Do environmental markets improve allocative efficiency? Evidence from U.S. air pollution

Kyle C. Meng and Vincent Thivierge*

September 17, 2025

Abstract

The central appeal of environmental markets – efficient allocation of emission reductions – has been difficult to establish empirically. We develop a framework linking the theoretical change in allocative efficiency following a market-based policy to a quasi-experimental estimator. We apply this framework with administrative data to two major U.S. air pollution markets. We find allocative efficiency for pollution improved by 3.3 percentage points annually under California's RECLAIM program but do not detect a change under the U.S.' NO_x Budget Trading Program. These results are supported by greater heterogeneity in baseline facility characteristics under RECLAIM and differences in program implementation.

*Meng: Bren School, Dept. of Economics, and emLab, UC Santa Barbara and NBER (email: kmeng@bren.ucsb.edu) Thivierge: Dept. of Economics, and Institute of the Environment, University of Ottawa (email: vincent.thivierge@uottawa.ca). We thank Spencer Banzhaf, Chris Costello, Olivier Deschênes, Antony Millner, Bernard Salanie, Gonzalo Vazquez-Bare, and members of the UCSB environmental economics research group, and seminar participants at the Université de Montréal, University of Ottawa, Georgia Tech, Resources for the Future, UC Berkeley, the 7th Annual LSE/Imperial Workshop on Environmental Economics, HEC Montréal, Carleton University, and the 11th Canadian PhD and Early Career Workshop in Environmental and Resource Economics. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau's Disclosure Review Board and Disclosure Avoidance Officers have reviewed this information product for unauthorized disclosure of confidential information and have approved the disclosure avoidance practices applied to this release. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2581. (CBDRB-FY22-P2581-R10061, CBDRB-FY22-P2581-R10140, CBDRB-FY22-P2581-R10141, CBDRB-FY23-P2581-R10769, CBDRB-FY24-P2581-R11386, CBDRB-FY24-P2581-R11664).

1 Introduction

Many economic policies introduce markets in settings where prices are traditionally absent. Such market-based policies have been designed to improve allocative efficiency in domains as wide ranging as education, healthcare, and food provision.[1] However, determining whether there are actual efficiency improvements from these policies is inherently challenging. When markets are missing, so too are the typical measures for allocative efficiency, making it difficult to determine any misallocation before the policy and changes thereafter.

This is a central challenge when evaluating environmental markets, which have been long viewed as a solution to environmental externalities (Coase, 1960; Arrow, 1969). Theoretical results from the 1960s first formalized that environmental markets can be cost-effective: they achieve an aggregate pollution target at minimum total cost through efficient pollution allocation (Kneese, 1964; Crocker, 1966; Dales, 1968; Baumol and Oates, 1971; Montgomery, 1972).[2] Furthermore, this can be achieved without a regulator knowing polluters' abatement costs, which can be hard to truthfully obtain. Later work noted that such efficiency gains may be dampened, or even reversed, in second best settings, putting the onus on empirical assessment (Malueg, 1990; Goulder et al., 1999; Fullerton and Metcalf, 2001; Godby, 2002; Harstad and Eskeland, 2010; Klenow, Pasten and Ruane, 2024; Kim, 2025). Still, the promise of cost-effectiveness continues to motivate the adoption of market-based policies across many environmental settings – in fisheries, groundwater, local air pollution, and greenhouse gas emissions – even as empirical evidence on realized efficiency gains remains limited.

This paper develops a quasi-experimental framework for estimating the change in allocative efficiency following an environmental market. Theory establishes our estimand: the change in misallocation cost across two policies. Allocative efficiency requires the marginal product of an input be equal across producers. Distortions create wedges between marginal products, leading to misallocation (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Restuccia and Rogerson, 2013). We leverage a theoretical link between a producer's unobserved input distortions to its observed average revenue per emissions through a first order condition. This relationship informs a difference-in-differences research design that isolates facility-level distortions and tests whether their dispersion changes following the introduction of a market. We show that under certain assumptions, our quasi-experimental estimator recovers a lower bound on our estimand.

Three features of our framework are well suited for the study of environmental markets. First, our theory allows for arbitrary input allocations before and after the policy, without assuming that the market-based policy necessarily improves efficiency. This generality accommodates a wide range of institutional contexts and ensures that our empirical test is two-sided: a market-based intervention can increase or decrease misallocation, consistent with second-best theory. Second, we allow for changes in the total quantity of the input across policies. This is essential for "cap-and-trade" environmental markets which lower total emissions in addition to reallocating emissions across polluters. Third, our quasi-experimental research design addresses several key challenges from the misallocation literature such as facility heterogeneity, endogenous elasticities, and macroeconomic changes over time.

We apply our framework to the introduction of two major U.S. markets for nitrogen oxides (NO_x): the

[1] See, for example, Ladd (2002); Epple, Romano and Urquiola (2017) on education, Roth, Sönmez and Ünver (2007); Agarwal et al. (2019) on healthcare, Prendergast (2022) on food banks, and Milgrom and Segal (2020) on spectrum auctions.

[2] Throughout the paper, we use allocative efficiency and cost-effectiveness interchangeably to describe the least-cost allocation of pollution across facilities under an aggregate pollution limit. This differs from social efficiency of pollution which balances marginal abatement cost with marginal pollution damages. Our analysis does not consider pollution damages.

southern California's Regional Clean Air Incentives Market (RECLAIM) and the eastern U.S. NO_x Budget Trading Program (NBP). The emission impacts of both programs have been extensively studied (Fowlie, Holland and Mansur, 2012; Deschenes, Greenstone and Shapiro, 2017). To estimate efficiency consequences, we construct a novel panel of facility-level average revenue per emissions, linking restricted-access revenue data from the U.S. Census Bureau with emissions data from state and federal environmental agencies for manufacturing facilities.

We find that RECLAIM and the NBP lowered manufacturing NO_x emissions by 24% and 23% respectively, roughly in line with prior studies. Applying our quasi-experimental estimator, we estimate that RECLAIM improved allocative efficiency consistently after its introduction at a rate of 3.3 percentage points annually, with the strongest effects for oil refineries and metal manufacturing facilities. By comparison, we do not detect allocative efficiency changes under the NBP. These contrasting results are consistent with differences in baseline facility heterogeneity: across a range of characteristics correlated with marginal pollution abatement costs, RECLAIM facilities exhibit greater variation than NBP facilities. Within policies, we find suggestive, though statistically imprecise, evidence that allocative efficiency improves when there are fewer frictions to emissions trading across facilities and when the market-based policy is more durable over time.

Our approach addresses limitations on previous methods to quantify the allocative efficiency impact of environmental markets. In theory, a polluter's abatement cost is the difference in optimal profit under no-policy and policy scenarios. In practice, most prior studies rely on the cost minimizing dual of this problem by assuming a particular cost function.[3] These studies must argue that all relevant inputs and prices are observed and vary exogenously. They must also assume that polluters do not alter output in the counterfactual, restricting a potentially important abatement margin.[4] Crucially, this prior approach implicitly assumes that a market-based policy necessarily improves efficiency, restricting researchers to the task of determining by how much.[5] Our approach starts with the profit maximization problem, using its first order condition to inform an observable proxy for marginal product of emissions.[6] We then build a two-sided test: an environmental market may increase or decrease misallocation based on changes in the dispersion of this proxy. Furthermore, our approach does not require the researcher to specify each polluters' marginal abatement cost curve, recognizing an informational limit that makes environmental markets appealing to begin with. In doing so, this paper expands the quasi-experimental literature on environmental markets to studying arguably its most important rationale: allocative efficiency.[7]

Our approach comes with limitations. First, we are unable to determine whether a market-based environmental policy achieved allocative efficiency, only that it led to more or less misallocation. Second, in contrast to studies that estimate a cost or production function, we can not analyze facilities' specific abatement decisions following a market-based policy to shed light on the behaviors that alter misallocation costs

[3]Seminal applications of this approach include ex-ante studies that forecast the allocative efficiency gains of hypothetical market-based policies (Gollop and Roberts, 1983, 1985; Carlson et al., 2000) and ex-post studies that quantify efficiency gains of realized policies (Keohane, 2006; Chan et al., 2018; Greenstone et al., 2025).

[4]More recent structural approaches that model both the demand and supply of environmental goods avoid these limitations through added structure (Rafey, 2023; Aronoff and Rafey, 2023; Hagerty, 2024).

[5]In ex-ante studies, a cost minimizing algorithm is applied to quantify the counterfactual market-based policy. In some ex-post studies, the counterfactual uniform pollution standard is modeled as an extra constraint on the cost minimization problem, which necessarily increases total costs relative to the market-based policy.

[6]This is similar to Anderson and Sallee (2011) which also use a first order condition to construct the unobserved shadow price of environmental regulation from observable objects.

[7]Prior quasi-experimental studies have focused on how environmental markets affect aggregate costs (Petrick and Wagner, 2014; Calel and Dechezleprêtre, 2016; Meng, 2017; Calel, 2020; Greenstone et al., 2025), aggregate benefits and their distribution (Fowlie, Holland and Mansur, 2012; Murray and Rivers, 2015; Deschenes, Greenstone and Shapiro, 2017; Lawley and Thivierge, 2018; Hernandez-Cortes and Meng, 2022; Colmer et al., 2022), or both aggregate costs and benefits (Ayres, Meng and Plantinga, 2021).

(Linn, 2008; Fowlie, 2010; Chan et al., 2018). Finally, our theory relies on theoretical assumptions to facilitate a mapping between our quasi-experimental estimator and the change in allocative efficiency. While these assumptions are employed elsewhere in the misallocation literature, they are nonetheless difficult to validate.

Finally, this paper draws from the misallocation literature. Input misallocation within an economy has been shown to be a strong determinant of aggregate productivity differences across economies (i.e., the indirect approach) (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Restuccia and Rogerson, 2013). More recently, researchers have turned to quasi-experimental approaches to examine the causes of misallocation (i.e., the direct approach) (Restuccia and Rogerson, 2017), with a focus on capital market liberalization policies (Bau and Matray, 2023; Sraer and Thesmar, 2023). As with Bau and Matray (2023), we argue that a quasi-experimental estimator can address potential concerns about measurement error (Bils, Klenow and Ruane, 2021) and misspecification (Asker, Collard-Wexler and De Loecker, 2014; Haltiwanger, Kulick and Syverson, 2018).

The rest of the paper is structured as follows. Section 2 provides background on environmental markets. Section 3 presents our conceptual framework, linking theory with our empirical research design. Section 4 discusses our data. Section 5 presents our main results. Section 6 concludes the paper. Appendix A, B, C, and D offer additional theoretical proofs, data, figures, and tables.

2 Background

2.1 Environmental markets and allocative efficiency

Environmental markets grew out of two strands of economic thinking over fifty years ago. The first was an institutionalist view, exemplified by Coase (1960), that excessive pollution arose from a lack of property rights to pollute. The second was Arrow (1969)'s notion from general-equilibrium theory that externalities (and thus pollution) can be regarded as a case of missing markets. Both views suggested a correction through a market for pollution. Building on these foundations, environmental economists recognized that environmental markets can in theory achieve an aggregate environmental target at minimal cost by allocating the environmental good across heterogeneous agents efficiently. This cost-minimization property was articulated in early proposals for markets in water quality (Kneese, 1964) and air pollution (Crocker, 1966; Dales, 1968) and theoretically established soon after (Baumol and Oates, 1971; Montgomery, 1972).[8]

Cost-effectiveness remains a central appeal for modern environmental markets. Sometimes called "cap-and-trade", such programs establish a limit (or cap) on total emissions by issuing a fixed supply of emission permits. Regulated facilities are then either given, or must purchase through auction or trade with other facilities, permits to cover their emissions. Cost-effectiveness has motivated the adoption of environmental markets in nearly every environmental domains. Today, pricing policies cover 20% of global fisheries catch (Costello and Ovando, 2019), account for over \$36 billion in global ecosystem service payments (Salzman et al., 2018), govern 20% of global greenhouse gas (GHG) emissions (World Bank, 2021), and govern many local air pollutants.

