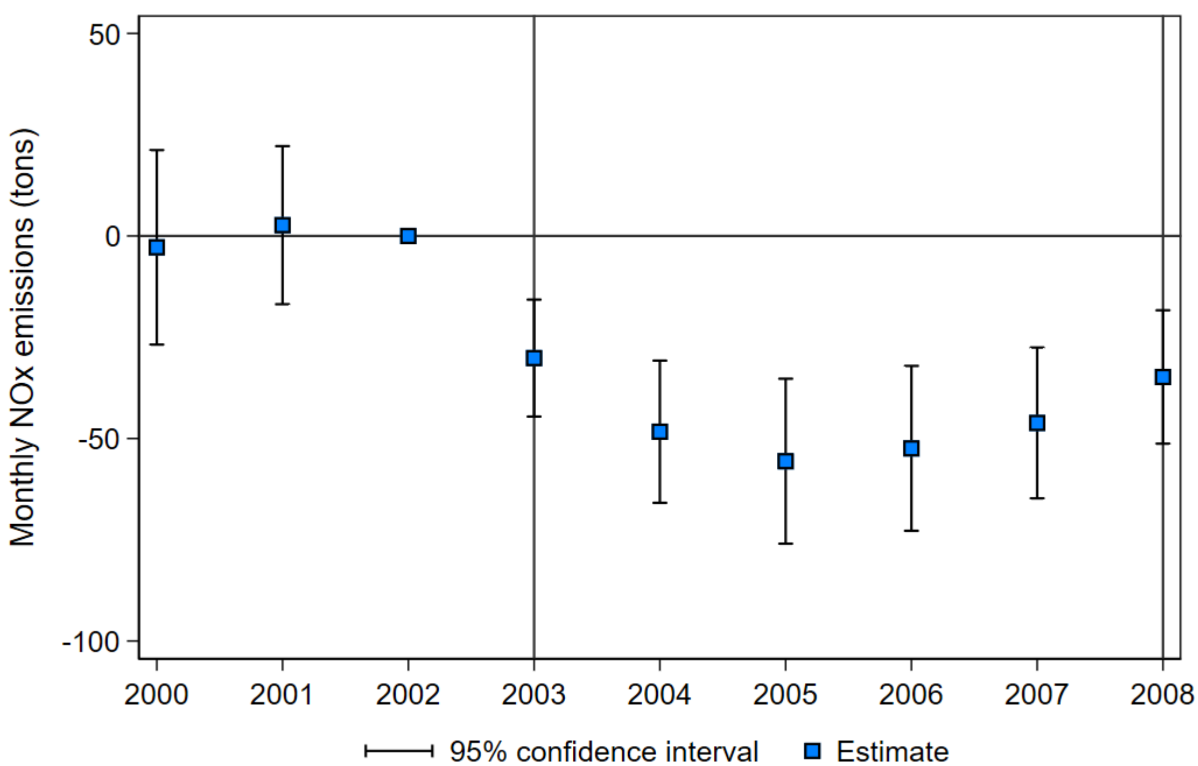
where Y_{icmt} represents unit-level monthly emissions unit i, located in county c, in month m and year t. Equation 1 includes unit fixed effects, μ_i, to control for time-invariant unit characteristics, such as fuel, age, and capacity. The seasonal nature of the NBP cap allows for the inclusion of county-by-year fixed effects (γ_{ct}) that control for local economic shocks at the county level that may affect emissions independently of the NBP regulation, as well as for electricity market regulatory incentives for abatement technology that vary by state. Seasonality and trends in emissions that vary across counties are absorbed by the county-by-month fixed effects, δ_{cm}, and month-by-year fixed effects, η_{mt}. The latter accounts for trending drivers of the outcomes that are common across all units but varying by year and month. Standard errors are two-way clustered at the power plant level and state-by-year level to account for potential serial correlation within power plants over time and common to all units in a state-year.

It is important to qualify the interpretation of the estimate of the parameter of interest, β. When treatment effects are constant across treated units and follow common trends, we can interpret $\hat{\beta}$ as the unbiased estimate of the average treatment effect (ATE) of the NBP market on unit-level monthly emissions. However, the recent econometric literature shows that when treatment effects are not constant, $\hat{\beta}$ may be a biased estimate of the ATE (see de Chaisemartin and D'Haultfoeuille 2023 for a review). Since the roll-out of the NBP occurred over 2 consecutive years, the issues associated with staggered adoption designs are of lesser concern in our setting. In addition, we present our approach to estimating heterogeneous treatment effects below.

There were other environmental regulations or markets in place in the early 2000s that could affect emissions of NO_x by electricity generating units regulated by NBP. In particular, the CAA ozone non-attainment designations, California's RECLAIM market, and New England's Ozone Transport Commission Market (which operates prior to the NBP) all can influence investments, production decisions, and emissions at power plants. Therefore $\hat{\beta}$ should be interpreted as the additional impact of the NBP market on unit emissions above and beyond the impact of the other markets and regulations. This interpretation requires the assumption that these other policies did not change differentially across the ozone and non-ozone seasons by NBP treatment status. We provide evidence supporting this assumption in the section below.

Event-study and triple difference estimates. We begin the empirical analysis by reporting event-study estimates of the difference in monthly NO_x emissions between units in the NBP market region (compared to those outside), in the ozone season months (compared to the rest of the year), separately for each year between 2000 and 2008, with the difference in 2002 normalized to zero. Figure 2 reports these estimates (shown by the blue squares), along with their 95% confidence intervals. The estimates in the pre-NBP year are close to zero (and statistically indistinguishable from zero), consistent with the hypothesis of no pre-NBP differences in relative trends in NO_x emissions. In contrast, the estimates for the years 2003–2008 show the impact of the NBP market on NO_x emissions for regulated units. In all those years, the estimates indicate fairly consistent reductions in NO_x emissions of about 50 tons per ozone season month per unit, with smaller impacts in 2003 (recall that several states enter the NBP regulation in 2004) and in 2008 as the uncertainty regarding CAIR was starting to emerge.

Figure 2: Event Study Estimates for Monthly NO_x Emissions, 2000–2008



Notes: This figure shows the estimates of the difference in monthly NO_x emissions by NBP regulatory status and ozone season status over the period 2000 and 2008, with the difference in 2002 normalized to zero.

Table 2: **Triple Difference Estimates of the Impact of NBP on Monthly NO_x and SO_2 Emissions**

	All Units (1)	All Units (2)	All Units (3)	Coal Units (4)	Gas Units (5)	Oil Units (6)
NO_x Emissions	-40.42^{***}	-50.28^{***}	-47.30^{***}	-108.94^{***}	4.22^{***}	8.78
	(7.87)	(9.19)	(8.86)	(16.37)	(1.13)	(6.49)
SO_2 Emissions	-3.75	-8.02	-7.10	-22.30^{*}	0.55	51.44
	(5.46)	(6.31)	(6.70)	(10.83)	(1.36)	(27.58)
Observations	309,459	259,061	367,478	98,108	178,176	33,097
Plant-Unit FE	Yes	Yes	Yes	Yes	Yes	Yes
County-year FE	Yes	Yes	Yes	Yes	Yes	Yes
County-month FE	Yes	Yes	Yes	Yes	Yes	Yes
Sample	2000-08	2000-08	1997-08	2000-08	2000-08	2000-08
Positive Dep. Variable	No	Yes	No	No	No	No

Notes: This table reports estimates of the impact of NBP on monthly emissions. Standard errors in parenthesis are two-way clustered at the power plant level and state-by-year levels. Statistical significance: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Table 2 reports estimates of β from equation 1 for monthly NO_x and SO_2 emissions. Emissions of co-pollutants such as SO_2 could have been impacted by NBP if, for example, NO_x abatement technologies also generate SO_2 abatement, or if the market led to a shift in electricity production from coal units to natural gas units. Columns (1) to (3) in Table 2 estimate the effect of NBP across all units, while columns (4) to (6) estimate the effect separately by input fuel. Overall we find that the NBP market led to sizable reductions in ozone season NO_x emissions but had small and statistically insignificant effect on SO_2 emissions. Comparing the estimates in columns (1) to (3) shows that specification choices in terms of fixed effects (at plant or unit level), sample period (2000–2008 versus 1997–2008), and dropping units with zero NO_x emissions have no meaningful impact on the estimates of the average effect of NBP on NO_x emissions. For the average unit, NBP caused a reduction in NO_x emissions of 40 to 50 tons per month, or roughly 28-36% of baseline average monthly emissions (all the estimates are precise with $p < 0.001$).

Columns (4) to (6) report fuel-specific estimates of the effect of NBP by fitting equation 1 separately by unit fuel. Our estimates indicate that NBP led to reductions in NO_x emissions largely by impacting coal units: for the average coal unit, the reduction is 108.9 tons per ozone season month. We estimate small *increases* in monthly NO_x emissions from natural gas units, and no detectable effect on NO_x emissions from oil units. Our estimates of the fuel-specific effects of NBP on SO_2 are not precisely estimated.

3.2 Treatment Effect Heterogeneity and Counterfactual NO_x Emissions

In the next step of the analysis we estimate the counterfactual emissions of each electricity generating unit regulated by NBP in absence of the NBP during the months of the ozone season. The difference between observed and counterfactual emissions provides a unit-specific estimate of the impact of the NBP market on emissions, which we interpret as the abatement specific to each unit and induced by the NBP market. Since a key feature of a market-based policy is that it allows for regulated units to heterogeneously change their emissions, we want to preserve this treatment effect heterogeneity in the construction of the counterfactual.

The recent econometric literature documents the issues that arise when estimating multi-way fixed effect models with dimensions of treatment heterogeneity that are pre-specified by the researcher, proposing machine learning-based approaches as alternatives (Wager and Athey, 2018; Chernozhukov et al., 2018). These approaches, however, can be computationally intensive to implement and difficult to replicate. Instead, we implement a new approach motivated by Verdier (2022) and following Borusyak et al. (2024), which allows for the estimation of treatment effect heterogeneity without pre-specifying the observable covariates by which the treatment responses vary. To proceed, let i denote unit, and t now denote a month-year.[12] The parallel trends assumption in our setting presumes the following about the evolution of monthly emissions Y_{it} for untreated units:

[12] We ignore the subscript for counties for simplicity.

$$E[Y_{it}(0) - Y_{it-1}(0) \mid D_i, \mathbf{x}_{it}] = E[Y_{it}(0) - Y_{it-1}(0) \mid \mathbf{x}_{it}] \tag{2}$$

where D_i is a vector of treatment history with elements $d_{it} \in 0, 1$ for unit i $\forall$ $t \in T$, and where T is the set of all month-years from 2000–2008; $\mathbf{x}_{it}$ is a vector of observable covariates (that can be time-varying). $Y_{it}(0)$ denotes the potential outcome of unit i at time t in the non-treated status (i.e., when NBP does not cap ozone season NO_x emissions), and $Y_{it}(1)$ denotes the potential outcome of unit i at time t under the treated status (i.e., when NBP caps ozone season NO_x emissions). Alternatively, we can state (2) following Borusyak et al. (2024) as:

$$E[Y_{it}(0)|\mathbf{x}_{it}] = \alpha_i + \lambda_t + \mathbf{x}_{it}'\delta_t \tag{3}$$

Under equation (3), we can estimate unit-level $Y_{it}(0)$ by estimating α_i, λ_t, and δ_t using non-treated observations. Importantly, we note that units treated by the NBP are also observed under the no-treatment status in two subsets of data: (1) in all months of the year before the NBP was in effect; (2) in months outside of the ozone season months during the period when NBP is in effect. Additionally, units located outside of the NBP region also contribute to identifying the counterfactual in all months and years. This information is critical in estimating the fixed effects underlying α_i, λ_t, and δ_t.

Estimating counterfactual emissions. We begin by estimating a regression of NO_x emissions Y_{it} on unit fixed effects, state-year fixed effects, state-month fixed effects, and covariates using only the *non-treated* sub-sample of observations. We then use these estimates to predict counterfactual emissions outcomes, $\hat{Y}_{it}(0)$ for units regulated by NBP, during the market period (2004–2008) and during the ozone season.[13] These predictions represent the counterfactual emissions for treated units absent the policy. We can then define the unit-

[13] While the policy started earlier in some states, we focus on counterfactual analysis on the years during which all 20 states and D.C. were regulated.

specific treatment effect (or unit-specific abatement) $\hat{\beta_{it}}$ as: $Y_{it}(1) - \hat{Y}_{it}(0)$.