This promise of cost-effectiveness has also been critiqued theoretically and empirically. A theoretical second-best literature shows that markets may lead to smaller allocative efficiency gains, or even losses,

[8]For excellent reviews of this intellectual history, see Tietenberg (2010*a*), Tietenberg (2010*b*), Berta (2017), Banzhaf (2020) and Banzhaf (2023).

in the presence of existing distortions. Such distortions include market power in output markets (Malueg, 1990; Godby, 2002), complementary policies (Bohi and Burtraw, 1992; Fowlie, 2010), input taxation (Goulder et al., 1999; Fullerton and Metcalf, 2001), and distortions that come with the environmental market itself such as market power in permit markets (Hahn, 1984; Godby, 2002), transaction costs (Stavins, 1995), noncompliance (Malik, 1990), and rent-seeking (Harstad and Eskeland, 2010). From these second-best considerations emerged a more modest view on cost-effectiveness, namely that in real-world settings where other distortions exist, whether an environmental market improves allocative efficiency is an empirical question (Stavins, 1995), a point that echoes Demsetz (1969)'s critique of government interventions and indeed was raised initially in Coase (1960).

The empirical critique of whether environmental markets result in allocative efficiency is of a more epistemic nature. Early scholars of environmental markets recognized a fundamental informational limit to implementing cost-effective environmental policy (Samuelson, 1954). Setting such policies requires regulators to know the marginal abatement cost curves of every polluter, which are either unobserved or hard to truthfully obtain by a regulator. The impracticality of this informational requirement led researchers to recognize a secondary appeal of market-based policies: in (first-best) theory, an economy-wide environmental objective can be met at minimum cost without the regulator needing to know anything about polluters' marginal abatement cost curves. But within this feature lies an inherent tension with empirical validation: if it is challenging for a regulator to know marginal abatement cost curves, is it reasonable to assume that researchers can estimate such curves when trying to evaluate the allocative efficiency consequences of environmental markets?

This question arises with cost function estimation, the literature's prevailing approach for quantifying allocative efficiency impacts. First, in any cost function estimation, the researcher must argue that she observes all inputs and prices and that each varies exogenously. Second, for estimated cost functions to be valid for counterfactual policies, duality theory requires that output be unchanged in the counterfactual, restricting a potentially important abatement option (Malueg, 1990). Third, many cost function studies implicitly assumes that a market-based policy would necessarily lead to greater allocative efficiency than the policy it replaces. For example, in ex-ante studies, a cost minimizing algorithm is employed when characterizing the counterfactual market-based policy (Gollop and Roberts, 1983, 1985; Carlson et al., 2000). While in some ex-post studies, the counterfactual non-market policy is modeled as an extra constraint on the cost minimization problem, which necessarily increases total costs relative to the market-based policy (Keohane, 2006; Chan et al., 2018; Greenstone et al., 2025). More recent structural approaches that explicitly model the supply and demand of environmental goods overcome some of these previous limitations, but the added structure also imposes greater informational requirements for the researcher, in conflict with the core empirical critique (Rafey, 2023; Aronoff and Rafey, 2023; Hagerty, 2024).[9] Overcoming these concerns requires a different approach, one that is informative of allocative efficiency impacts but recognizes the informational limits that make environmental markets appealing to begin with.

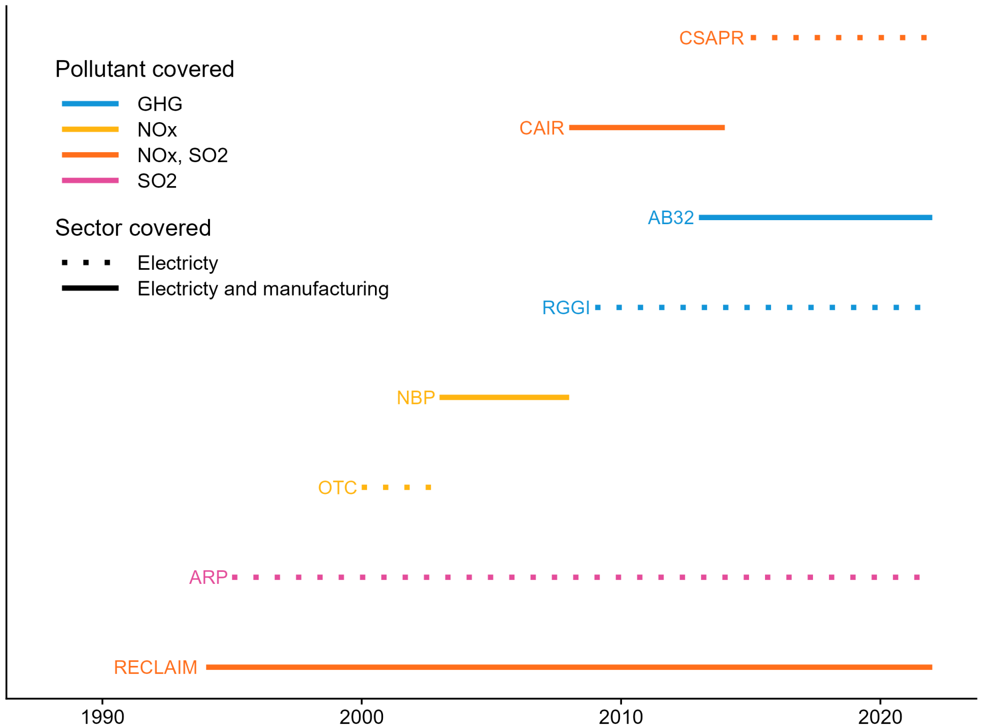
2.2 U.S. air pollution markets

U.S. air pollution is arguably the environmental domain where market-based policies have been most influential. The first pollution trading scheme - a market that offsets pollution reductions required by the U.S.

[9]Another approach is to estimate a distance output function following Färe et al. (1989, 1993). Because distance output, as a ratio of observed outputs to potential output under efficiency, is unobserved, its value relies heavily on functional form assumptions on how inputs and outputs map onto distance output, and exogeneity of these variables. Coggins and Swinton (1996), Swinton (2002), and Swinton (2004) conduct ex-post analyses of a market-based policy using this approach.

Clean Air Act (CAA) - was introduced in 1976 (see Shapiro and Walker (2024)). Other market-based interventions followed, leading eventually to regional- and national- air pollution cap-and-trade programs.[10]

Figure 1: U.S. air pollution cap-and-trade programs



Notes: Timeline of major local air pollution and greenhouse gas cap-and-trade markets in the U.S. from 1990 to 2020. The length of the line represents the start to end dates for each market. SO_2 and NO_x markets under CAIR and CSARP are bundled together for visual ease. Covered pollution and sectors also shown.

Figure 1 displays U.S. air pollution markets over the last three decades. For each market, we show its duration, the pollutants regulated, and sectors covered. In selecting markets for this study, we employ two criteria necessitated by our theoretical framework in Section 3. First, because we assume profit-maximizing facilities, we exclude programs, or facilities within programs, from the electricity sector given the presence of vertically-integrated electric utilities. This rules out the SO_2 Acid Rain Program (ARP), the Regional Greenhouse Gas Initiative (RGGI), the Ozone Transport Commission (OTC), and the Cross-State Air Pollution Rule (CSAPR).[11] Second, because our framework is static, we only evaluate programs that forbid or extensively limit the use of permit banking. This rules out the Clean Air Interstate Rule (CAIR) and California's AB32 greenhouse gas program, both of which does not forbid nor restrict the use of banked permits for future compliance. Two air pollution markets remain after applying our criteria, both for nitrogen oxides (NO_x): southern California's Regional Clean Air Incentives Market (RECLAIM) and the eastern U.S. NO_x Budget Trading Program (NBP).

RECLAIM The REgional CLean Air Incentives Market (RECLAIM) is a mandatory NO_x emission cap-and-trade program in southern California that was introduced in 1994 by the South Coast Air Quality

[10]See Carlin (1992) for a history of early air pollution markets.

[11]While the OTC NO_x market covered some manufacturing facilities, it covered only 30% of NBP facilities.

Management District (SCAQMD). The program mainly targets NO_x emissions, a precursor pollutant to ground-level ozone.[12] RECLAIM replaced over 40 prescriptive command-and-control (CAC) regulations imposed by the SCAQMD, many of which mandated adoption of specific pollution abating technologies, such as low-NO_x emitting industrial boilers. Under RECLAIM, facilities no longer needed to meet such controls beyond U.S. Clean Air Act (CAA) requirements.[13]

RECLAIM covers facilities emitting more than four tons of NO_x emissions annually. This includes nearly 400 facilities in Los Angeles, Orange, Riverside, and San Bernardino counties across the manufacturing, electricity, and oil and gas extraction sectors. Within the 300 manufacturing facilities, RECLAIM covers facilities across a wide range of industries, from food, cement, and metal manufacturing to petroleum refining. About 80% of observations are in 30 different 3-digit SIC sectors. In contrast to other cap-and-trade air pollution markets that typically monitors only emissions from large boilers, RECLAIM covers all on-site emissions, which potentially expands abatement options available to a facility.[14]

Annual permits to regulated facilities are freely allocated based on historical 1989-1992 emissions. Banking of permits for future use is prohibited with unused permits expiring at the end of each annual compliance period. Because prevailing winds blow inland from the coast, to reduce pollution hot spots, RECLAIM restricts the purchase of inland emission permits by coastal facilities, but not the other way around. As a result, coastal facilities face more trading restrictions than inland facilities.

Importantly for our analysis, while the market was introduced in 1994, the aggregate NO_x emission cap did not start binding until 2000. Indeed, covered emissions were below aggregate permit allocations during the early years of the program as shown in Figure A1. We follow previous RECLAIM studies and consider the treatment period starting when the cap first binds in 2000 (Fowlie, Holland and Mansur, 2012; Grainger and Ruangmas, 2018; Mansur and Sheriff, 2021). These prior studies find RECLAIM lowered NO_x emissions but not their distribution across nearby demographic groups (Fowlie, Holland and Mansur, 2012; Grainger and Ruangmas, 2018; Mansur and Sheriff, 2021) and that there is no statistically discernible relationship between facility-level permit allocation and emissions (Fowlie and Perloff, 2013).

NOx Budget Trading Program The NO_x Budget Trading Program (NBP) was a NO_x emission cap-and-trade market operated by the U.S. EPA from 2003 to 2008 designed to help states comply with ozone standards under the 1990 Clean Air Act Amendments. It was introduced in addition to, not instead of, previous CAA command-and-control requirements. In particular, NBP-covered facilities were required through earlier regulation to install Reasonably Available Control Technologies, a requirement that was maintained under the NBP (U.S. Environmental Protection Agency, 2007).

The NBP assigned each state a NO_x emission cap during summer months, when ground-level ozone concentrations were highest, across all large-emitting facilities.[15] Each state then determined how to al-

[12] RECLAIM also covers SO_x emissions in a separate permit market though the SO_x part of the program covered far fewer facilities. RECLAIM covered 30 to 40 SO_x emitting facilities compared with nearly 400 NO_x emitting facilities (South Coast Air Quality Management District, 2006).

[13] Anecdotally, the U.S. EPA noted several RECLAIM-covered facilities canceled orders for control technologies required under the previous SCAQMD CAC regulation following the introduction of RECLAIM, such as selective catalytic reduction units (U.S. Environmental Protection Agency, 2002).

[14] RECLAIM covers on-site NO_x emissions of facilities point sources ranging from heat input capacity of 2 MMBTU/hour to greater than 250 MMBTU/hour (South Coast Air Quality Management District, 2005). Emissions from larger sources are typically monitored using continuous emissions monitoring systems, whereas emissions from smaller sources are measured using engineering models and emissions factors. Emissions from smaller point sources account for roughly one-sixth of total covered NO_x emissions under RECLAIM (U.S. Environmental Protection Agency, 2002).

[15] In contrast to RECLAIM, the NBP regulated emissions only from major sources within a facility, such as a boiler, heat exchanger, or turbines with heat input capacity greater than 250 MMBTU/hour and not from other on-site point emissions sources.

locate permits across covered facilities within its cap. Altogether, the program covered over 700 facilities across 20 eastern states.[16] About 100 of these facilities were in manufacturing with the rest being power plants. Within covered manufacturing facilities, 90% were in four 4-digit NAICS sectors: pulp and paper manufacturing, chemical manufacturing, petroleum refineries, and primary metal manufacturing.[17] The NBP only operated between May and September. Permit banking, while allowed, was severely restricted: banked permits were discounted by 50 percent if used for future compliance, and only when the aggregate bank exceed the aggregate cap by 10% (U.S. Environmental Protection Agency, 2006). The result is that in a typical year, discounted banked permits used for compliance accounted for less than 1% of the yearly cap (U.S. Environmental Protection Agency, 2009). The aggregate cap was binding throughout the program's duration: Figure A2 shows the close tracking between aggregate covered emissions and the cap under the NBP. The NBP ended in 2009 when it was replaced by the Clean Air Interstate Rule (CAIR).

Prior NBP studies have largely focused on NBP impacts on electricity sector abatement costs (Fowlie, Knittel and Wolfram, 2012), capital and technology adoption (Linn, 2008; Fowlie, 2010; Popp, 2010), and emissions and resulting health consequences (Deschenes, Greenstone and Shapiro, 2017). The effects of different permit allocation rules within the NBP have also been studied (Lange and Maniloff, 2021). A few papers have examined NBP impacts on manufacturing facilities. Shapiro and Walker (2018) find that the implicit NO_x pollution tax level facing regulated manufacturing facilities doubled following the NBP. Curtis (2018) finds the NBP lowered manufacturing employment.

3 Conceptual framework

This section details our framework for estimating the change in allocative inefficiency following the introduction of an environmental market. Section 3.1 presents a stylized example for building intuition behind our approach. Section 3.2 presents a model of environmental policy leading to our estimand: a theoretical measure capturing the change in allocative inefficiency of emissions across two arbitrary policies. Section 3.3 links theory to data showing how our estimand can be recovered using a quasi-experimental estimator.

3.1 Stylized example

We start with a 2-facility example to illustrate the empirical challenge of estimating the change in allocative efficiency following the introduction of a market-based policy. Figure 2 shows emissions on the horizontal axis and its (shadow) price on the vertical axis. Facility 1 has a steeper marginal product of emissions curve than facility 2.[18] For a given allowable total emissions, E, there is a unique emissions allocation that minimizes total cost, indicated in Panel (a) by the sum of the shaded areas across the facilities which occurs when the marginal product of emissions is equalized across facilities at the economy-wide emissions price $\lambda(E)$. At this allocation, Facility 1 engages in less abatement than Facility 2.

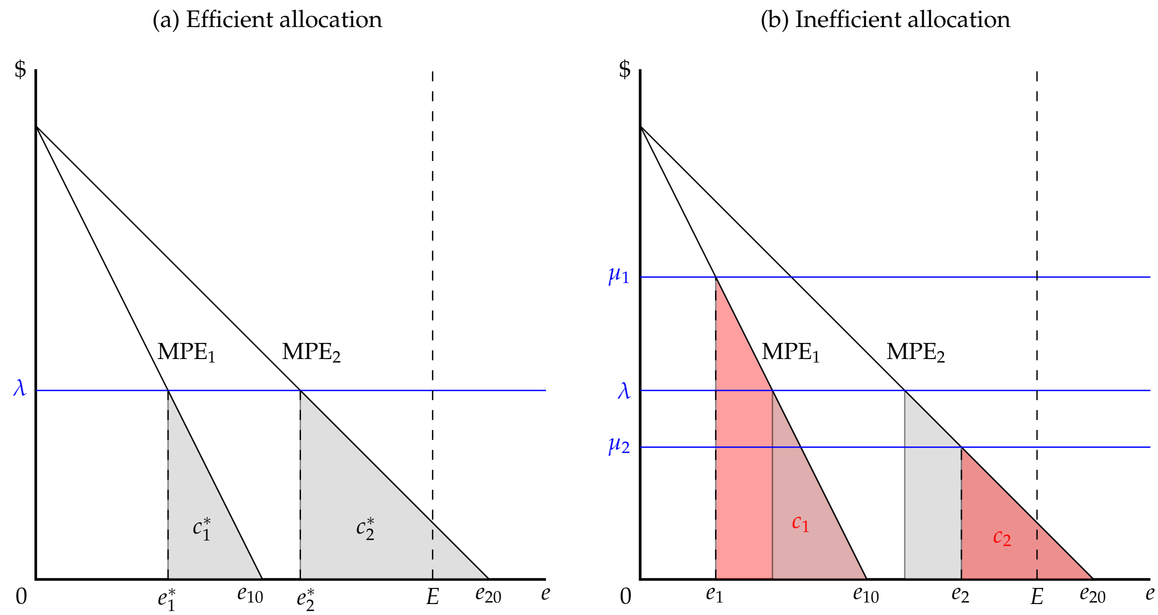
Next, consider when total emissions E is not efficiently allocated across facilities, as shown in Panel (b). The marginal product of emissions is no longer equalized with each facility facing its own emissions

[16] NBP-participating states include Alabama, Connecticut, Delaware, Illinois, Indiana, Kentucky, Maryland, Massachusetts, Michigan, Missouri, New Jersey, New York, North Carolina, Ohio, Pennsylvania, Rhode Island, South Carolina, Tennessee, Virginia, and West Virginia, and Washington, DC.

[17] 3-digit SIC classification is roughly at the same level of detail as 4-digit NAICS.

[18] The horizontal axes in Figure 2 indicates emissions rather than abatement in order to illustrate emissions levels when the emissions price is zero. When presented in terms of emissions abatement relative to the no-policy scenario, the marginal product of emissions curve becomes the marginal abatement cost curve.

Figure 2: Environmental policy and allocative (in)efficiency



Notes: Panels illustrate allocative efficiency in emissions for a 2-facility economy. Horizontal axes indicate emissions. Vertical axes indicate emissions (shadow) price. In Panel (a), total emissions E is allocated at minimum total cost (gray area) with facilities equating their marginal product of emissions to the economy-wide emissions price $\lambda(E)$. In Panel (b), facilities face separate emissions prices, resulting in misallocation and increased total cost (red area).

price, μ_i. There is too much abatement in one facility and not enough abatement in the other, leading total cost to increase. How much does the policy in Panel (b) deviate from allocative efficiency? In theory, the additional cost due to inefficient allocation in policy (b) is simply the difference in the total areas under the curves between panels (b) and (a). But that calculation is predicated on knowing each facility's marginal production of emissions curves, which face estimation challenges noted in Section 2.2.

There is an alternative approach. Rather than explicitly estimate each facility-level marginal product of emissions curve, perhaps the dispersion in input prices across facilities may be informative. The misallocation literature often draws on this insight when using the dispersion in distortion-inclusive input prices to quantify the aggregate productivity consequences of input misallocation (Restuccia and Rogerson, 2008; Hsieh and Klenow, 2009; Restuccia and Rogerson, 2013). To make progress, we adapt methods from the misallocation literature to estimates changes in total abatement cost from the introduction of market-based environmental policies. We now turn to such a framework.

3.2 Theory

Let $i = 1, ..., N$ index facilities using emissions e_i and input z_i in production function $q_i(e_i, z_i)$. Let $p(q_i)$ denote output price and w be price of input z_i. Let $\mathbf{e}^o = \{e^o_1, ..., e^o_N\}$ denote the vector of facility-level emissions in the absence of policy with $E_o = \sum_i e_{io}$. An environmental policy state s is similarly defined by two components: the vector of facility-level emissions $\mathbf{e}_s = \{e_{1s}, ..., e_{Ns}\}$ and total emissions across facilities, $E_s = \sum_i e_{is} \leq E_o$. Importantly, we place no restrictions on $\mathbf{e}_s$: any policy s - including the market-based policy - can have any allocation of emissions $\mathbf{e}_s$. This generality allows the ensuing empirical test in Section 3.3 to estimate allocative efficiency changes under any policy change, regardless of which policy – market-based or command-and-control – pre-existed or is being introduced.

In what follows, for a given policy s, we first define total abatement cost of going from E_o to E_s first under efficient allocation and then under $\mathbf{e}_s$. The cost of misallocation under policy s is the ratio of the latter over the former. We then compare this ratio across policies.

Total abatement cost under allocative efficiency To measure the cost of allocative inefficiency for policy s, one must first establish the total abatement cost of going from E_o to E_s under efficient allocation. Following Montgomery (1972), this is the solution to the regulator's problem of allocating E_s emissions across facilities to maximize total profit. That problem is

$$\begin{aligned}\Pi_i^* &= \max_{e_i, z_i} \ \sum_i p(q_i) q_i(e_i, z_i) - wz_i \\ &\qquad s.t. \ \ \sum_i e_i = E_s \\ &= \max_{e_i, z_i} \ \sum_i p(q_i) q_i(e_i, z_i) - wz_i - \lambda_s (\sum_i e_i - E_s) \end{aligned} \tag{1}$$

where $\lambda_s(E_s)$ is the economy-wide (shadow) emissions price on the total emissions constraint when facility-level emissions are allocated efficiently, henceforth denoted as λ_s. Under efficient allocation, the total abatement cost of going from E_o, total emissions in the absence of policy, to E_s is

$$\begin{aligned}\Delta\Pi_s^* &= (E_s - E_o)\Pi'_{s_{|E_o}} + (1/2)(E_s - E_o)^2 \Pi''_{s_{|E_o}} + \mathcal{O}^3 \\ &= (E_s - E_o)\lambda_o + (1/2)(E_s - E_o)(\lambda_s - \lambda_o) + \mathcal{O}^3 \\ &\approx (1/2)(E_s - E_o)\lambda_s \end{aligned} \tag{2}$$

where the first line applies a Taylor expansion around E_o. The second line observes that via the envelope theorem the derivative of optimized aggregate profit with respect to emissions is the aggregate shadow price. The third line notes that the shadow price at E_o is zero and then uses the first two terms of the Taylor series as an approximation.

Total abatement cost under a policy Next, we consider total abatement cost of going from total emissions E_o to E_s under a particular policy allocation $\mathbf{e}_s$. Optimal profit for facility i is

$$\begin{aligned}\pi_{is}(e_{is}) &= \max_{e_i, z_i} \ \ p(q_i) q_i(e_i, z_i) - wz_i \\ &\qquad s.t. \ \ e_i = e_{is} \\ &= \max_{e_i, z_i} \ \ p(q_i) q_i(e_i, z_i) - wz_i - \lambda_s \phi_{is}(e_i - e_{is}) \end{aligned} \tag{3}$$

where, following the misallocation literature, ϕ_{is} is a facility-level distortion, or wedge, that potentially breaks the equivalence between the aggregate shadow price under efficient allocation, λ_s, and the facility-level shadow price at e_{is}. Intuitively, the policy induces an efficient allocation of emissions when there are no distortions, $\phi_{is} = 1 \ \forall i$. Allocative inefficiency arises when distortions generate dispersion in facility-level shadow prices. Observe that eq. 3 encompasses a wide range of regulatory environments. For example, under a command-and-control regulation, one can view the regulator as explicitly setting each facility's $\lambda_s \phi_{is}$ in order to achieve a prescribed e_{is}. Under an emissions trading policy, the aggregate emissions cap would determine λ_s while ϕ_{is} would capture any facility-specific trading frictions.

Under policy s, the total abatement cost of going from the no-policy vector of emissions, $\mathbf{e}^o$, to the policy s vector of emissions, $\mathbf{e}^s$, is

$$\begin{aligned}\Delta\Pi_s &= \sum_i \Delta\pi_{is}(e_{is}) \\ &= \sum_i (e_{is} - e_{io})\pi'_{s_{|e_{io}}} + (1/2)(e_{is} - e_{io})^2 \pi''_{s_{|e_{io}}} + \mathcal{O}^3 \\ &= \sum_i (e_{is} - e_{io})\lambda_o\phi_{io} + (1/2)(e_{is} - e_{io})^2 \lambda_s\phi_{is} + \mathcal{O}^3 \\ &\approx \sum_i (1/2)(e_{is} - e_{io})\lambda_s\phi_{is} \end{aligned} \tag{4}$$

where the second line applies a Taylor expansion around e_{io}. The third line observes that by the envelope theorem the derivative of optimized profit with respect to emissions is the facility-level shadow price. The fourth line notes that the shadow price at e_{io} is zero and uses the first two terms of the Taylor series as an approximation.