We implement this approach in a setting where some units produce zero monthly NO_x emissions, especially gas and oil units. For example, units with zero emissions may be peakers that are rarely fired, units on scheduled maintenance, or units that are shutdown in response to the NBP regulation. To address the mass-point at zero in NO_x emissions we use a Poisson regression model, implemented as a quasi or pseudo maximum likelihood estimator (PPML) to estimate α_i, λ_t, and δ_t from equation (3).[14]

In practice, we specify the PPML model using a similar set of fixed effects as in the DDD analysis of the average impact of NBP in equation (1) (that is, unit fixed effects, state-by-year and state-by-month fixed effects, and month-by-year fixed effects). Combined with these fixed effects, we include observable predictors that best represent the process for emissions under the no-NBP counterfactual at each unit. These include load in the NERC region (varying monthly), and temperature and precipitation in the county where the unit is located (varying monthly). We also control for electricity market regulatory status with an indicator for units operating in regulated (non-restructured) electricity markets as of the year 2003.[15] Importantly, as we show in the Appendix Figure 3, the choice between PPML and OLS—where OLS aligns most directly with the parallel-trends assumption in equations 2 and 3—has no material effect on the counterfactual emissions estimates.[16] For example, a regression of PPML-predicted on OLS-predicted counterfactual emissions estimates has an slope of 1 and an R-squared of 0.98. See Appendix Figure 3.

We estimate the parameters of equation (3) using the *ppmlhdfe* routine in Stata, applied to non-treated observations (units outside of the NBP market, units inside the NBP market

[14] See e.g., Gourieroux, Monfort, and Trognon (1984), Santos Silva and Silvana (2008), and Wooldridge (2010)

[15] The EIA Form-860 data on restructuring status does provide annual data on electricity market status for the years 2001–2008, excluding the year 2004 and 2005. Since in our empirical strategy we do not attempt to identify time-varying treatment effects, we include only the status as of the year 2003, and thus our identified effect should be interpreted with respect to electricity regulatory status at that time.

[16] The unit fixed effects control for many attributes that predict emissions such as unit fuel, age, and capacity. The cross-validation exercise described below justifies our choice of state-by-year and state-by-month fixed effects instead of county-by-year and county-by-month.

region in years before the market is operating, and units inside the NBP market region, but during months outside of the ozone season). We then use these estimates, including the estimated fixed effects parameters, to predict monthly emissions of treated units under the counterfactual case where NBP is not implemented, but where all other regulations and drivers of emissions remain the same. The PPML estimator ensures that all predicted values for counterfactual emissions are non-negative, and we also verified that the largest values for predicted monthly emissions are within the range of observed values in the data.

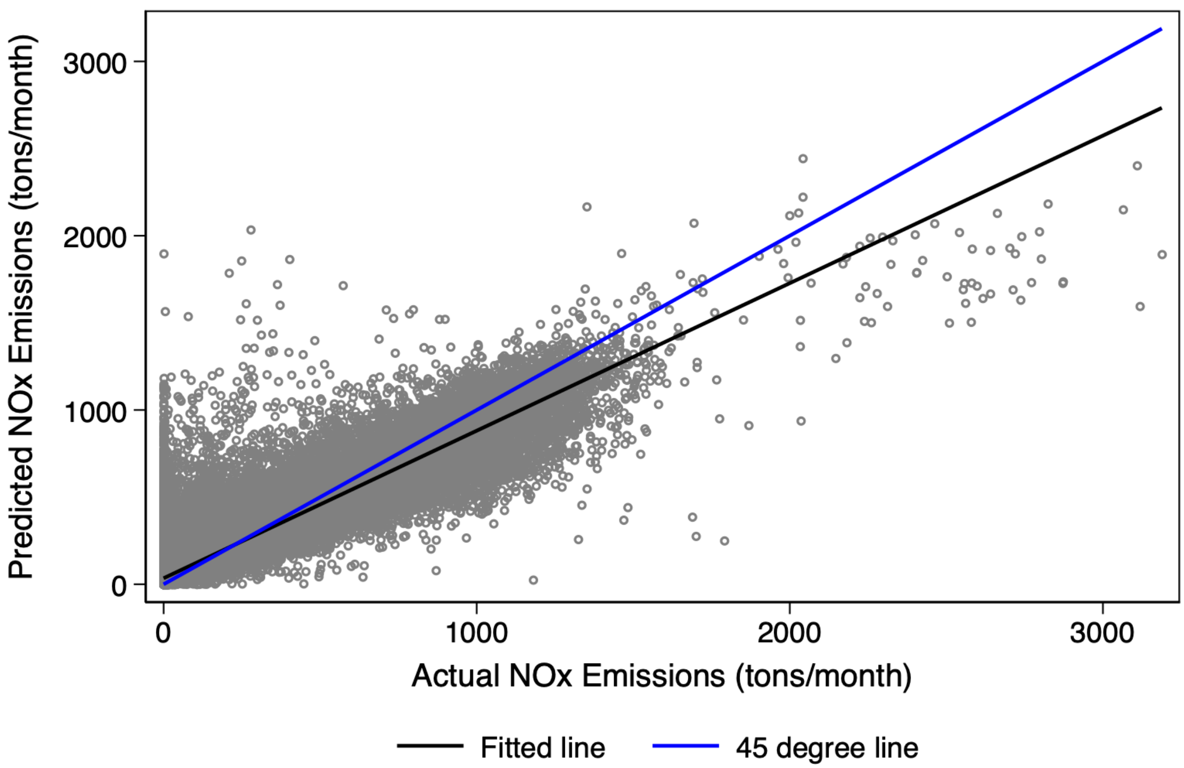
Figure 3 summarizes the predictive accuracy of our approach using a five-fold cross-validation. Specifically, the predicted monthly NO_x emissions at all untreated units over the period 2000–2008 (scaled on the y-axis) are plotted against the observed values for the same units, in the same months. The predictions are obtained by reserving a sample of 80% of observations for untreated units for a given fuel, estimating the parameters of equation (3) using PPML separately by fuel, and then predicting the level of the dependent variable out-of-sample over the remaining 20% of observations. We repeat this exercise 5 times so that each actual observation receives 1 predicted value. As the figure highlights, the data points shown in gray are mostly clustered around the 45-degree line (shown in magenta). The fitted regression line across the five cross-validation samples (shown in black) is slightly below the 45-degree line, indicative of a small degree of over-prediction of NO_x emissions.[17] The cross-validated R-squared is 0.86.[18]

Figure 4 continues the analysis of our estimated counterfactual emissions by reporting average monthly emissions for units operating outside of the NBP region, panel (a), and operating inside the NBP region, panel (b). For this analysis, we re-estimate the predictive PPML model in (3) separately by fuel using the sample of untreated observations, and predict emissions for both treated and untreated units. More specifically, Figure 4 reports the monthly average of actual (observed) emissions and counterfactual emissions for the

[17] The average prediction error is 0.4 ton per month, which represents 0.34% of the average observed monthly NO_x emissions (117.2 tons).

[18] We note that out of a sample of 260,423 non-treated observations, 2,070 observations have a missing predicted value. These are small units, that represent 0.2% of NO_x emissions among untreated units.

Figure 3: Accuracy of Predictive PPML Model for Counterfactual Emissions, 2000–2008



Notes: This figure illustrates the predictive accuracy of the PPML model for counterfactual emissions in equation (3) for all untreated units over the period 2000–2008. Predicted counterfactual monthly NO_x emissions are on the y-axis and observed monthly NO_x emissions are on the x-axis. The predictions are performed using out-of-sample five-fold cross-validation so that each observation receives 1 predicted value. The 45-degree line is shown in magenta and the fitted regression line is shown in black. The cross-validated R-squared is 0.86.

untreated units in Panel (a) and treated units in Panel (b). The full lines (gray) show the average predictions and the dashed lines (blue) show the actual observations.

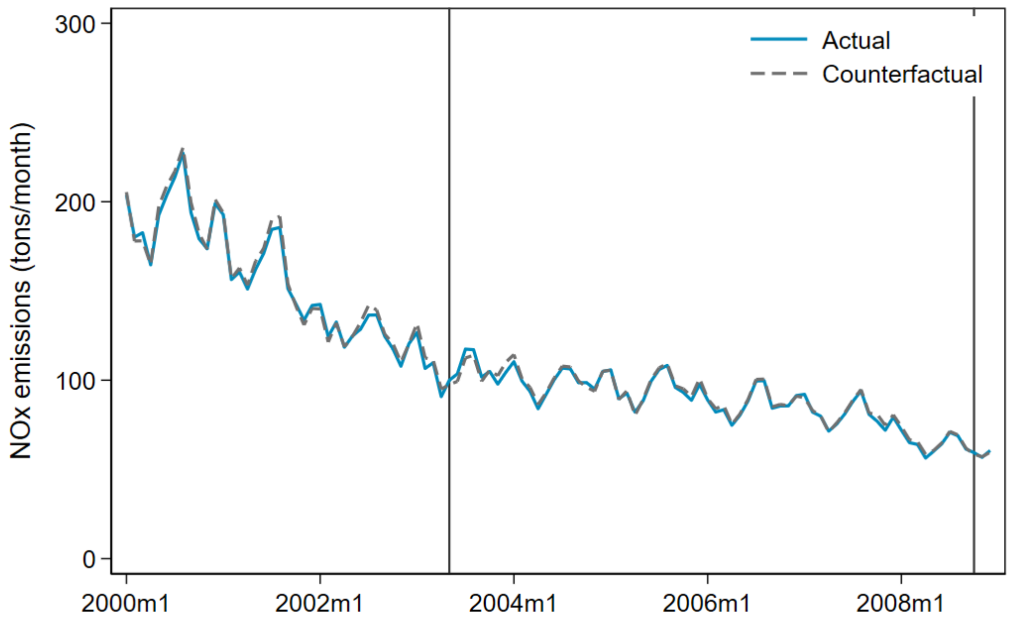
Panel (a) in Figure 4 shows that our approach for predicting counterfactual NO_x emissions accurately replicates (average) actual observed emissions for non-regulated units. Importantly, the predicted emissions track actual emissions for unregulated units equally well in the pre-NBP and during the NBP periods, as well as during the ozone season and the rest of the year. Overall the errors shown by discrepancies between the full and dashed lines appear random and evenly distributed across 2000–2008.

Panel (b) reports a similar degree of accuracy for the average predicted emissions for NBP units in the pre-NBP period when NO_x emissions were uncapped. Further, emissions are uncapped during the non-ozone season in the NBP period, and predicted emissions are remarkably similar to observed emissions during those months. Finally, Panel (b) provides a visual representation of the impact of NBP on NO_x emissions. During the ozone season months, from 2003 (especially starting in 2004) to 2008, observed emissions are notably smaller than predicted counterfactual emissions, especially between 2004–2007. The average treatment effect, which can be computed by comparing observed and counterfactual emissions among NBP units, during the ozone season, between 2003 and 2008 is -32.3, slightly smaller than the DDD estimates reported in Table 2.