Allocative inefficiency under a policy What is the cost of emissions misallocation under policy state s? For a given total emissions E_s, one can examine the ratio of total abatement cost under the policy to total abatement cost under allocative efficiency. Combining eqs. 2 and 4, this measure is

$$\theta_s = \frac{\sum_i (e_{io} - e_{is})\lambda_s\phi_{is}}{(E_o - E_s)\lambda_s} = \sum_i a_{is}\phi_{is} \tag{5}$$

where $a_{is} = \frac{e_{io} - e_{is}}{E_o - E_s}$ are weights capturing facility-level shares of total abatement with $\sum_i a_{is} = 1$.

Eq. (5) presents a difficulty. Observe that under allocative efficiency, $\phi_{is} = 1\ \forall i$ implies $\theta_s = 1$. However, the reverse is not in general true: $\theta_s = 1$ does not imply allocative efficiency. To make progress, we turn to two additional assumptions.

Assumption 1. *Facility distortions are distributed* $\phi_{is} \sim \mathcal{LN}(0, \sigma_s^2)$

Assumption 2. *Facility abatement share* a_{is} *is increasing in (a)* ϕ_{is} *and (b)* $\phi_{is}\frac{\partial a_{is}}{\partial \phi_{is}}$.

Assumption 1 imposes a lognormal distribution used elsewhere in the misallocation literature. This restriction ensures that $\phi_{is} > 0$ so that the facility-level shadow price of emissions is never negative.[19] Assumption 2 captures the idea that emission abatement is increasing in distortions but not excessively so.[20] This leads to our first proposition establishing our measure of allocative efficiency.

Proposition 1. *Under Assumptions 1, 2a and 2b, (a)* $\theta_s = 1$ *implies allocative efficiency, or* $\phi_{is} = 1\ \forall i$ *when* $\sigma_s^2 = 0$ *and (b)* θ_s *is increasing in* σ_s^2.

That is, not only does $\theta_s = 1$ imply allocative efficiency, but efficiency losses are increasing in the dispersion of distortions. Appendix A.1 provides the proof.

[19] We take a super-population perspective whereby our set of N facilities is drawn from a super-population of facilities with lognormally distributed distortions. Expectations are therefore taken over sampling uncertainty when drawing from the super-population.

[20] In Appendix A.2, we discuss how Assumption 2b also implies weakly positive abatement shares.

Allocative inefficiency change across policies Consider now two arbitrary policy states $s \in \{b, m\}$, where b indicates the baseline policy and m indicates a new policy, we can construct

$$\frac{\theta_m}{\theta_b} = \frac{\sum_i a_{im}\phi_{im}}{\sum_i a_{ib}\phi_{ib}} \tag{6}$$

When policy m is more efficient than policy b, we have $\frac{\theta_m}{\theta_b} < 1$. When policy m is less efficient, $\frac{\theta_m}{\theta_b} > 1$. Again, observe the generality behind eq. (6). The two policies compared can have any total emissions, E_s, and any allocation of emissions across facilities, $\mathbf{e}_s$. This allows one to study policy changes that lower (or increase) total emissions (i.e., a binding a cap) and does not presume market-based policies necessarily result in allocative efficiency gains.

3.3 From theory to empirics

Our estimand is $\frac{\theta_m}{\theta_b}$. Unfortunately, it has two unobserved elements: facility-level abatement shares a_{is} and facility-level distortions ϕ_{is}. Making progress on facility-level abatement shares is difficult as by definition it involves facility-level emissions and total emissions in the absence of the policy.[21] But suppose one could measure distortions, and in particular the ratio of expected distortions across policies, $\frac{E[\phi_m]}{E[\phi_b]}$. If so, the following proposition establishes a bounding argument:

Proposition 2. *Under Assumptions 1 and 2b, $\frac{E[\phi_m]}{E[\phi_b]}$ is a lower bound on $\frac{\theta_m}{\theta_b}$. That is, $\frac{\theta_m}{\theta_b} - \frac{E[\phi_m]}{E[\phi_b]} < 0$ if $\frac{\theta_m}{\theta_b} < 1$ and $\frac{\theta_m}{\theta_b} - \frac{E[\phi_m]}{E[\phi_b]} > 0$ if $\frac{\theta_m}{\theta_b} > 1$.*

Appendix A.2 provides the proof. Proposition 2 states that if policy m is more efficient than policy b, $\frac{E[\phi_m]}{E[\phi_b]}$ will understate that efficiency gain or $\frac{\theta_m}{\theta_b} < \frac{E[\phi_m]}{E[\phi_b]} < 1$. Likewise, if policy m is less efficient than policy b, then $\frac{E[\phi_m]}{E[\phi_b]}$ will understate the efficiency loss with $\frac{\theta_m}{\theta_b} > \frac{E[\phi_m]}{E[\phi_b]} > 1$. The argument utilizes a natural link between $\frac{E[\phi_m]}{E[\phi_b]}$ and the change in the variance of distortion across policies, an observation made elsewhere in the misallocation literature (Hsieh and Klenow, 2009).[22]

So how does one estimate $\frac{E[\phi_m]}{E[\phi_b]}$ when distortions are not directly observed? Our theory provides a proxy measure. Specifically, the first order condition for the facility problem in eq. (3) equates the marginal cost of emissions with its marginal revenue

$$\lambda_s\phi_{is} = (1+\xi_i)\kappa_i\frac{p_i q_{is}}{e_{is}} \tag{7}$$

where $\kappa_i = \frac{\partial q_i}{\partial e_i}\frac{e_i}{q_i} > 1$ is the output elasticity and $\xi_i = \frac{\partial p_i}{\partial q_i}\frac{q_i}{p_i}$ is the inverse price elasticity, both of which may be heterogeneous across facilities.[23] On the demand side, a growing literature documents heterogeneous markups, and thus demand elasticities, across facilities even within narrow sectoral definitions (Nevo, 2001; Hottman, Redding and Weinstein, 2016). On the supply side, facility-heterogeneity in output

[21] The possibility that an existing policy exists prior to a market-based policy suggests that e_{io} and E_o does not correspond to pre-change data.

[22] To see this, under Assumption 1

$$\frac{E[\phi_m]}{E[\phi_b]} = e^{\frac{\sigma_m^2}{2} - \frac{\sigma_b^2}{2}}$$

Since $\frac{1}{2}(var(\ln\phi_{im}) - var(\ln\phi_{ib})) = \frac{\sigma_m^2}{2} - \frac{\sigma_b^2}{2}$, $\frac{E[\phi_m]}{E[\phi_b]} > 1$ when $var(\ln\phi_{im}) - var(\ln\phi_{ib}) > 0$ and $\frac{E[\phi_m]}{E[\phi_b]} < 1$ when $var(\ln\phi_{im}) - var(\ln\phi_{ib}) < 0$.

[23] Profit maximization requires a firm to operate in the elastic portion of its demand curve such that $\frac{1}{\epsilon_i} > -1$.

elasticities provides the impetus for market-based environmental policies in the first place as they related to heterogeneity in abatement costs.

Rewriting eq. 7 as average revenue per emissions, $AR_{is} = \frac{p_i q_{is}}{e_{is}}$, and taking logs yields

$$\ln AR_{is} = \ln(1/(1+\xi_i)) - \ln \kappa_i + \ln \lambda_s + \ln \phi_{is} \tag{8}$$

Eq. (8) suggests a facility-level regression relating an observable variable, average revenue per emissions, with unobserved distortions. To bring this relationship to data, one needs a generalized version to address further empirical complications omitted thus far in the theory. First, there may be confounding macroeconomic changes that may alter the aggregate shadow price of emissions, λ_s or the dispersion of distortions, σ_s^2, that happens to coincide with the timing of the policy change.[24] These possibilities necessitate the use for a control group of facilities that are subject to the same macroeconomic changes but not the change in policy. Second, the first order condition in eq. (7) may be misspecified. For example, rather than being fixed, facility-specific demand and output elasticities may themselves be functions of distortions. If so, one wants to quantify misallocation as the combined consequence of both direct distortion effects and indirect effects mediated through changes in demand and output elasticities or other endogenous channels.

We address both empirical complications using a quasi-experimental approach, comparing treated with control facilities and estimating an effect of the policy, inclusive of all potential endogenous channels. Specifically, we employ a two-step quasi-experimental estimator. The first step estimates policy effects on mean parameters. The second step takes first-step residuals and estimates policy effects on the dispersion of residuals.

Letting $\mathcal{M}$ and $\mathcal{B}$ denote the set of treated and control facilities respectively, and t as year, our first step is an event study regression analog to structural equation (8)

$$\ln AR_{it} = \underbrace{\eta_i}_{\substack{\ln\left(\frac{1}{1+\xi_i}\right) - \ln \kappa_i \\ + \begin{cases} \ln \lambda_{b0} & \text{if } i \in \mathcal{B} \\ \ln \lambda_{m0} & \text{if } i \in \mathcal{M} \end{cases}}} + \underbrace{\gamma_t}_{\ln \lambda_{bt} - \ln \lambda_{b0}} + \sum_{\substack{-\underline{\tau} \leq \tau \leq \overline{\tau} \\ \tau \neq 0}} \underbrace{\alpha_\tau D_i \times \mathbf{1}(\tau = t)}_{\substack{(\ln \lambda_{mt} - \ln \lambda_{bt}) \\ -(\ln \lambda_{m0} - \ln \lambda_{b0})}} + \underbrace{\nu_{it}}_{\ln \phi_{it} + \varepsilon_{it}} \tag{9}$$

where D_i is a dummy variable that equals one for treated facility $i \in \mathcal{M}$ subject to the market-based policy. The facility-level fixed effect, η_i, captures captures facility-specific demand and supply side parameters, ξ_i and κ_i, respectively, as well as the aggregate shadow price for each respective group in the omitted year, or the last year before the policy change, $t = 0$. The year fixed effect, γ_t, captures any annual changes in the aggregate shadow price for the control group relative to the omitted year.[25] The coefficients α_τ, capture the difference in the aggregate price between treated and control facilities in each year τ relative to that difference in the omitted year. When $\tau < 0$, α_τ tests for the presence of pre-trends in the relative aggregate price. When $\tau > 0$, α^τ examines whether the aggregate shadow price changed due to the market-based policy. The residual ν_{it} in eq. (9) captures distortions, $\ln \phi_{it}$. It also contains any remaining error, ε_{it}, perhaps due to model misspecification or mismeasurement.

Eq. (9) is our most flexible specification, designed to detect the presence of pre-trends and time-varying

[24] For example, an increase in aggregate demand would drive up total emissions in the no-policy scenario, E_o, increasing $E_o - E_s$ and hence λ_s.

[25] The aggregate emissions price under efficient allocation is missing, and thus "shadow" for the control group and for the treatment group before a market-based policy. For the treated group after the policy, this price would equal the observed traded permit price only if the market-based policy resulted in efficient allocation of emissions.

policy change effects. To obtain an average treatment effect across the post change period, we also estimate differential trend-break and difference-in-differences versions of eq. (9)

$$\ln AR_{it} = \eta_i + \gamma_t + \alpha_1[D_i \times \mathbf{1}(\tau > 0)] + \alpha_2[D_i \times t] + \alpha_3[D_i \times \mathbf{1}(\tau > 0) \times t] + \nu_{it} \tag{9'}$$

$$\ln AR_{it} = \eta_i + \gamma_t + \alpha[D_i \times \mathbf{1}(\tau > 0)] + \nu_{it} \tag{9''}$$

In our second step, we square the predicted residuals $\widehat{\nu}_{it}$ after estimating eq. (9) and estimate a second-stage regression

$$\widehat{\nu}^2_{it} = \underbrace{\psi_i}_{\substack{\sigma_i^2 \\ +\begin{cases}\sigma^2_{b0} & \text{if } i \in \mathcal{B} \\ \sigma^2_{m0} & \text{if } i \in \mathcal{M}\end{cases}}} + \underbrace{\upsilon_t}_{\sigma^2_{bt}-\sigma^2_{b0}} + \sum_{\substack{-\underline{\tau} \le \tau \le \overline{\tau} \\ \tau \neq 0}} \underbrace{\beta_\tau D_i \times \mathbf{1}(\tau = t)}_{\substack{(\sigma^2_{mt}-\sigma^2_{bt}) \\ -(\sigma^2_{m0}-\sigma^2_{b0})}} + \epsilon_{it} \tag{10}$$

where the facility-level fixed effect, ψ_i, captures any baseline differences in the dispersion of distortions between treated and treated facilities in the omitted year. The year fixed effect, υ_t, captures annual changes in the dispersion of distortions for the control group relative to the omitted year.

Our main coefficients of interest are β_τ. The flexible function form of eq. (10) tests for pre-trends and time-varying policy change effects. When $\tau < 0$, β_τ examines pre-trends in the relative dispersion of distortions between treated and control facilities, relative to the omitted year. When $\tau > 0$, β_τ estimates the difference in the dispersion of distortions between treated and control facilities following the policy change, relative to the omitted year. This maps to the ratio of expected distortions across policies: $e^{\frac{\beta_\tau}{2}} = \frac{E[\phi_{mt}]}{E[\phi_{bt}]}$. For ease of exposition, we also report $(1 - e^{\frac{\beta_\tau}{2}}) \times 100 = (1 - \frac{E[\phi_{mt}]}{E[\phi_{bt}]}) \times 100$, the lower bound percentage point change in allocative efficiency. Observe that these reduced-form coefficients incorporate any endogenous changes in facility-level parameters - such as demand and output elasticities - in response to distortions and as such is inclusive of potential misspecification in these parameters in the first order condition contained in eq. (7).

As with our first stage estimation, we also consider differential trend-break and difference-in-difference versions of eq. (10)

$$\widehat{\nu}^2_{it} = \psi_i + \upsilon_t + \beta_1[D_i \times \mathbf{1}(\tau > 0)] + \beta_2[D_i \times t] + \beta_3[D_i \times \mathbf{1}(\tau > 0) \times t] + \epsilon_{it} \tag{10'}$$

$$\widehat{\nu}^2_{it} = \psi_i + \upsilon_t + \beta[D_i \times \mathbf{1}(\tau > 0)] + \epsilon_{it} \tag{10''}$$

Finally, across all first- and second-step estimating equations, we cluster standard errors at a broader jurisdictional level (i.e., zip code under RECLAIM and county under the NBP) to account for arbitrary forms of spatial correlation and serial correlation in the residual within facilities of that jurisdiction.[26]

4 Data

Our empirical framework requires a facility-year panel of emissions and revenue covering regulated and unregulated facilities, before and after the policy change. To achieve this, we construct a new link between restricted-access facility-level U.S. Census data from the Annual Survey of Manufactures (ASM) and the Census of Manufactures (CM) – henceforth ASMCM – with data on air pollution emissions and air pollu-

[26] We can not cluster at the county level for RECLAIM as there are only 4 counties in the treated group, which would lead to a over-rejection of the null (Carter, Schnepel and Steigerwald, 2017).

tion markets from state and federal environmental agencies.[27] This section details our pollution data, the ASMCM data, and their merge.

CARB data To study RECLAIM, we use the merged California Air Resources Board (CARB) and SCAQMD data compiled by Fowlie, Holland and Mansur (2012). This data provides facility-year NO_x emissions and facility characteristics across California for 1990, 1993, and annually from 1995 to 2005.[28] The data covers all facilities emitting above 10 tons of NO_x pollution per year.[29] Among CARB facility-level characteristics, we use facility name, address, SIC code, zip code, and county code for our ASMCM matching procedure (detailed below). We use SCAQMD data to define treated facilities regulated under RECLAIM. We use all other California manufacturing facilities in the CARB dataset as control facilities.

U.S. EPA data To study the NBP, we use facility-year NO_x emissions and facility characteristics data from the U.S EPA National Emissions Inventory (NEI), which reports emissions of criteria pollutants for point sources emitting at least 100 tons per year of one of the criteria air pollutants. NEI reports emissions every three years, which we use from 1996 to 2008.[30] Among NEI facility-level characteristics, we use facility name, address, NAICS code, zip code, and county code for our ASMCM matching procedure (detailed below). We also obtain cross-sectional measure of facility-level pollution and abatement capital expenditure from the 1999 restricted-use U.S. EPA Pollution Abatement Costs and Expenditures survey (PACE), accessed through the U.S. Census Bureau. NEI does not record the facility-level NBP treatment status. That information is contained in the U.S. EPA's Air Market Program Data (AMPD).[31] We link NEI and AMPD data using the common facility AMPD identifier available through U.S. EPA's Facility Registry Service (FRS). For NEI facilities without an AMPD identifier, we obtain NBP treatment status from Curtis (2018) which constructs NEI-AMPD links using other U.S. EPA FRS identifiers. We use all other manufacturing facilities in the NEI in the contiguous U.S. outside of California as control facilities.[32]

U.S. Census Bureau data Our restricted-access facility-level revenue measure is total value of shipment contained in the ASMCM. The CM is conducted in five year intervals in years ending with 2 and 7 and the ASM is conducted every non-CM year. The CM contains about 300,000 manufacturing facilities, while the ASM covers a subset of approximately 50,000 facilities. For ASM years, the 10,000 largest facilities by revenue are selected with certainty, with the remaining 40,000 randomly selected for a representative sample. We use the U.S. Census Bureau's Longitudinal Business Database (LBD) to create a panel of facilities linking ASM and CM data from 1990 to 2005 (Chow et al., 2021). We use the LBD facility identifier as our unique facility identifier for facility fixed effects in the analysis.[33] The LBD identifier also allows us to merge

[27] Previous papers have matched panel of US facility-level pollution to a single year of ASM data (Shapiro and Walker, 2018) or used third-party facility-level data as a proxy for facility revenue. For example, Cherniwchan (2017); Cui, Lapan and Moschini (2016) match facility-level data from the privately-constructed National Establishment Time-Series (NETS) to facility-level pollution data from U.S. EPA. Facility-level revenue in NETS is imputed as the product of facility-level employment and industry-level sales per employee. We directly use facility-level revenue reported to ASMCM and match that to facility-level pollution over multiple years.

[28] Emissions for the years 1991, 1992, and 1994 are not available.

[29] Facility-level emissions data is also collected by the U.S. EPA. But because the federal reporting threshold is higher – at 100 tons per year – U.S. EPA data covers fewer NO_x-emitting facilities than CARB data.

[30] Due to a budget cut in 2005, about 1/3 of facilities reported the 2002 emissions for 2005 (Cui, Lapan and Moschini, 2016). We drop these facilities from our sample.

[31] The AMPD also provides annual facility emissions but only for sample facilities that are regulated by the NBP and not those that are unregulated by the NBP. Thus, any use of AMPD data for treated facilities would require use of NEI data for control facilities. We avoid issues with different data quality across treatment and control groups by using only NEI emissions data.

[32] We omit California facilities in our study of the NBP effects to avoid complications with RECLAIM effects entering into our NBP control group.

[33] The LBD identifier has been cleaned and scrutinized by U.S. Census Bureau researchers over several decades (Chow et al., 2021).

facility names and address from the U.S. Census Bureau Standard Statistical Establishment List (DeSalvo, Limehouse and Klimek, 2016). We retain the NAICS and SIC industry classifiers, zip code, and FIPS county code from the LBD in our merged ASMCM panel.

Data linkage We develop a new linking algorithm to merge facility-year state and federal pollution data with facility-year ASMCM revenue data. Our procedure builds on established Census Bureau procedures for processing data, identifying potential matches, and resolving multiple matches (Massey and O'Hara, 2014; Cuffe and Goldschlag, 2018). Our algorithm uses several non-unique identifiers: facility name, facility address, industry classifiers, zip code, and county codes. Appendix B provides further details.

For RECLAIM, we match about 70% of treated manufacturing facilities reporting to CARB to the ASMCM, and about 40% of control facilities.[34] For the NBP, we match over 90% and about 70% of our treated and control facilities with emissions data, respectively, to the ASMCM.

5 Results

This section presents our RECLAIM and the NBP results. For each program, we first report average effects on facility emissions, confirming results from earlier studies. Next, we discuss results from the first-stage of our estimation procedure, which capture impacts on the average (shadow) price of emissions. We then present our second-stage estimation results, showing program impacts on the variance of first-stage residuals, our proxy measure of allocative efficiency changes. Section 5.1 presents RECLAIM results while Section 5.2 presents the NBP results. Section 5.3 explores potential mechanisms behind the programs' different effects.

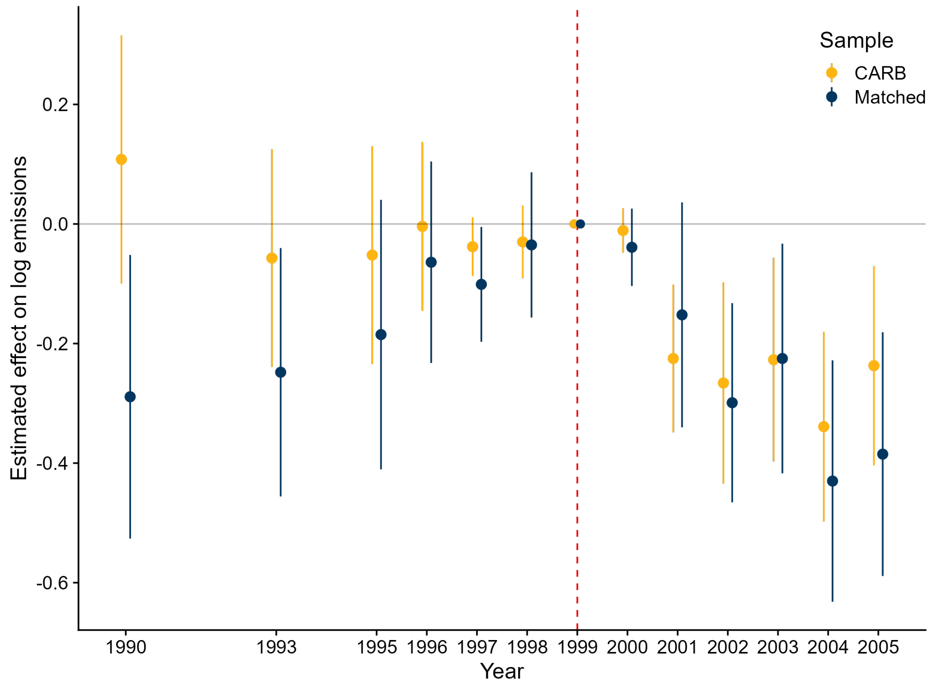
5.1 RECLAIM

To build confidence in our research design, we first examine RECLAIM effects on NO_x emissions, an outcome that has been previously studied. Figure 3 shows the effect of RECLAIM on emissions using the event-study specification in eq. 9 but with log NO_x emissions as the outcome. We show results both for the full sample of manufacturing facilities in CARB's emissions dataset (in gold) and the subsample of CARB-ASMCM matched facilities (in blue) used in subsequent results that depend on revenue data.