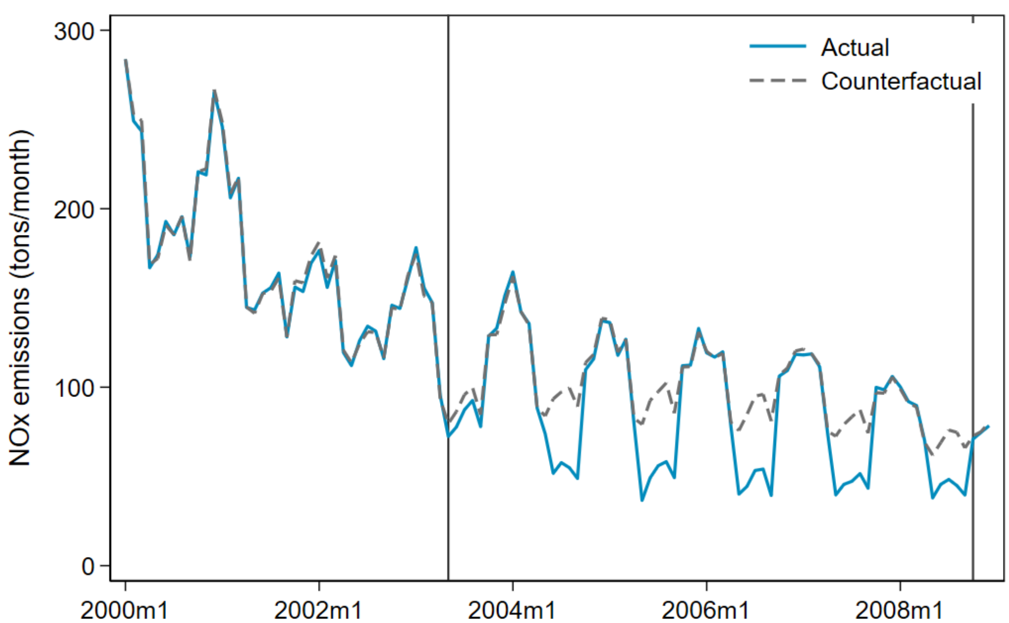
Heterogeneous treatment effects. As discussed above, the mean difference between predicted counterfactual and actual emissions over all N regulated units, $\frac{1}{N}\sum_i^N Y_{it}(1) - \hat{Y_{it}}(0)$ during the regulated months, corresponds to the average treatment effect of the NBP on monthly emissions during the ozone season among regulated units. We now analyze the distributions of the estimated unit-specific treatment effects (or unit-specific abatement). Figure 5 plots the average of these unit-specific effects by NBP regulatory status and fuel. The top panel shows the distribution of estimated treatment effects for coal units in the NBP market region; the average is -116.5 tons of NO_x per month, similar to the estimates

Figure 4: Actual and Counterfactual Emissions by NBP Status, All Units

(a) No NBP Regulation



(b) NBP Regulation



Notes: The solid blue lines show the average observed monthly NO_x emissions for untreated units (panel a) and for units in the NBP market region (panel b). The dashed gray lines show the average predicted counterfactual NO_x emissions based on the PPML model in equation (3). The vertical line correspond to the beginning of the NBP market regulatory period in May 2003.

of -108.9 in Table 2. The NBP impact estimates for natural gas and oil are smaller in magnitude in line with the results in Table 2, with average treatment effects in NBP months and states equal to -2.6 and -2.4 tons per month, respectively. These findings are consistent with policy literature that suggests that coal units were most impacted by the NBP (e.g., U.S. EPA (2011)).

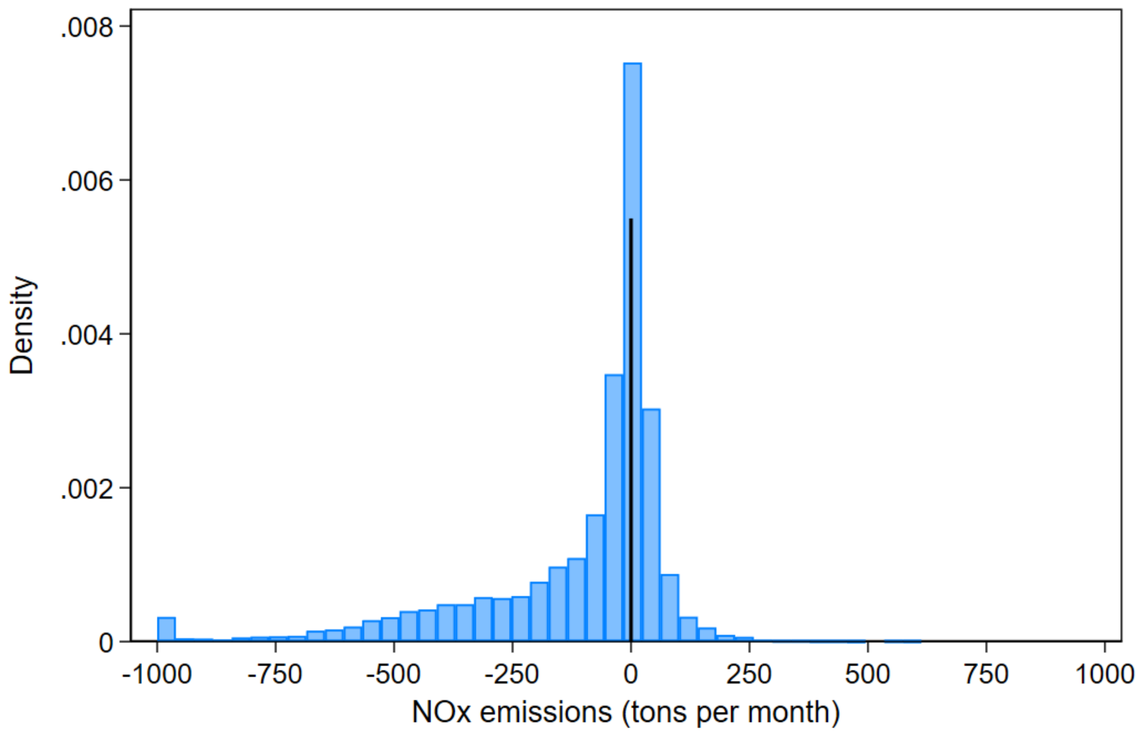
While the average of the unit-specific estimates of the impact of NBP on NO_x emissions are similar to the DDD estimates in Table 2, the most striking feature of the estimates in Figure 5 is the large degree of heterogeneity in the impact of NBP among regulated units, even within fuel source. In particular, the standard deviation in the estimated effect of NBP across coal units is 244.1, almost twice the magnitude of the average treatment effect; the 5th to 95th percentile range in estimated impact of NBP is -596.5 to 66.4, expressed in tons per month in the ozone season. Likewise, the standard deviation across gas and oil units are 14.1 and 28.9, several times larger than the average effect of the regulation on emissions from those units. This important degree of heterogeneity in the response to the regulation across units, ignored by the standard DDD estimator, will allow us to precisely characterize the impact of NBP on ambient $PM_{2.5}$ concentrations.

We next explore the drivers of the treatment effect heterogeneity documented in Figure 5 by regressing the unit-specific treatment effects, $Y_{it}(1) - \hat{Y}_{it}(0)$, on observable unit covariates and year fixed effects among the sub-sample of treated units, thus only including units in NBP regulated states in regulated years during the ozone season. These estimates indicate how much of the variation in treatment effect differences across units is explained by the respective unit characteristics. The results are reported in Figure 6, which shows the estimated regression coefficients and 95% confidence intervals for a specification that only includes regulated coal units, the set of units most impacted by NBP. The covariates include electricity market regulatory status and binary indicators for the each of the different categories of NO_x control technology.[19] In addition, we include controls for unit age (years) and

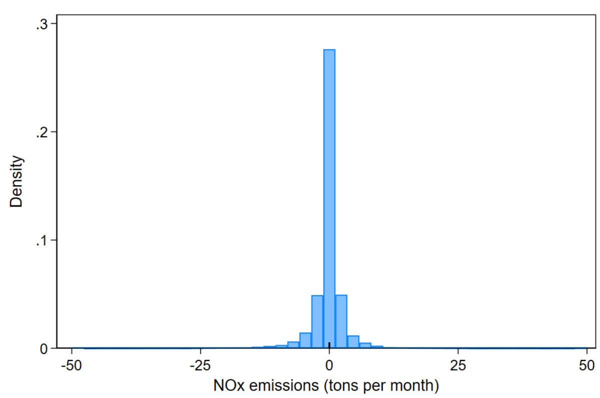
[19] Some units have multiple control technologies (up to five), and some technologies are only used for specific fuels. This specification assumes that the impact of control technologies on the estimated treatment effect

Figure 5: Distribution of Abatement Effects by Fuel

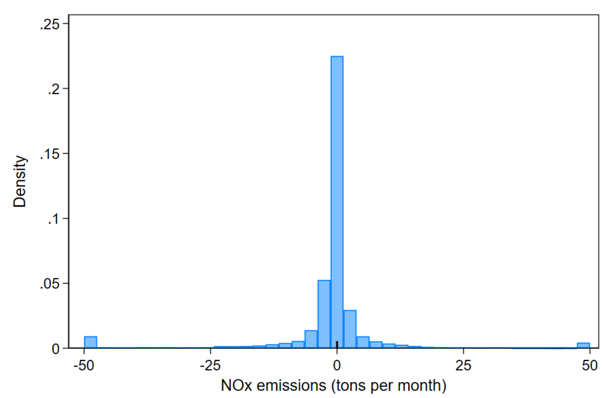
(a) Coal



(b) Natural Gas



(c) Oil



Notes: These histograms show the distribution of estimated unit-level abatement in tons per month of the ozone season (or unit-level treatment effect) by fuel type. This is calculated as the average of $Y_{it}(1) - \hat{Y}_{it}(0)$ over all NBP-regulated units over ozone season months between 2004-2008.

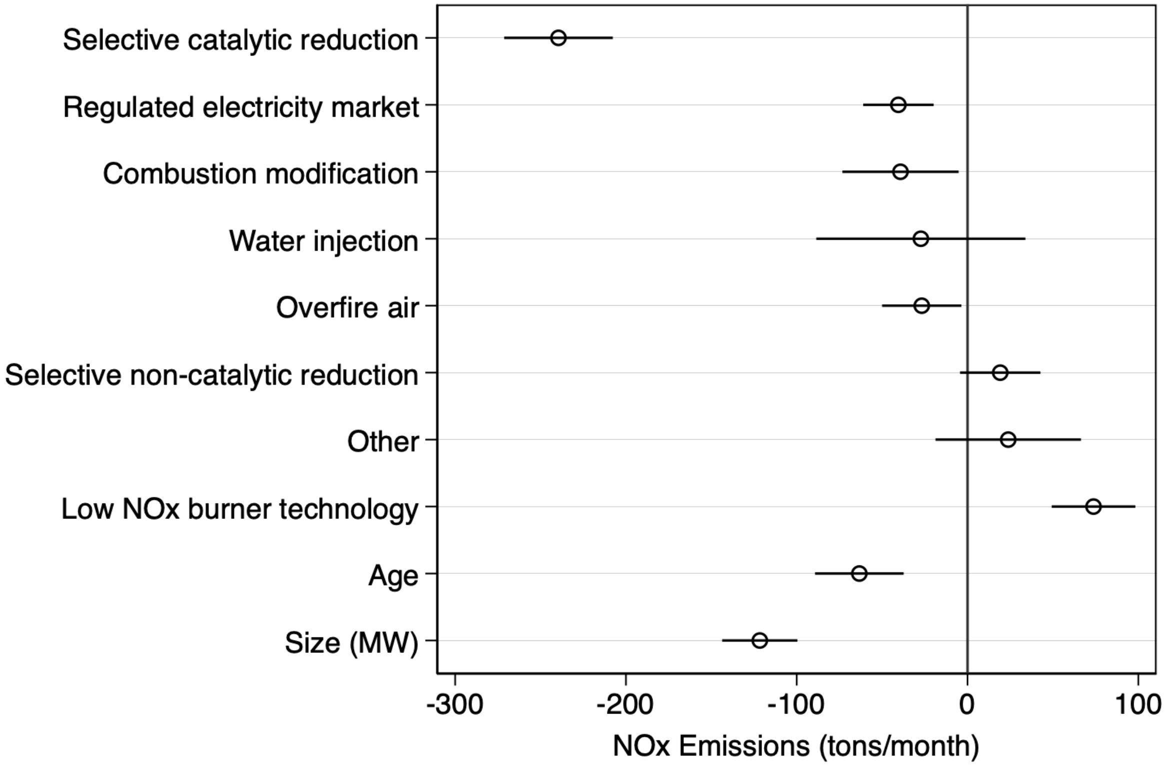
size (MW).[20]

The estimates in Figure 6 show units with selective catalytic reduction processes generate the largest treatment effects, leading to an additional 230 tons of NO_x emission abatement per month, after controlling for unit age, size, electricity market regulatory status, and other NO_x control technologies. By comparison, the average estimated treatment effect for coal units overall has a mean of -111.8 tons per month during the ozone season. The heterogeneous treatment effects by regulated electricity market status persist even after controlling for NO_x control technology choices, indicating that not only does electricity market regulatory status impact NO_x control technology choices as found in Fowlie (2010), but also potentially the intensive margin uses of these technologies after adoption, or another margin of emissions reduction. Finally, older and larger coal units also produce larger NO_x abatement on average, holding other predictors constant.

is additive.

[20] We standardize the estimated coefficients for unit age and size to facilitate comparison.

Figure 6: Predictors of Treatment Effect Heterogeneity for Coal Units



Notes: This figure shows the estimated coefficients and 95% confidence intervals from a regression of the estimated unit-level treatment effects (abatement) on observable covariates among treated coal units during the treated ozone season months. The continuous predictors (unit age and size) are standardized for better comparability. Standard errors are clustered by unit.

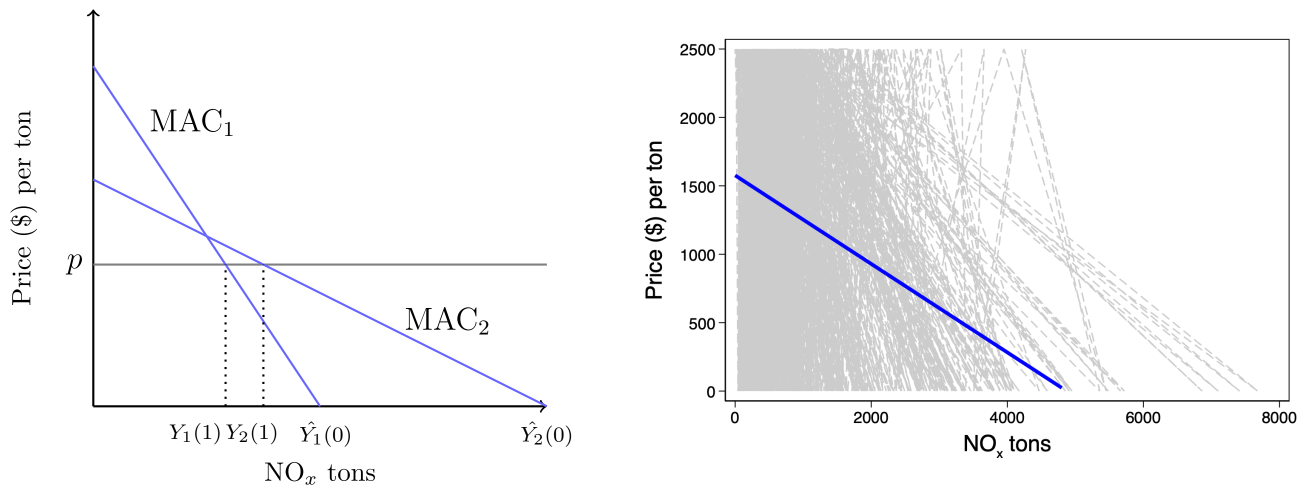
3.3 Marginal and Total Abatement Costs

Abatement costs. We use observed unit-level NO_x emissions under the regulation, $Y_{it}(1)$, and estimated of emissions absent the regulation, $\hat{Y}_{it}(0)$, together with the average observed NO_x emissions permit price to construct each unit's implied marginal abatement cost curve.[21] The construction follows from a simple model of competitive firm profit-maximizing behavior when regulated by an environmental market, shown in the left-panel of Figure 7. The left panel shows two representative units with observed monthly average regulated ozone season emissions, $Y_1(1)$ and $Y_2(1)$. Section 3.2 shows how we estimate counterfactual ozone season

[21] We follow Deschenes et al. (2017) and use an average permit price of $2,523 ($2015).

emissions in absence of the NBP market, $\hat{Y}_1(0)$ and $\hat{Y}_2(0)$, which correspond on the graph to a setting with a price of NO_x emissions equal to zero. We estimate each source's marginal abatement cost curve by computing the slope between the two points $(\hat{Y}_i(0), 0)$ and $(Y_i(1), P)$. The right panel of Figure 7 shows the resulting estimated MAC curves for all coal units in the gray dashed lines, with an average estimated MAC in blue. Appendix Figure 6 shows the estimate MAC curves for natural gas and oil units.[22]

Figure 7: Estimated Firm-Specific Marginal Abatement Costs: Illustrative Example (Left) and Empirical Estimates (Right)



Notes: The left panel shows the construction of unit-specific MAC curves from observed regulated emissions, $Y_i(1)$, and estimated no-market emissions, $\hat{Y}_i(0)$. The right panel shows all estimated MAC curves for regulated coal units–gray dashed lines, and the blue line represents the average estimated coal-unit MAC curve.

We use the estimated MAC curves to construct unit-specific total abatement costs in an average regulated ozone season—the triangle formed by $(\hat{Y}_i(0), 0), (Y_i(1), P), (Y_i(1), 0)$—and find aggregate abatement costs for an average ozone season by summing over all regulated units. We estimate total abatement costs from regulated units of $413 million per ozone season, with median unit-specific costs of $466,000 ($2015).[23] Prior estimates of the monetized benefits from the NBP—$800 million from reduced pharmaceutical purchases and $1.3

[22] Meng and Thivierge (2025) point out that due to transaction costs, the price paid for a permit may fall above a regulated source's true MAC at that emissions level. Our approach implicitly assumes negligible transaction costs.

[23] Deschenes et al. (2017) use the average permit price to calculate an upper bound of total abatement costs in the market of $(US 2015) 1,076 million.

billion from avoided mortality (Deschenes et al. (2017))—total \$2.1 billion (\$2015). Comparing these 2 monetized benefits alone with our estimated abatement costs implies that the benefits of the NBP are roughly five times its costs.

Aggregate abatement costs under uniform standard. We next turn our analysis to comparing the aggregate abatement costs under the NBP market to the aggregate abatement costs under a counterfactual policy that regulates NO_x emissions with a command-and-control style approach. Specifically we consider abatement when all regulated coal units reduce emissions by a uniform share, where the percentage is chosen such that the total annual NO_x emissions abatement among regulated units is the same as in the NBP.[24] We focus on coal units only when analyzing this counterfactual policy given our earlier analysis found that coal units were responsible for 98% of the total abatement induced by the NBP market.

We calculate the unit-specific abatement costs under this 'uniform reduction' policy using the previously estimated MAC curves, and the new abatement cost triangle for the uniform reduction policy (U) formed by $(\hat{Y}_i(0), 0), (Y_i(U), P), (Y_i(U), 0)$, where $Y_i(U)$ is calculated as: emissions absent the market, $\hat{Y}_i(0)$, less the required emissions reduction under the uniform reduction policy. We find that the total estimated abatement costs under the uniform reduction policy are around \$2.5 billion, which is more than 6 times the corresponding estimate under the NBP market, as shown in Table 6. This finding demonstrates the value of flexible, market-based approaches that allow polluters to respond to heterogeneous abatement opportunities and minimize aggregate abatement costs.

[24] We identify the share, τ, by solving equation: $R = \sum_{i=1}^{N}\{\hat{Y}_i(0)(1-\tau)\}$, where R is the total NO_x emissions abatement in an average regulated summer ozone season and $\hat{Y}_i(0)$ is the estimated no-NBP average summer ozone season emissions from coal unit $i = 1, ..., N$.

Table 3: Total Abatement Costs in the NBP Market and Counterfactual Uniform Reduction Policy

	NBP Market	Uniform Reduction
Total abatement costs	412,791,875	2,536,124,735
Median unit costs	466,493	638,925

Notes: The top row of this table reports our estimates of total abatement costs under the observed market-based NO_x Budget Program as well as under a counterfactual policy that achieves the same aggregate NO_x emissions reductions by requiring each coal generating unit to reduce its emissions by an equal share. For the uniform reduction costs estimates units with positive estimated MAC curves are assigned the median estimated MAC curve over all regulated coal units.

4 Distribution of Changes in Ambient Air Pollution due to the NBP

Next we evaluate the *distribution* of benefits from improved air quality induced by the policy, which requires a source-specific approach to estimating policy impacts. In particular, this approach must identify the marginal benefits of avoided NO_x attributable to each regulated unit and characterize the populations exposed to the pollution resulting from each unit's NO_x emissions. As explained in the introduction, these benefits and their distribution will be determined by three primary factors: (1) the spatial distribution of regulated sources; (2), the abatement induced by the market at each regulated source; and, (3) the marginal damages of pollution released by each regulated source. Factor (1) is observed from the locations of the regulated units, and we estimated (2) in the preceding section. Here we discuss our approach to estimating factor (3) and describe the results.

Theoretical predictions for distributional outcomes. Before proceeding to the estimation, we review the predictions from a simple model of environmental markets. The left panel of Figure 7 illustrates predicted abatement from regulated firms where emissions are regulated with an environmental market. In this example, we see that firm 2 abates more than firm 1: $Y_2(0) - Y_2(1) > Y_1(0) - Y_1(1)$. The difference in abatement reflects firm 2's flatter MAC curve—smaller slope in absolute magnitude—compared to firm 1. Extending that logic, one can characterize relative abatement quantities across regulated firms with a ranking of their MAC curve slopes: a ranking of firms by MAC curve slope from flattest to steepest would provide a ranking of predicted market-induced abatement from largest to smallest. Moreover, populations downwind of firms with flatter MAC curves are likely to gain more from an emissions market, since those firms are predicted to abate more, leading to less downwind ambient air pollution. In contrast, populations downwind of firms with steeper MAC curves are expected to experience smaller air quality benefits. This intuition follows from canonical environmental markets theory in Baumol and Oates (1988) and is discussed in more recent empirical work in Hernandez-Cortes and Meng (2023). Our paper is the first to pair this theoretical prediction with empirical estimates of source-specific MAC curves.

Estimating the distribution of ambient $PM_{2.5}$. Next, we estimate unit-specific marginal benefits of avoided NO_x emissions. The benefits of avoided NO_x stem from the two primary air quality impacts of NO_x: its impact on ambient $PM_{2.5}$ concentrations and its impact on ozone formation. To estimate the impact of a given unit's NO_x emissions on spatially-explicit $PM_{2.5}$ concentrations, we make use of a leading air pollution transport model that uses a reduced-complexity approach to estimate a source-receptor matrix for NO_x emissions and $PM_{2.5}$. NO_x contributions to ambient ozone formation is a much more complex chemical process than its contribution to PM concentrations, and a comparable reduced-complexity model does not exist. Full complexity models of NO_x impact on ozone formation are highly nonlinear and sensitive to environmental factors, in particular, the presence of volatile organic compounds (VOCs) (Liu and Shi, 2021). These models are prohibitively computationally and time intensive for the purposes and geographic granularity required for this project.[25] Thus, we proceed with the distributional analysis focusing on the

[25] Examples of these models include: CMAQ (Community Multiscale Air Quality), CAMx (Comprehensive Air Quality Modeling with Extensions), and WRF-Chem (WRF-Chem is the Weather Research and Forecasting model coupled with Chemistry).

marginal benefits from avoided $PM_{2.5}$ alone.[26]

To identify the contribution of a NBP-regulated unit's NO_x emissions to downwind $PM_{2.5}$ concentrations across the contiguous United States, we make use of the Intervention Model for Air Pollution (InMAP), a geographically fine-scale atmospheric transport model for air pollution (Tessum et al., 2017). InMAP was created using spatially explicit, annual-average physical and chemical information from a frontier chemical transport model, combined with simplifying assumptions regarding atmospheric chemistry to estimate the air pollution impacts of marginal changes in emissions.[27] As a result, InMAP provides elements of a source receptor matrix, henceforth ISRM, which quantifies how primary emissions of NO_x from a given unit contribute to ambient concentrations of $PM_{2.5}$ at receptor locations. The ISRM allows us to map unit-level NO_x emissions changes induced by the NBP to resulting changes in ambient $PM_{2.5}$ concentrations. We use this to derive avoided mortality from NBP-induced NO_x emissions changes, as well as to estimate source-specific marginal damages from NO_x emissions contribution to $PM_{2.5}$ concentrations.

The inputs we provide to the ISRM are emissions from electricity-generating units and unit-level smokestack characteristics from EIA Form-860. It then computes ambient $PM_{2.5}$ concentrations attributable to each unit's NO_x emissions across gridded locations in the contiguous United States.[28] More specifically, the ISRM is composed of elements ω_{ij} of a source receptor-matrix, which quantifies how primary emissions of NO_x from unit i contribute to concentrations c_j of $PM_{2.5}$ in location j. We use the ISRM to calculate the contributions to ambient concentrations of $PM_{2.5}$ in location j from all NBP-regulated units as:

[26] In the analysis below, we compare the ordering of units by their marginal damages against an ordering based on their abatement quantity. To the extent that the spatial distribution of ozone formation from a given unit's NO_x emissions follows a comparable pattern as the spatial distribution of $PM_{2.5}$ due to that unit's NO_x emissions, our distributional analysis will be informative about the combined ozone and $PM_{2.5}$ effects of the NBP regulation. However, to the extent that a unit's NO_x emissions generate different spatial patterns of ozone compared to $PM_{2.5}$ concentrations, our distributional analysis applies only to the marginal benefits from avoided $PM_{2.5}$ induced by the NBP.