We find that NO_x emissions fall steadily for regulated facilities after 1999 in both estimating samples. Pre-treatment coefficients do not differ statistically from zero in the full CARB sample. For the matched sample, there are a few pre-treatment coefficients that differ from zero but the direction of the pre-trend is in the opposite direction of the post-treatment trend. These effects are confirmed by the differential trend-break specification of eq. (9′) shown in Columns (1) and (2) of Panel (a) in Table 2. By 2005, NO_x emissions are 24% (0.28 log points) and 34% (0.42 log points) lower than in 1999 across these two samples, implying an annual decline rate of 4.1% and 5.7% respectively. For comparison, using a sample that further includes power plants and oil and gas extraction facilities, Fowlie, Holland and Mansur (2012) find a reduction of 26% (0.30 log points) between average 1990-1993 emissions and 2004-2005 emissions. Estimates from the difference-in-differences specification of eq. (9″) shown in Columns (1) and (2) of Panel (b) indicate that NO_x are 17% (0.18 log points) and 11% (0.12 log points) lower on average during 2000-2005 respectively for

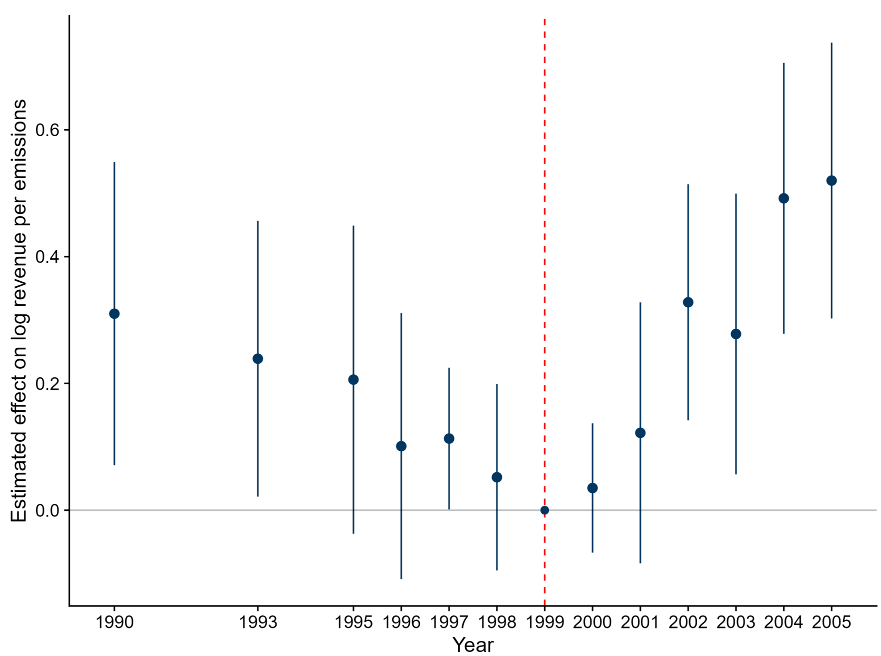
[34]One reason for the differential match rate is that the average RECLAIM facility emits more than the average control facility in California, and smaller emitting facilities, if they also have less revenue, are less likely to be sampled in the ASM.

Figure 3: Event-study estimates of RECLAIM effect on emissions



Notes: Point estimates and 95% confidence intervals of the annual effect of RECLAIM on log NO_x emissions relative to 1999 using eq. (9). Estimates for the full sample of manufacturing facilities in CARB shown in gold and for the CARB-ASMCM matched sample shown in blue. Standard errors are clustered at the zip code level.

Figure 4: Event-study estimates of RECLAIM effect on revenue per emissions

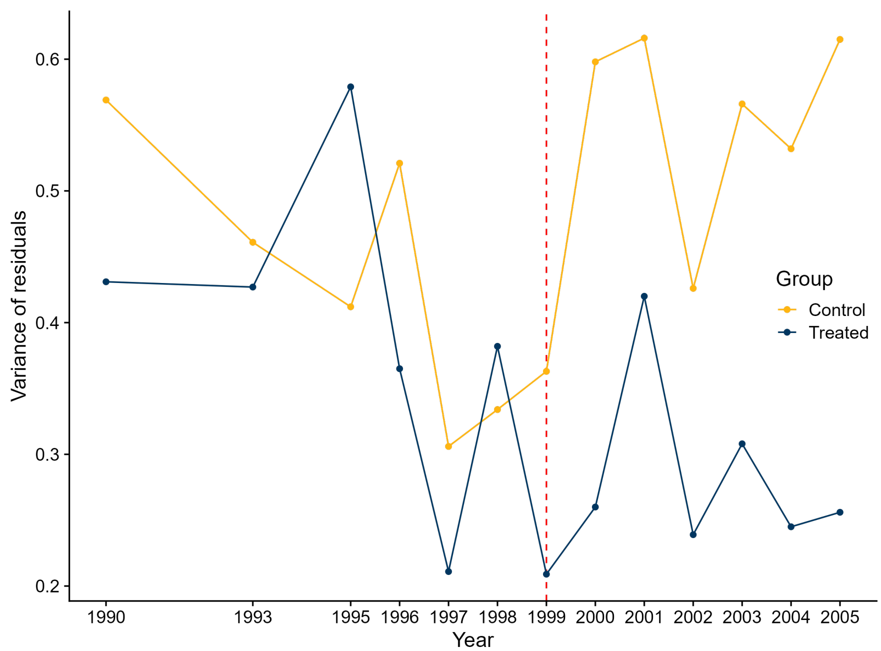


Notes: Point estimates and 95% confidence intervals of the yearly effect of RECLAIM on log revenue per emissions relative to 1999, or $\widehat{\alpha}^{\tau}$ using eq. (9). Standard errors are clustered at the zip code level.

the two samples.[35] The fall in average - and thus total – emissions reinforces the importance of a framework that allows total emissions target to change across policies, as we allow in Section 3.

Turning now to estimation informed by our framework. We begin with our first stage estimates of the RECLAIM effect on revenue per emissions, which our model interprets as the effect on the aggregate shadow price of NO_x emissions for regulated facilities under efficient allocation, or λ_s. If RECLAIM lowered emissions, one should expect this price to rise. Figure 4 shows estimates of α_τ from eq. 9, or the difference in the emissions price between treated and control facilities for each year, relative to their difference in 1999. Estimates of α_τ rise steadily after RECLAIM. While there is a pre-trend in estimates of α_τ, they are in the opposite direction of the post-trend. The differential trend-break specification of eq. (9′) shown in Column (3) of panel (a) in Table 2 find that the aggregate emissions price increased by 68% (0.52 log points) in 2005 relative to 1999, or at an annual rate of 11.3%. Estimates from the difference-in-differences specification in eq. (9′′) in Column (3) of Panel (b) indicate that the aggregate emissions price fell by 15% (0.142 log points) on average during 2000-2005.

Figure 5: Variance in predicted residuals for treated and control facilities under RECLAIM



Notes: Annual variance of the predicted residual, $\widehat{v}_{it}$, from eq. 9 for treated (blue) and control facilities (gold).

Our theory-informed estimator of the change in allocative efficiency under RECLAIM from Section 3 examines whether the program widened or narrowed the variance of distortions, interpreted as regression residuals after statistically accounting for other determinants. To build confidence in our allocative efficiency result, we first provide two examinations of these residuals.

In Figure 5, we take regression residuals from our first-step estimation of eq. (9) on revenue per emission, $\widehat{v}_{it}$, and construct their variances separately for treated and control facilities and for each year of our sample. Residual variances for treated and control facilities roughly track each other before RECLAIM and then diverge thereafter with lower residual variances for treated facilities, consistent with greater allocative

[35] RECLAIM estimates between the trend-break and difference-in-differences models across outcomes differ because the trend-break model allows for a flexible pre-trend whereas the difference-in-differences model does not.

efficiency for treated facilities. This gap also appears to grow over time after RECLAIM.

Table 1: Heterogeneity in RECLAIM effects by pre-treatment distortion

	(1) Residual	(2) Log emissions
RECLAIM X Post	-1.247^{***}	0.181^{***}
	(0.286)	(0.067)
RECLAIM X Post X Low distortion	1.384^{***}	-0.645^{***}
	(0.230)	(0.112)
Observations	11,500	11,500

Notes: Estimates of the differential effect of RECLAIM on residuals (Column 1) and log emissions (Column 2) by whether a RECLAIM facility had below-median pre-RECLAIM average residual. Specification based on eq. 9 and includes facility and year fixed effects. Robust standard errors clustered at the zip code level in (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

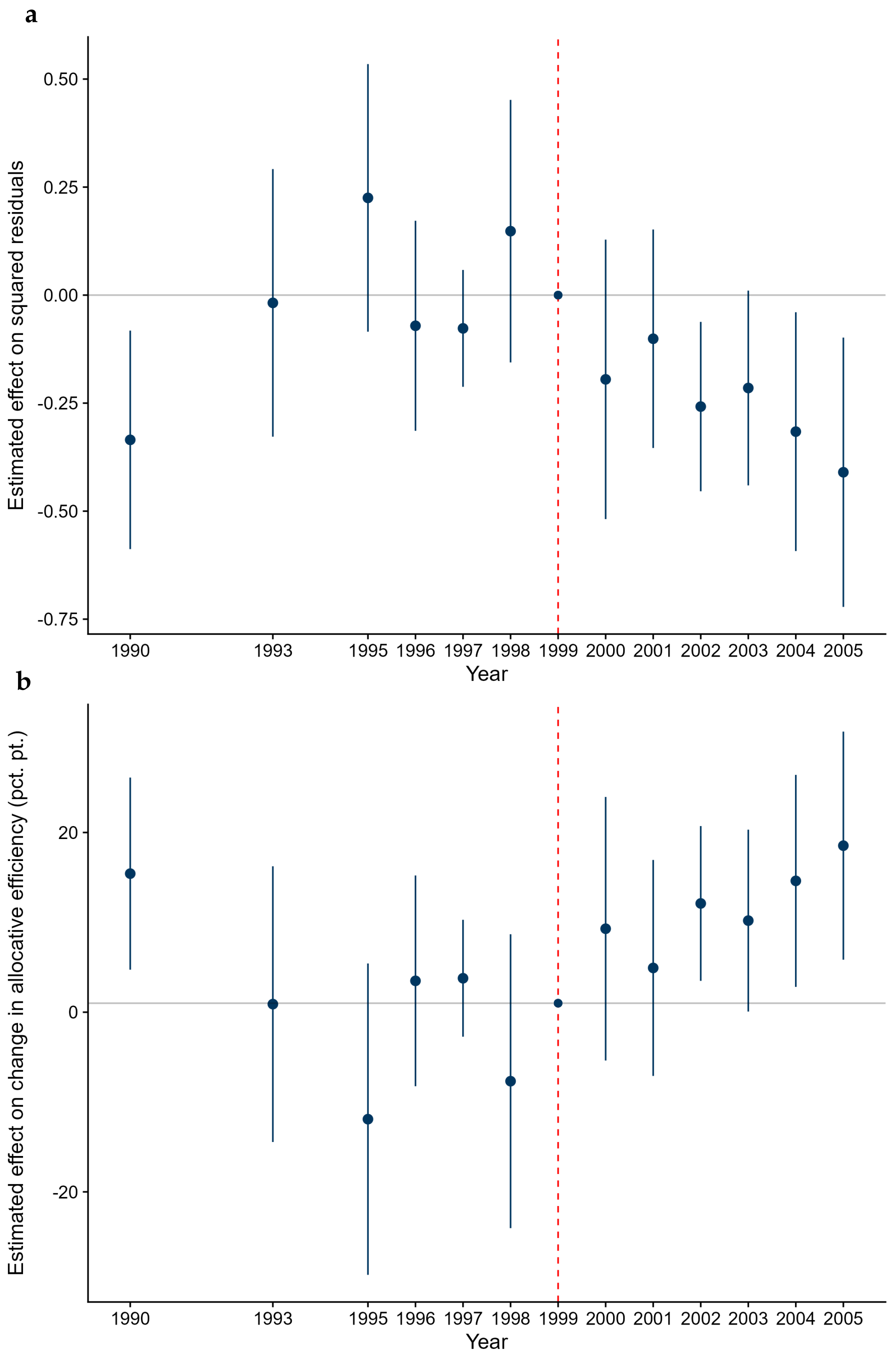
What is driving this decrease in dispersion at the facility level? Returning to our illustrative two-firm example in Figure 2, an improvement in overall allocative efficiency should have previously high-distortion facilities experience a lower distortion and thus higher emissions following the market introduction, and vice versa for previously low-distortion facilities. Following Bau and Matray (2023), Table 1 conducts this test by augmenting our difference-in-differences specification with an indicator variable for whether a facility has below-median pre-RECLAIM residual where the median is defined across regulated facilities and pre-treatment years. Column (1) shows that RECLAIM facilities with above-median pre-RECLAIM residuals experience a drop in residuals after the program while RECLAIM facilities with below-median pre-RECLAIM residuals have increased residuals.[36] Column (2) indicates this is linked to emissions: facilities with above-median pre-RECLAIM residuals experience an emissions decrease while facilities with below-median pre-RECLAIM residuals have their emissions rise. Taken together, Figure 5 and Table 1 point to RECLAIM narrowing the dispersion of residuals for regulated facilities in a manner consistent with a more efficient allocation of emissions.