[27] InMAP was created using annual average atmospheric conditions as of the year 2005 (Goodkind et al., 2019), which is in the middle of our study period, 2003–2008.

[28] The geographic grid in InMAP is as small as one by one kilometer in densely populated areas, and as large as 48 by 48 kilometers in rural areas Gallagher et al. (2023).

$$c_j(0) = \sum_i \omega_{ij} \hat{Y}_i(0), \qquad c_j(1) = \sum_i \omega_{ij} Y_i(1) \tag{4}$$

where $\hat{Y}_i(0)$ denotes the counterfactual NO_x emissions from unit i in absence of the NBP market and $Y_i(1)$ denotes the observed NO_x emission of unit i under the NBP market. As explained earlier, for each regulated unit, we focus on the average counterfactual and observed NO_x emissions during the average ozone season over 2004–2008. Accordingly, $c_j(0)$ denotes the contribution to ambient $PM_{2.5}$ at location j of all units under the counterfactual no-NBP scenario, whereas $c_j(1)$ denotes the contribution of all units under the observed NBP scenario. The difference between $c_j(0)$ and $c_j(1)$ can be interpreted as the impact of the NBP market on ambient concentrations of $PM_{2.5}$ at location j.

Locations in the ISRM are defined by a grid covering the contiguous U.S. We link the ISRM grid with population and demographics using the 2000 U.S. Census of Population. We layer the ISRM's gridded output onto the 2000 census tract polygons and compute the area-weighted average $PM_{2.5}$ concentration based on grid cells within or intersecting census tracts. This generates NBP-regulated units' census-tract level contributions to ambient $PM_{2.5}$ concentrations under the observed (NBP) and counterfactual (no-NBP) scenarios. Thus, j effectively denotes a census tract, noting this additional computational step to project the ISRM grid results onto census tracts. We assemble census tract-level demographic characteristics from the 2000 Census using the IPUMS-NHGIS database: specifically, population counts by tract for Black, Hispanic, and White populations, and the number of people below the poverty line. We focus on the distribution of population in 2000 since it cannot reflect the possible impact of the NBP market on population mobility.

We use our estimates of $c_j(1)$ and $c_j(0)$ to calculate predicted deaths under the observed (NBP) and counterfactual (no-NBP) ambient $PM_{2.5}$ concentrations. We use the standard "health impact function" approach that maps a *change* in ambient air pollution into predicted change in deaths based on estimates of baseline mortality rate, population, and estimated concentration-mortality function for $PM_{2.5}$:[29]

$$\Delta Deaths_j^d = Population_j^d \times MR_0^d \times [exp(\gamma \times (c_j(1) - c_j(0)) - 1] \tag{5}$$

where $d \in \{Black, White, Hispanic, Asian, Below\ Poverty\}$ and γ is the concentration-mortality parameter from Krewski et al. (2009). We use demographic group (d)-specific baseline mortality rates, MR_0^d, in 2000 from CDC Wonder. We assume that γ applies to all age groups and all demographic groups.

The results are presented in Table 4. Column (1) shows predicted mortality by demographic group from NO_x emissions contributions to ambient $PM_{2.5}$ in our counterfactual scenario without the NBP market for an average regulated ozone season. Column (2) presents the same metric with the NO_x emissions observed under the NBP market, and column (3) shows the difference (column (1) minus (2)), which provides an estimate of the NBP market's impact on predicted mortality. Column (4) reports avoided mortality from the NBP as a rate, per 100,000 population of each respective demographic group, to account for population size differences. Finally, column (5) shows avoided deaths by demographic group under the counterfactual uniform policy that achieves the same aggregate amount of ozone season NO_x emissions abatement as the NBP market but where each coal unit reduces emissions by an equal share.

Three key insights emerge from this table, focusing on the impact of NBP on avoided deaths per 100,000–the measure of benefits that normalizes for population differences across demographic groups. We find that the distribution of benefits across race, ethnicity, and povery status range from 0.109 to 0.001, with the White population seeing the largest benefits and Hispanic and Asian populations seeing the smallest. Second, the total amount of avoided mortality from NO_x contribution to $PM_{2.5}$ in the observed NBP market and simulated uniform policy is very similar: 303 and 305 avoided deaths per ozone season, respec-

[29] We use the same formula to compute predicted mortality associated alternaitve distributions of ambient $PM_{2.5}$ concentrations in the analysis below (i.e, the $c_j()$ that results from our counterfactual uniform reduction scenario).

tively. This outcome is perhaps particularly notable given the large difference in abatement costs across the policies found in Table 3, with the uniform policy costing 6 times as much as the NBP market. Third, the distribution of benefits from the NBP market and the uniform policy in terms of avoided mortality across race, ethnicity, and poverty status is also very similar, which is observed by comparing columns (3) to (5). Based on these results, we find that overall the NBP and uniform policy yield similar benefits, in terms of both total avoided mortality and the distribution of avoided mortality across demographic groups.

Table 4: Avoided Mortality Benefits by Demographic from the NBP Market and Counterfactual Uniform Policy

	(1) Predicted Deaths No NBP	(2) Predicted Deaths NBP	(3) Δ Pred. Deaths (NBP–no-NBP)	(4) Δ Pred. Deaths per 100,000 (NBP–no-NBP)	(5) Δ Pred. Deaths (Uniform–no-NBP)
All	704	401	-303	-0.109	-305
White	605	342	-263	-0.093	-263
Black	83	48	-35	-0.014	-36
Hispanic	11	7	-4	-0.001	-5
Asian	5	3	-2	-0.001	-2
Below poverty line	77	44	-33	-0.013	-33

Notes: Column (1) shows predicted mortality by demographic group from NO_x emissions contribution to ambient $PM_{2.5}$ without the NBP market during an average regulated ozone season. Column (2) shows the same metric with the NBP, and column (3) reports the impact of the NBP on $PM_{2.5}$ -related mortality, column (1) less (2). Column (4) shows the impact of NBP on mortality as a rate per 100,000 people of the respective demographic group. Column (5) reports the impact of the counterfactual uniform reduction policy (each coal unit reduces its emissions by an equal share with aggregate abatement the same as NBP) on predicted $PM_{2.5}$ -related deaths per 100,000.

Mechanisms driving the distribution of air quality benefits. Next, we explore the underlying mechanisms generating the results in Table 4. To do so, we use the ISRM in a similar manner as used for Table 4, but now focus on identifying the marginal contributions of individual regulated unit to ambient $PM_{2.5}$ and the resulting predicted mortality effects. To estimate these marginal impacts, we use the ISRM to compute the $PM_{2.5}$ concentrations that result from a series of NO_x "pulse" simulations. Each pulse simulation includes a discrete increase in NO_x emissions from one individual unit. Sequentially for each unit, we set its NO_x emissions equal to the aggregate NO_x emitted during an average regulated ozone season in the NBP market.[30] Then, as done in the earlier analysis, we map receptor locations from the ISRM to the census tract level to incorporate population counts and calculate the predicted mortality that results from the pulse, using a health impact function as in equation (5). To compare marginal damages with marginal abatement costs, we could monetize predicted mortality using a VSL. However, because scaling mortality by any positive constant does not affect our unit-specific comparison of damages with MAC curve steepness, we report damages in units of predicted mortality associated with a pulse.

Before reporting our results, we review the distributional predictions from a simple model of environmental markets. Figure 7 illustrates that a market will induce sources with flatter MAC curves to abate more compared to sources with steeper MAC curves. Consequently, a market is expected to produce larger avoided damages in empirical settings where emissions from sources with flatter MAC curves (higher abaters) systematically generate more damages than emissions from sources with steeper MAC curves (lower abaters) compared to a setting where the opposite is true—emissions from flatter-MAC curves generate *less* damages than emissions from sources with steeper MAC curves. The figures below rank units from most to least abatement. An empirical setting where a negative relationship exists between source-specific abatement rank and marginal damages will see more benefits (avoided damages) from an environmental market compared to an empirical setting with a positive relationship between abatement rank and marginal damages.

Figure 8 reports our estimates of unit-specific marginal damages of NO_x emissions generated by the pulse analysis on the y-axis, plotted against abatement rank on the x-axis, which is our ordering of abatement quantity under the NBP. As discussed earlier, this ranking is also an ordering of units based on the slope of their MAC curve. The 373 coal units for which we estimate a positive quantity of abatement are on the left of the x-axis, ordered

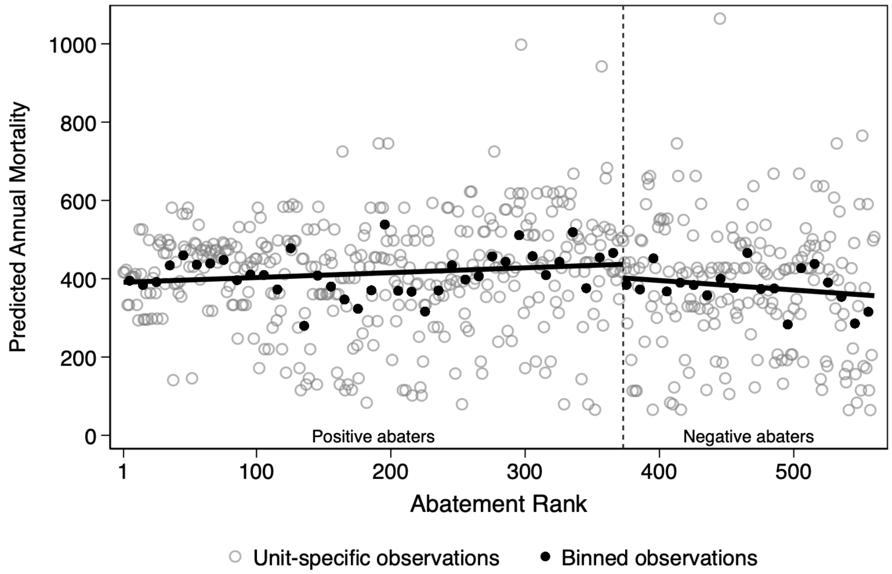
[30] Through repeated InMAP runs, we found that $PM_{2.5}$ concentrations scale approximately linearly with NO_x tons. For ease of interpretation, in this analysis we assume that each unit emits NO_x tons equivalent to the total NO_x tons emitted in an average regulated ozone season, roughly 445,000 tons. To scale our results to correspond to marginal damages from one ton of NO_x , one can scale by $1/(445 * 10^3)$. Since the analysis in this section is concerned with relative benefits and not total, our results are agnostic to the level of NO_x emitted by each unit, so long as that quantity is the same across units.

from most to least *abatement*. The 178 units for which we estimate no abatement or an increase in emissions, are on the right side of the x-axis, ordered from least to most emissions *increase* (negative abatement). The predicted mortality and abatement rank of each unit is represented by the circles, with two fitted linear regression lines summarizing how marginal damages correlate with relative abatement rank for the abating and non-abating units.

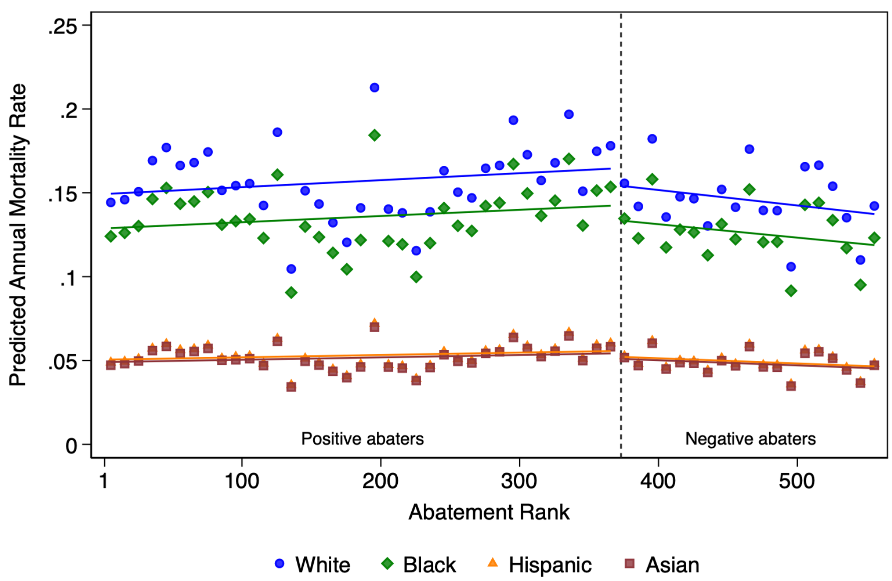
The estimated relationship for abaters is positive, with a slope coefficient of 0.12; the p-value for the null hypothesis that the slope is equal to zero is 0.04. Thus, we find that NO_x emissions from units that abate more have *smaller* marginal damages than units predicted to abate less. This weakly increasing relationship between abatement rank and marginal damage indicate that the added avoided mortality benefits of the market versus a non-market approach will be relatively small. For units that do not abate, i.e., increase emissions in response to NBP, the estimated slope is -0.22 and is statistically insignificant. Overall, the slopes for both abaters (ranks 1–373) and negative abaters (ranks 374–560) in Figure 8 are relatively small. For example, this means that the difference in the predicted number of annual deaths from a pulse of NO_x emissions between the smallest and the largest abater (predicted mortality at rank 373 minus rank 1) is 46. This outcome provides clear intuition for the results identified in Table 4. For example, larger slopes (in absolute value) in Figure 8 would generate starker differences in total benefits from a market compared to the uniform policy. The modest slopes linking marginal damages and abatement rank imply that aggregate predicted avoided mortality across market and non-market policies will be similar.

The second mechanism driving the results in Table 4 is the spatial distribution of demographic groups relative to each unit's abatement quantity. Demographic groups that are more likely to reside downwind of high-abatement units will receive greater benefits from a market than those residing downwind of low-abatement units. If demographic groups do not differentially reside downwind of higher- versus lower-abating units, differences in benefits across groups should be minimal. Figure 8b shows this empirical relationship, again using the results of the pulse analysis. That is, we use the estimates of the pulse analysis to gen-

Figure 8: Estimated Marginal Damages by Unit Abatement Rank



(a) All population



(b) By demographic group

Notes: Panel (a) shows the predicted number of deaths from NO_x emissions among units regulated by the NBP (y-axis). The x-axis ranks units from most to least estimated NBP induced abatement. Units in ranks 1–373 have positive estimated abatement, ordered from most to least abatement. Units 374–560 have negative estimated abatement (emissions increase), ordered from least to most increase. Panel (b) shows mortality rates (per 100,000) from each unit, with the same x-axis as in panel (a). The solid markers in panel (a) represent values average in bins of 10 units; all markers in panel (b) represent 10-unit binned values.

erate predicted mortality by demographic group. The results are shown as mortality rates (per 100,000 people) to account for differences in population sizes.

The results by demographic group indicate weakly positive slopes among abaters, with p-values from tests of the slope significance of about 0.08, providing no systematic evidence that marginal damages across demographic groups are correlated with abatement rank. In other words, we fail to reject the null of no relationship (slope of 0) between abatement rank and demographic group-specific damages from NO_x emissions of units regulated by NBP. The p-values for regressions of demographic group-specific predicted annual mortality rate on abatement rank among non-abaters are even larger, around 0.4, so we again fail to reject the null of no relationship in this group of units.

To compare the relationship between predicted annual mortality rate from a pulse and abatement rank across demographic groups, we regress predicted annual mortality rate on abatement rank, first among the 373 abating units. We estimate the regression separately by demographic group and use a Seemingly Unrelated Estimation (SUR) approach to compare coefficients from separate regressions.[31] We use Stata's *suest* routine to combine the four demographic group-specific models, and estimate a joint variance-covariance matrix for the coefficients. Under this joint system, we conduct Wald tests of the null hypothesis that the estimated slopes are equal across demographic groups. P-values from these tests range from 0.1–0.05. Thus, we do not find strong evidence that the slope linking predicted annual mortality rate and abatement rank varies by demographic group among abating units. We implement the same SUR approach for the non-abating units. The p-values are sufficiently large that we we cannot reject the null of no slope difference across demographic-groups.

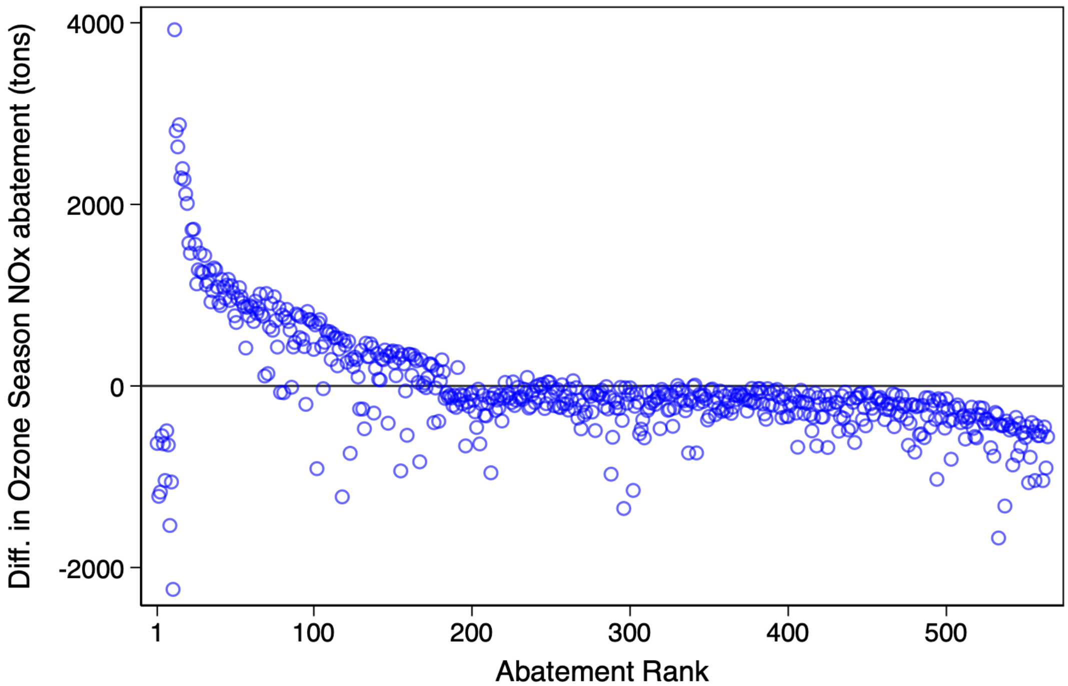
Figure 8 holds unit-specific abatement fixed. A final mechanism that can generate differences in the distribution of air quality benefits between market and non-market policies is unit-level variation in the level of abatement across the two policies. The magnitude of this mechanism depends on which units are predicted to exhibit large differences in abatement

[31] Because predicted mortality is demographic group-specific, a pooled regression with interaction terms is not workable.

across the two policies, and on the characteristics of the downwind populations affected by the emissions from those units. Figure 9 plots unit-level differences in abatement quantity in the NBP market compared to the uniform policy on the y-axis, by market abatement rank on the x-axis, where units are ordered along the x-axis in the same manner as in Figure 8. The largest differences in abatement between NBP and the counterfactual uniform policy occurs for units with abatement rank 1–150, with smaller differences for the remaining units. As we showed earlier in Figure 8, there is no evidence of systematic differences in the relationship between marginal damages and abatement rank when comparing this group of "high" abaters (ranks 1–150) compared to the other, "smaller" abaters (ranks 151–373).

Together, Figures 8a, 8b, and 9 demonstrate the mechanisms underlying the results in Table 4, which show that the NBP market and an aggregate abatement-equivalent non-market policy produce very similar benefits in terms of avoided mortality. Because marginal benefits and their demographic distribution show little systematic relationship with abatement rank, the market and uniform policy produce only small differences in both total and distributional benefits. These differences seem particularly small when comparing the differences in abatement costs, with the NBP market costing roughly one sixth of the cost of the uniform reduction policy.

Figure 9: Unit-Level Difference in NO_x Emissions in the NBP Market Compared to the Uniform Policy



Notes: This figure shows the difference in abatement in NO_x tons for coal units under the NBP market compared to a counterfactual policy that requires all coal units reduce NO_x emissions by an equal share. Values greater than 0 indicate more abatement in the market compared to the uniform reduction policy; values less than 0 indicate more abatement in the uniform reduction policy compared to the market. The x-axis ranks units from most to least abatement, as in Figure 8.

5 Conclusion

This paper develops an empirical framework to recover source-level abatement behavior and the associated distribution of costs and benefits under market-based environmental regulation. We leverage institutional variation from the NO_x Budget Program and high-frequency emissions data from U.S. electricity generating units to estimate source-specific abatement responses, construct marginal abatement cost curves, and connect source-level emissions changes to receptor-specific damages using an atmospheric transport model. Our methodological approach allows us to compare the distribution of abatement costs and benefits across regulated sources and across populations, both under the observed cap-and-trade system and a counterfactual non-market policy.

Several empirical findings emerge: first, total abatement costs under the NBP are substantially less—roughly one sixth—of previous empirical estimates of the benefits of the program. Second, a counterfactual policy that achieves the same amount of NO_x emissions abatement through a non-market regulatory regime is roughly six times as costly as the market. Third, the impact of the market on the distribution of air quality is largely uncorrelated with the demographic characteristics of receptor locations. We show that the joint distribution of marginal abatement cost curve slopes and marginal damages across regulated sources is sufficient to predict the distribution of air-quality benefits in a market-based pollution control system. We empirically characterize this distribution and find no evidence of a systematic relationship between MAC curve slopes and marginal damages from $PM_{2.5}$ in the NBP market.

Our results have implications for the evaluation and design of future environmental markets. We demonstrate a tractable approach for estimating source-specific treatment effects (abatement) induced by the policy using frontier econometric methods that requires data on source-specific outcomes during a regulated period, but not permit bids. For market-based regulatory regimes, we have shown how these estimates can be used to construct source-specific MAC curves, which enable the researcher to estimate not only costs of the regulation, but costs under alternative regulatory schemes as well.

This paper and previous literature discuss that the distributional outcomes of a market depend on the joint distribution of empirical objects, and there is no theoretical reason why results in one market should be generalizable to another. However, we show that ex-ante predictions are possible: if one observes or can estimate source-specific MAC curve slopes and marginal damages, one can predict the distributional outcomes of a market, which would be useful when designing future environmental markets. If regulated sources with flatter MAC curves account for a large share of damages to vulnerable and minoritized populations, then an environmental market can help reduce environmental disparities. Conversely, if sources with steeper MAC curves are responsible for most of the damages experienced by these populations, the market would not be expected to improve their outcomes. In the case of non-uniform pollutants, the relationship between MAC curve slope and marginal damages over all populations can predict the efficiency of the market and the extent to which trading ratios (Muller and Mendelsohn, 2009) or market decentralization (Doyle et al., 2014) should be pursued to improve efficiency.

References

W. J. Baumol and W. E. Oates. *The theory of environmental policy.* Cambridge university press, 1988.

K. Borusyak, X. Jaravel, and J. Spiess. Revisiting event-study designs: Robust and efficient estimation. *The Review of Economic Studies*, 91(6):3253–3285, 2024.

V. Chernozhukov, M. Demirer, E. Duflo, and I. Fernandez-Val. Generic machine learning inference on heterogeneous treatment effects in randomized experiments, with an application to immunization in India. Technical report, National Bureau of Economic Research, 2018.

J. Colmer, R. Martin, M. Muûls, and U. J. Wagner. Does pricing carbon mitigate climate change? firm-level evidence from the european union emissions trading system. *The Review of Economic Studies*, 92(3):1625–1660, 05 2024. ISSN 0034-6527. doi: 10.1093/restud/rdae055. URL `https://doi.org/10.1093/restud/rdae055`.

A. Dechezleprêtre, D. Nachtigall, and F. Venmans. The joint impact of the european union emissions trading system on carbon emissions and economic performance. *Journal of Environmental Economics and Management*, 118:102758, 2023. ISSN 0095-0696. doi: https://doi.org/10.1016/j.jeem.2022.102758. URL `https://www.sciencedirect.com/science/article/pii/S0095069622001115`.

O. Deschenes, M. Greenstone, and J. S. Shapiro. Defensive investments and the demand for air quality: Evidence from the NO_x Budget Program. *American Economic Review*, 107 (10):2958–89, 2017.

M. W. Doyle, L. A. Patterson, Y. Chen, K. E. Schnier, and A. J. Yates. Optimizing the scale of markets for water quality trading. *Water Resources Research*, 50(9):7231–7244, 2014.

U. S. E. P. A. EPA. NO_x Budget Trading Program. Technical Report EPA-430-R-08-008, Washington, DC: Environmental Protection Agency, December 2008.