Turning now to our main allocative efficiency result, Panel (a) of Figure 6 plots the estimates $\hat{\beta}_\tau$ from eq. (10). Recall from Section 3.3 that our theory interprets $e^{\frac{\beta_\tau}{2}} = \frac{E[\phi_{mt}]}{E[\phi_{bt}]}$, which under Proposition 1 serves as a lower bound on our estimand $\frac{\theta_m}{\theta_b}$, the ratio of misallocation cost across policies. For a more straightforward interpretation, Panel (b) plots $(1 - e^{\frac{\hat{\beta}_\tau}{2}}) \times 100 = (1 - \frac{E[\phi_{mt}]}{E[\phi_{bt}]}) \times 100$, the lower bound percentage point change in allocative efficiency. RECLAIM consistently improved allocative efficiency following its introduction. By 2005, the improvement was 20 percentage points relative to 1999 as estimated by the differential trend break model in eq. (10′) in Column (4), Panel (a) of Table 2. This is an annual improvement of 3.3 percentage points. The difference-in-differences estimates of eq. (10″) indicate a 10 percentage point improvement on average in the six years after 1999.

Some RECLAIM-regulated facilities were owned by firms with facilities in the control group, which may result in a violation of Stable Unit Treatment Value Assumption (SUTVA) if such firms reallocate production to unregulated facilities in response to RECLAIM. As robustness check, Column (1) of Table A1 presents difference-in-differences estimates of eq. (10″) using a restricted sample of firms with facilities only in either

[36]Note that this is not a conclusive test that the program has narrowed the dispersion of distortions. Consider a two firm scenario. The result in Column 1 is consistent with a firm with above-median pre-program residual and a firm with below-median pre-program distortion flipping positions with no change in the dispersion of distortions.

Figure 6: Event-study estimates of RECLAIM effect on allocative efficiency



Notes: Panel (a) shows point estimates and 95% confidence intervals of the yearly effect of RECLAIM on squared residuals relative to 1999, or $\widehat{\beta}_\tau$ using eq. (10). Bottom panel shows the implied percentage point change in allocative efficiency, or $(1 - e^{\frac{\hat{\beta}_\tau}{2}}) \times 100$. Standard errors are clustered at the zip code level.

treated or control groups. We find a similar improvement in allocative efficiency using this subsample of firms that cannot reallocate production across treated and control facilities.

Table 3 shows separate estimates for the percentage point change in allocative efficiency for each 2-digit SIC manufacturing sector by interacting our treatment variable with industry indicators in our difference-in-differences specification in eq. 10″. There are allocative efficiency gains in each sector, though they are notably stronger for petroleum refineries and metal manufacturing facilities.

Table 2: RECLAIM effects

	(1) Log emissions	(2) Log emissions	(3) Log revenue per emission	(4) Squared residual
	Panel (a): Trend-break model			
RECLAIM X Post	-0.016	0.022	-0.070	-0.168
	(0.044)	(0.060)	(0.074)	(0.151)
RECLAIM X Trend	-0.010	0.032^{**}	-0.034^{**}	0.033^{*}
	(0.013)	(0.014)	(0.014)	(0.017)
RECLAIM X Post X Trend	-0.034	-0.106^{***}	0.133^{***}	-0.080^{*}
	(0.024)	(0.026)	(0.028)	(0.044)
RECLAIM effect in 2005	-0.280^{***}	-0.420^{***}	0.522^{***}	-0.445^{***}
	(0.095)	(0.110)	(0.117)	(0.136)
Allocative efficiency chg in 2005 (pct. pt)				19.95 [7.18, 30.96]
	Panel (b): D-i-D model			
RECLAIM X Post	-0.182^{***}	-0.116^{*}	0.142^{*}	-0.215^{**}
	(0.049)	(0.062)	(0.073)	(0.092)
Allocative efficiency chg in 2005 (pct. pt)				10.19 [1.72, 17.94]
Sample	CARB	Matched	Matched	Matched
Observations	27,000	11,500	11,500	11,500

Notes: Panel (a) shows differential pre-trend estimates for log emissions (Columns (1) and (2)), log average revenue per emissions (Column 3 from eq. 9′), and predicted residuals (Column 4 from eq. 10′). Modeled predicted effect for 2005 shown. Column (4) also shows the implied percentage change in allocative efficiency in 2005, or $(1 - e^{\frac{\hat{\beta}}{2}}) \times 100 = (1 - \frac{\widehat{E[\phi_m]}}{\widehat{E[\phi_b]}}) \times 100$. All models include facility-specific and year-specific dummy variables. Column (1) uses the full CARB sample of manufacturing facilities and the CARB facility identifier for facility fixed effects. Columns (2)-(4) uses the matched CARB-ASMCM sample and the LBD facility identifier for facility fixed effects. Panel (b) shows analogous difference-in-differences estimates eqs. 9″ and 10″. Robust standard errors clustered at the zip code in parentheses, and 95% confidence interval in brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

Table 3: Allocative efficiency effect of RECLAIM by industry

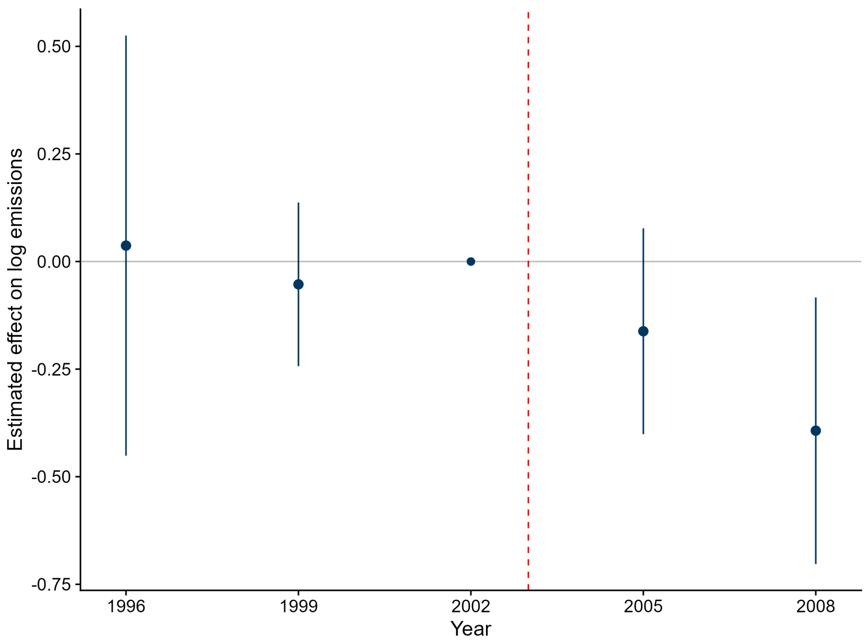
Industry	$(1 - e^{\frac{\hat{\beta}}{2}}) \times 100$	95% CI
Petroleum refineries (SIC 29)	17.1	[4.4, 28.1]
Primary metal manufacturing (SIC 33)	14.3	[7, 21.1]
Other manufacturing	8.3	[-3.5, 18.7]
Cement and glass manufacturing (SIC 32)	7.6	[-0.3, 14.9]
Secondary metal manufacturing (SIC 34)	6.2	[-13.1, 22.1]
Food manufacturing (SIC 20)	3.5	[-9.3, 14.9]

Notes: Industry-specific point estimates and 95% confidence interval of percentage point change in allocative efficiency effect, $(1 - e^{\frac{\hat{\beta}}{2}}) \times 100 = (1 - \frac{\widehat{E[\phi_m]}}{\widehat{E[\phi_b]}}) \times 100$, from difference-in-differences model in eq. 10″. Robust standard errors are clustered at the zip code level.

5.2 NOx Budget Trading Program

This section presents our results for the NBP market. Figure 7 shows the NBP effects on NO_x emissions applying the event study specification in eq. (9).[37] Manufacturing emissions fell consistently over the five years that the NBP was implemented. We do not detect pre-trends. A differential trend-break model shown in Column (1), Panel (a) of Table 5 find that emissions were 31% (0.375 log points) lower in 2008, five years after the program's introduction. The difference-in-differences estimate in Column (2), Panel (b) of Table 5 indicate a drop of 23% (0.257 log points) on average between 2003 to 2008. For comparison, Deschenes, Greenstone and Shapiro (2017) find that the NBP lowered electricity power plant NO_x emissions by 30% (0.36 log points).

Figure 7: Event-study estimates of the NBP effect on emissions



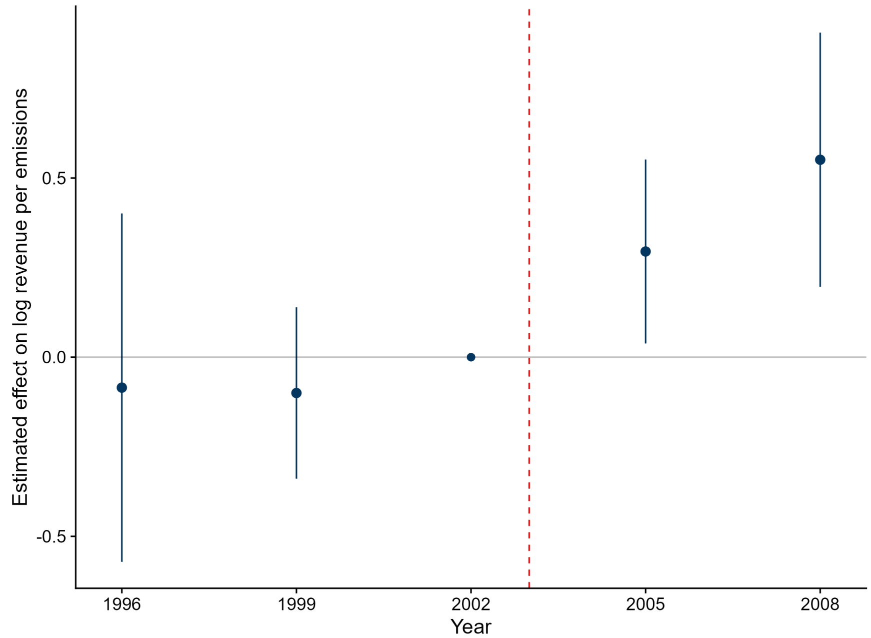
Notes: Point estimates and 95% confidence intervals of the yearly effect of the NBP on log NO_x emissions relative to 2002 using eq. (9). Standard errors are clustered at the county level.

Figure 8 shows the NBP effect on revenue per emissions using eq. (9), α_τ, which our theory interprets as the NBP effect on the aggregate NO_x (shadow) price. The NBP increased the NO_x price over its five years, with no pre-trend. Differential trend-break estimates of eq. (9′) in Column (2), Panel (a) of Table 5 indicate that the aggregate NO_x (shadow) price was 76% (0.57 log points) higher in 2008 relative to 2003. Difference-in-differences estimates of eq. (9″) shown in Column (2), Panel (b) of Table 5 find that the NBP increased the aggregate price by 60% (0.47 log points) on average between 2003 and 2008. By comparison, using industry-level data and a triple-difference research design applied to a model-based measure of the aggregate NO_x price, Shapiro and Walker (2018) find that the NBP increased the aggregate NO_x price for regulated manufacturing facilities by 1.2 log points.