U. S. E. P. A. EPA. Power sector emissions data, 2022a. URL `https://campd.epa.gov`.

U. S. E. P. A. EPA. Power sector facility attributes data, 2022b. URL `https://campd.epa.gov`.

M. Fowlie. Emissions trading, electricity restructuring, and investment in pollution abatement. *American Economic Review*, 100(3):837–69, 2010.

M. Fowlie, S. P. Holland, and E. T. Mansur. What do emissions markets deliver and to whom? Evidence from Southern California's NO_x trading program. *American Economic Review*, 102(2):965–93, 2012.

C. L. Gallagher, T. Holloway, C. W. Tessum, C. M. Jackson, and C. Heck. Combining satellite-derived pm2. 5 data and a reduced-form air quality model to support air quality analysis in us cities. *GeoHealth*, 7(5):e2023GH000788, 2023.

A. L. Goodkind, C. W. Tessum, J. S. Coggins, J. D. Hill, and J. D. Marshall. Fine-scale damage estimates of particulate matter air pollution reveal opportunities for location-specific mitigation of emissions. *Proceedings of the National Academy of Sciences*, 116 (18):8775–8780, 2019.

C. Grainger and T. Ruangmas. Who wins from emissions trading? Evidence from California. *Environmental and Resource Economics*, 71(3):703–727, 2018.

M. Greenstone, R. Pande, N. Ryan, and A. Sudarshan. Can pollution markets work in developing countries? experimental evidence from india. *The Quarterly Journal of Economics*, 140(2):1003–1060, 2025.

D. Hernandez-Cortes and K. C. Meng. Do environmental markets cause environmental injustice? evidence from california's carbon market. *Journal of Public Economics*, 217: 104786, 2023.

D. Hernandez-Cortes, K. C. Meng, and P. Weber. Decomposing trends in us air pollution: Disparities from electricity. *NBER Chapters*, 2022.

S. P. Holland, E. T. Mansur, N. Z. Muller, and A. J. Yates. Decompositions and policy consequences of an extraordinary decline in air pollution from electricity generation. *American Economic Journal: Economic Policy*, 12(4):244–74, November 2020. doi: 10.1257/pol.20190390.

C. Liu and K. Shi. A review on methodology in O_3-NO_x-VOC sensitivity study. *Environmental pollution*, 291:118249, 2021.

R. Martin, L. B. de Preux, and U. J. Wagner. The impact of a carbon tax on manufacturing: Evidence from microdata. *Journal of Public Economics*, 117:1–14, 2014. ISSN 0047-2727. doi: https://doi.org/10.1016/j.jpubeco.2014.04.016. URL `https://www.sciencedirect.com/science/article/pii/S0047272714001078`.

K. C. Meng and V. Thivierge. Do environmental markets improve allocative efficiency? evidence from us air pollution. Technical report, National Bureau of Economic Research, 2025.

N. Z. Muller and R. Mendelsohn. Efficient pollution regulation: getting the prices right. *American Economic Review*, 99(5):1714–1739, 2009.

R. Schmalensee and R. N. Stavins. The SO_2 allowance trading system: The ironic history of a grand policy experiment. *Journal of Economic Perspectives*, 27(1):103–22, 2013.

J. S. Shapiro and R. Walker. Where is pollution moving? environmental markets and environmental justice. In *AEA papers and proceedings*, volume 111, pages 410–414. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203, 2021.

J. S. Shapiro and R. Walker. Is air pollution regulation too lenient? evidence from us offset markets. *American Economic Review*, 115(9):3058–3080, 2025.

C. W. Tessum, J. D. Hill, and J. D. Marshall. Inmap: A model for air pollution interventions. *PloS one*, 12(4):e0176131, 2017.

U.S. Energy Information Administration. Emissions allowance prices for SO_2 and NO_x remained low in 2011, February 2012. URL `https://www.eia.gov/todayinenergy/detail.php?id=4830`.

U.S. Energy Information Administration. Form EIA-860 detailed data with previous form data (EIA-860A/860B), 2022. URL `https://www.eia.gov/electricity/data/eia860/`.

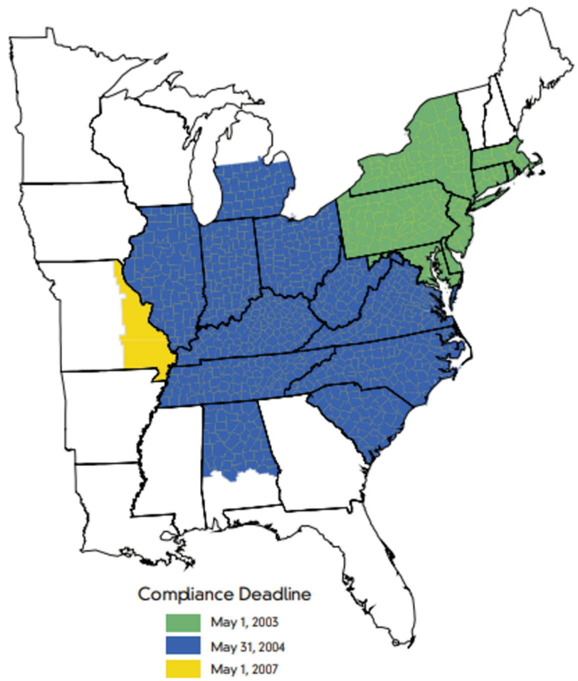
U.S. EPA. Clean Air Interstate Rule, Acid Rain Program, and Former NO_x Budget Trading Program, Progress report 2011, 2011.

V. Verdier. Estimating treatment heterogeneity by projection. *working paper*, 2022.

S. Wager and S. Athey. Estimation and inference of heterogeneous treatment effects using random forests. *Journal of the American Statistical Association*, 113(523):1228–1242, 2018.

Appendix

Figure 1: NBP Regulatory Status Over Time



Notes: Source: EPA (2008). This map shows regions in the United States regulated by the NBP, with colors denoting variation in the first year of required compliance with NBP.

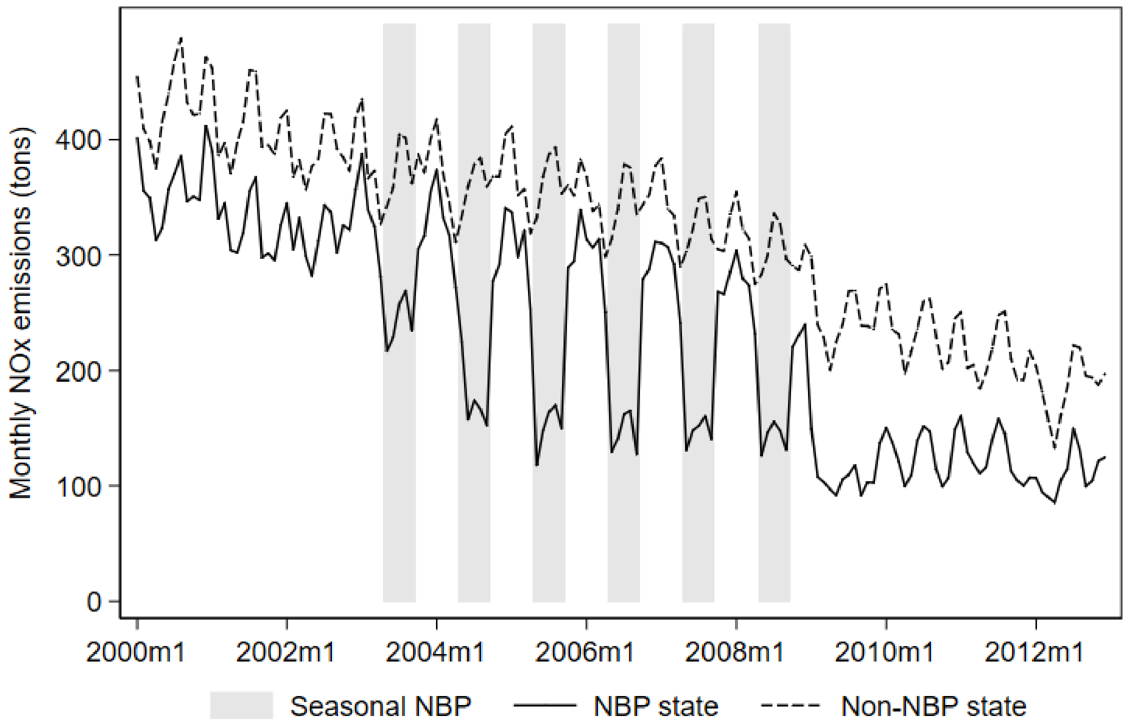
Table 1: States Regulated by NBP

Year	State	Count
2003	Delaware, Connecticut, Massachusetts, Maryland, New Jersey, New York, Pennsylvania, Rhode Island	8
2004	Alabama, Illinois, Indiana, Kentucky, Michigan, North Carolina, Ohio South Carolina, Tennessee, Virginia, West Virginia	11
2007	Missouri[1]	1
All years		20

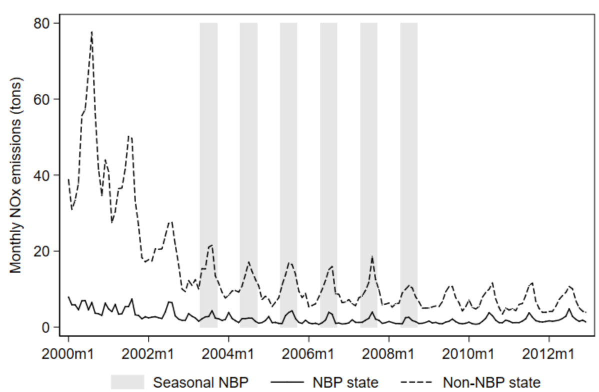
Notes: This table lists the states regulated by the NBP over the period 2003-2008, and the year in which the compliance began in ozone season months. (1) Only portions of Missouri were included in the program. The District of Columbia was also regulated under the NBP starting in 2004 Source: EPA (2008). The CAIR regulation added a seasonal NO_x emissions trading program for other states in 2009; those states are not considered as treated for the purposes of our study, which focuses on the period before CAIR was implemented.

Figure 2: Average Monthly NO_x of Electricity Generating Units by NBP Regulatory Status and Fuel Type

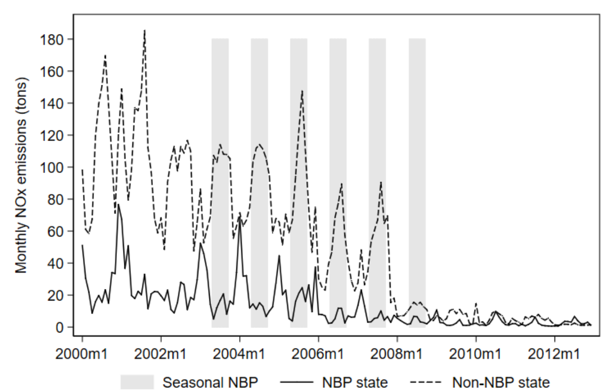
(a) Coal



(b) Natural Gas

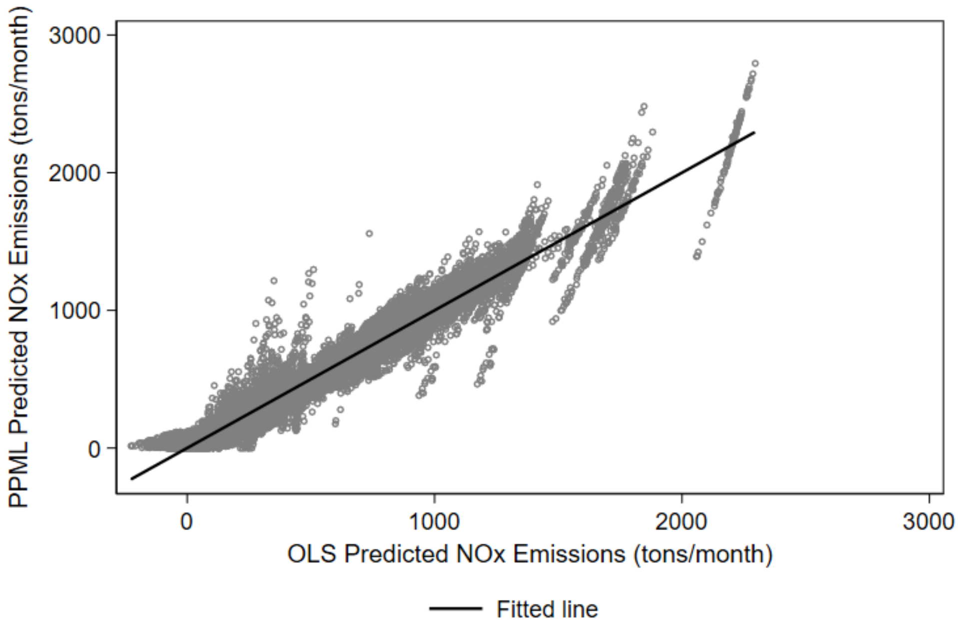


(c) Oil



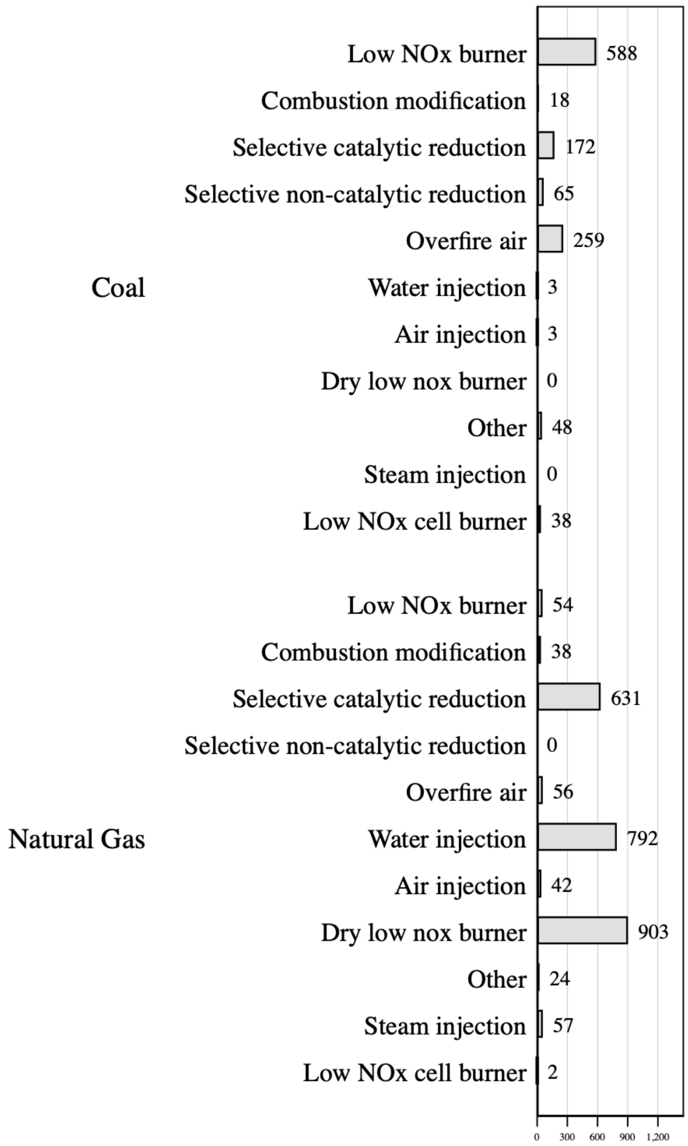
Notes: These figures are analogous to Figure 1, which plots average monthly NO_x emissions by NBP and non-NBP regulatory status, now separated by fuel type.

Figure 3: Comparing Predictions from OLS and PPML models



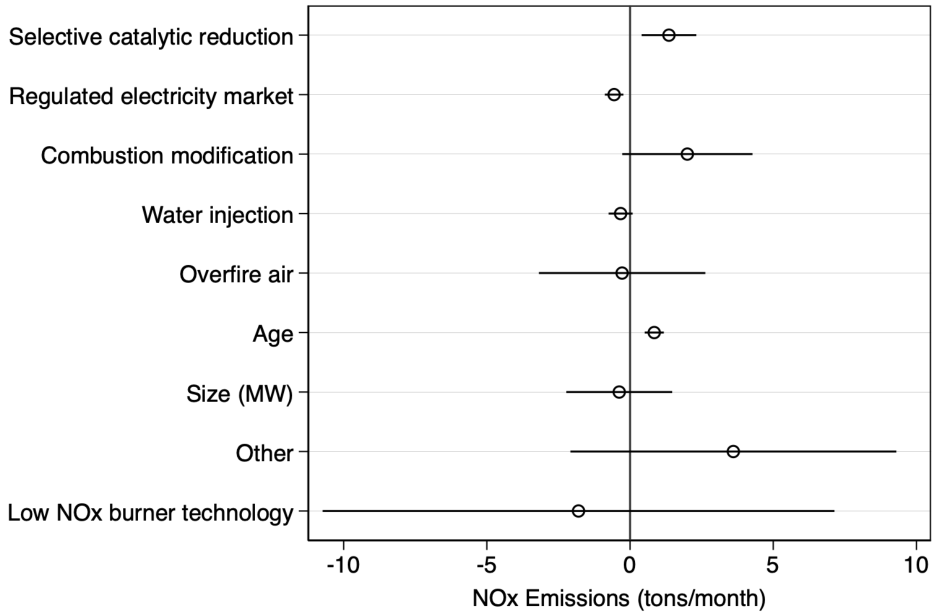
Notes: This figure plots monthly predicted NO_x emissions from our preferred PPML specification against the corresponding predictions from OLS. In both cases, the predicted emissions represent outcomes under the no-NBP counterfactual. The PPML and OLS models include state-by-year, state-by-month, and month-by-year fixed effects, along with observable predictors that best capture the determinants of NO_x emissions under the no-NBP scenario for each unit. These predictors include monthly load in the unit's NERC region and monthly temperature and precipitation in the county where the unit is located. We also include an indicator for whether a unit operated in a regulated (non-restructured) electricity market as of 2003.

Figure 4: NO$_x$ Control Technologies by Fuel, 2005

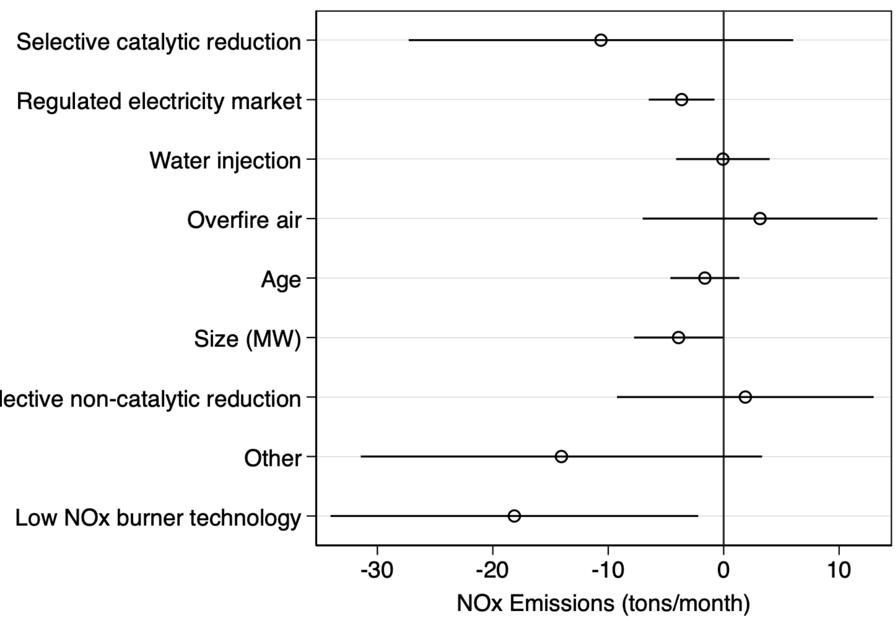


Notes: This figure shows total counts of NO$_x$ control technologies by fuel as of the year 2005, constructed using the EPA Clean Air Markets Data.

Figure 5: Heterogeneous Treatment Effects Estimates in the NO$_x$ Budget Program



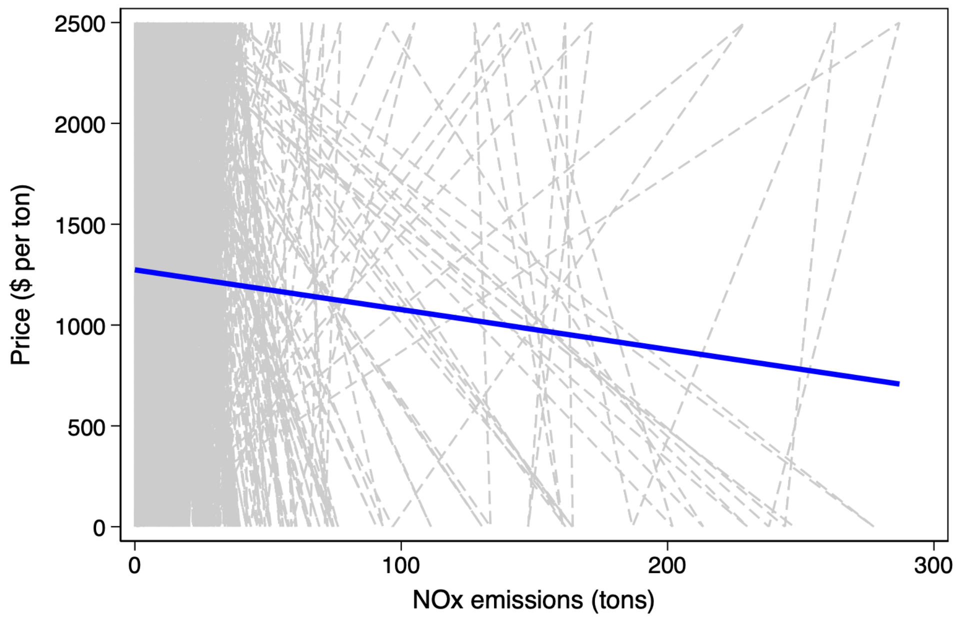
(a) Gas



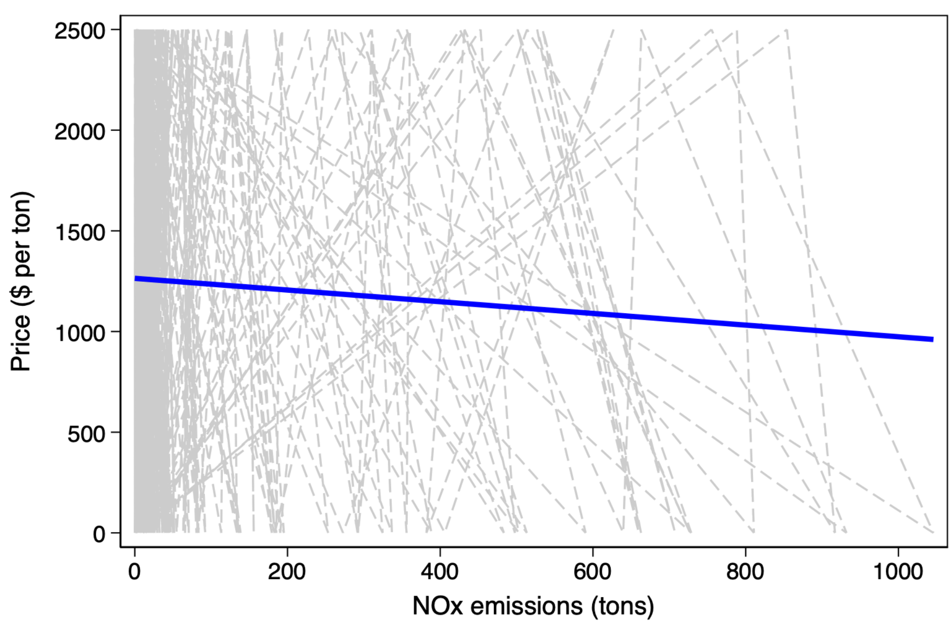
(b) Oil

Notes: These figures are analogous to Figure 6 but for natural gas and oil units, respectively: they shows the estimated coefficients and 95% confidence intervals from a regression of the estimated unit-level treatment effects (abatement) on observable covariates among treated coal units during the treated ozone season months. The continuous predictors (unit age and size) are standardized for better comparability. Standard errors are clustered by unit.

Figure 6: Estimated Unit-Specific Marginal Abatement Cost Curves by Fuel



(a) Gas



(b) Oil

Notes: These figures are analogous to Figure 7 panel (b) but for natural gas and oil, respectively: they show the distribution of estimated marginal abatement cost (MAC) curves by unit fuel type. Gray dashed lines correspond to unit-specific estimates, and the solid blue line represents the average MAC for an average unit of the respective fuel type.