Turning to residuals, Figure 9 plots the annual variance of residuals for treated and control facilities. Unlike with RECLAIM, we do not detect a divergence in residual variances between treated and control

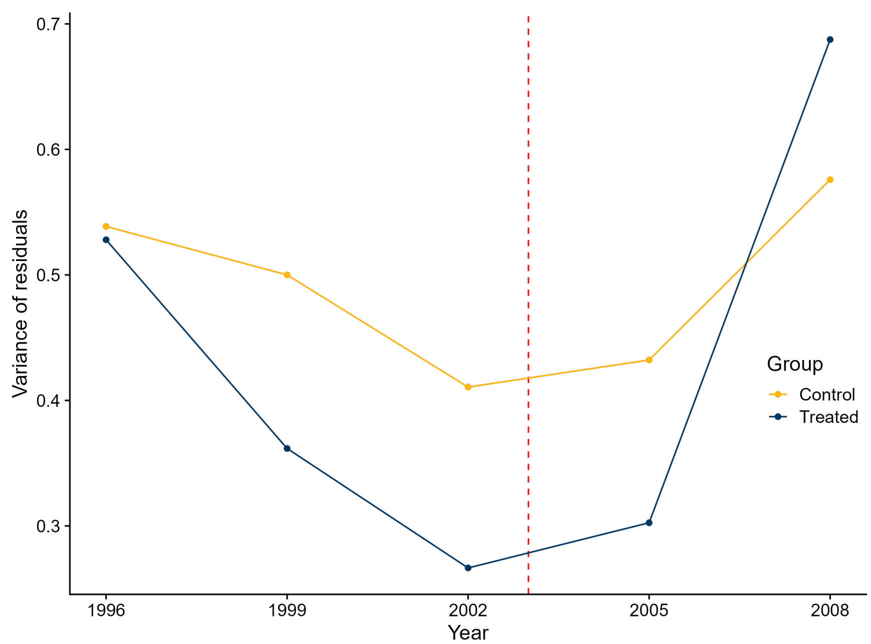
[37] Unlike with RECLAIM we cannot estimate an emissions effect on the full NEI dataset and compare that with estimates from the NEI-ASMCM matched dataset. This is because there is no unique time-invariant facility identifier in the NEI dataset. That identifier comes from the ASMCM.

Figure 8: Event-study model of the NBP effect on revenue per emissions



Notes: Point estimates and 95% confidence intervals of the yearly effect of the NBP on log revenue per emissions relative to 2002, or $\widehat{\alpha}_\tau$ using eq. (9). Standard errors are clustered at the county level.

Figure 9: Variance in predicted residuals for treated and control facilities under the NBP



Notes: Annual variance of the predicted residual, $\widehat{v}_{it}$, from eq. 9 for treated (blue) and control facilities (gold).

facilities following the NBP's introduction. Nor do we detect see a clear, statistically-precise pattern in the pre-NBP residual heterogeneity estimates in Table 4. Indeed, it shows residuals falling for both above- and below-median pre-NBP residual facilities.

Table 4: Heterogeneity in NBP effects by pre-treatment distortion

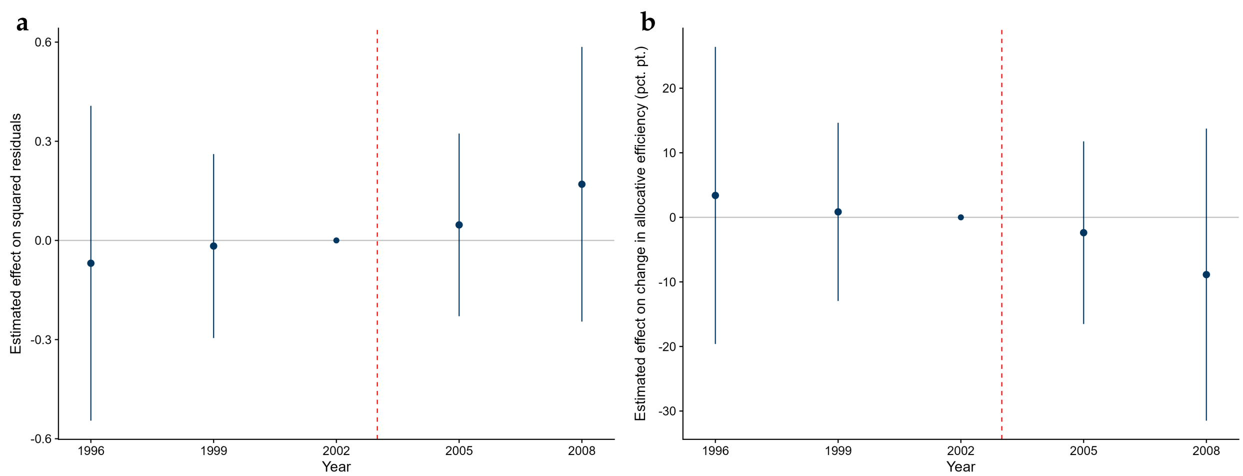
	(1) Residual	(2) Log emissions
NBP X Post	-4.696	0.009
	(5.291)	(0.074)
NBP X Post X Low distortion	1.848^{**}	-0.613^{**}
	(0.748)	(0.262)
Observations	32,500	32,500

Notes: Estimates of the differential effect of the NBP on residuals (Column 1) and log emissions (Column 2) by whether a NBP facility had below-median pre-NBP average residual. Specification based on eq. 9 and includes facility and year fixed effects. Robust standard errors clustered at the county level in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

These patterns are reflected in our estimate of the NBP's effect on allocative efficiency. Panel (a) of Figure 10 plots estimated coefficients $\hat{\beta}_\tau$ from eq. (10) or the effect of the NBP on the squared residual of average revenue per emissions. Coefficients before and after the NBP's introduction are not statistically significant. Panel (b) of Figure 10 presents the corresponding measure for the percentage point change in allocative efficiency, or $(1 - e^{\frac{\widehat{\beta_\tau}}{2}}) \times 100 = (1 - \frac{\widehat{E[\phi_m]}}{\widehat{E[\phi_b]}}) \times 100$, which also does not indicate a statistically precise change. This is reflected in the noisy trend-break and difference-in-differences estimates in Column (3), Panels (a) and (b) of Table 5.

Table 6 shows the estimated change in allocative efficiency for three different manufacturing sectors by re-estimating eq. (10) in which the treatment variable is interacted with industry indicators.[38] These results are statistically indistinguishable from zero in all sectors. In summary, the NBP lowered NO_x emissions, raised aggregate shadow emissions prices, but did not improve allocative efficiency.

Figure 10: Event-study estimates of the NBP effect on allocative efficiency



Notes: Top panel shows point estimates and 95% confidence intervals of the yearly effect of the NBP on squared residuals relative to 2002, or $\hat{\beta}_\tau$ using eq. (10). Bottom panel shows $(1 - e^{\frac{\hat{\beta}_\tau}{2}}) \times 100$. Standard errors are clustered at the county level.

[38]Due to disclosure requirements from the US Census Bureau on sample sizes, these three industry groups are the most disaggregated categories for which we could output results.

Table 5: NBP effects

	(1) Log emissions	(2) Log revenue per per emission	(3) Squared residual
		Panel (a): Trend-break model	
NBP X Post	0.090	0.056	-0.080
	(0.160)	(0.175)	(0.220)
NBP X Trend	-0.001	0.018	0.010
	(0.035)	(0.036)	(0.040)
NBP X Post X Trend	-0.076	0.067	0.031
	(0.051)	(0.051)	(0.065)
NBP effect in 2008	-0.375^{**}	0.565^{***}	0.166
	(0.154)	(0.183)	(0.203)
Allocative efficiency chg in 2008 (pct. pt)			-8.65 [-34.80, 12.42]
		Panel (b): D-i-D model	
NBP X Post	-0.257^{**}	0.471^{***}	0.124
	(0.123)	(0.151)	(0.134)
Allocative efficiency chg in 2008 (pct. pt)			-6.40 [-21.33, 6.70]
Observations	32,500	32,500	32,500

Notes: Panel (a) shows differential pre-trend estimates for log emissions (Column (1)), log average revenue per emissions (Column 2 from eq. 9′), and predicted residuals (Column 3 from eq. 10′). Modeled predicted effect for 2008 shown. Column (3) also shows the implied percentage change in allocative efficiency in 2008, or $(1 - e^{\frac{\hat{\beta}}{2}}) \times 100 = (1 - \frac{\widehat{E[\phi_m]}}{\widehat{E[\phi_b]}}) \times 100$. All models include facility-specific and year-specific dummy variables. Panel (b) shows analogous difference-in-differences estimates eqs. 9″ and 10″. Robust standard errors clustered at the county level in parentheses, and 95% confidence interval in brackets (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

5.3 Mechanisms

Why did RECLAIM improve allocative efficiency but the NBP did not? Where did allocative efficiency gains in RECLAIM come from? To shed light on these questions, we conduct a series of heterogeneity analyses across and within these markets.

Across markets We first focus on differences across RECLAIM and the NBP markets. In particular, we examine heterogeneity in proxy measures of facility-level marginal abatement costs. Intuitively, all things equal, a market-based policy covering facilities with greater pre-program heterogeneity in these proxies should induce larger efficiency gains.

Our investigation draws inspiration from Newell and Stavins (2003) which formally show that allocative efficiency gains from an environmental market depends on the coefficients of variation of marginal abatement cost curves. While not explicitly incorporated in their model, the authors highlight differences

Table 6: Allocative efficiency effect of the NBP by industry

Industry	$(1 - e^{\frac{\widehat{\beta}}{2}}) \times 100$	95% CI
Paper manufacturing (NAICS 322)	-6.7	[5.4, -20.2]
Refineries (NAICS 324 and 325)	-0.3	[12, -14.4]
Other manufacturing	-10	[11.6, -36.8]

Notes: Industry-specific point estimates and 95% confidence interval of percentage point change in allocative efficiency effect, $(1 - e^{\frac{\widehat{\beta}}{2}}) \times 100 = (1 - \frac{\widehat{E[\phi_m]}}{\widehat{E[\phi_b]}}) \times 100$, from difference-in-differences model in eq. 10''. Robust standard errors are clustered at the county level.

in location, age, size, and production technology as possible underlying characteristics that explain abatement cost heterogeneity.

Table 7 presents coefficients of variation (CV) of several such pre-program characteristics available in our data for regulated manufacturing facilities under RECLAIM and the NBP. This includes the pre-program CV of residuals from eq. 9, the total of value of shipment, NO_x emissions, total capital expenditures, total employment, facility age, pollution and abatement capital expenditures, and the number of unique NAICS-6 industries as a proxy for diversity of production technologies. Across these measures, Table 7 shows greater pre-program heterogeneity under RECLAIM than under the NBP, consistent with the allocative efficiency gains detected for RECLAIM but not for the NBP.

Table 7: Pre-program heterogeneity in facility characteristics

	$\widehat{v}_{it}$	TVS	NOx	Capital exp.	Emp.	Age	PACE	NAICS
RECLAIM	0.86	2.36	3.66	3.06	2.71	0.43	3.72	100
NBP	0.54	1.77	1.21	1.63	1.95	0.33	2.52	30
Statistic	CV	CV	CV	CV	CV	CV	CV	Count

Notes: The first seven columns show pre-program coefficients of variation across different regulated facility characteristics. $\widehat{v}_{it}$ are the predicted residuals from eq. 9. TVS = Total value of shipment. Capital exp. = Capital expenditures. Emp. = Total employment. PACE = Pollution and abatement capital expenditures. The last column shows the unique number of 6-digit NAICS industries. For RECLAIM, the pre-treatment period is before 2000. The NBP, the pre-treatment period is before 2003.

Other qualitative differences between RECLAIM and the NBP could also explain their different efficiency consequences. We speculate on three possible explanations. First, while RECLAIM replaced previous command-and-control regulations, the NBP was overlaid on top of pre-existing command-and-control regulations. Insofar as those regulations continued to bind for NBP-regulated facilities, improvements in allocative efficiency may be limited. Second, the NBP was a summer only pollution market which may have led to limited abatement options – and thus allocative efficiency gains – undertaken by regulated facilities that had to comply with command-and-control regulations during non-summer months. Third, RECLAIM cover all on-site facility emissions while the NBP covered only large boiler emissions. This allowed a larger set of pollution abatement options under RECLAIM than under the NBP.

Within market We now turn to heterogeneity analysis within each market. Unfortunately, given the relatively few number of treated facilities in each sample, we can report only a handful of heterogeneity analyses that meet U.S. Census Bureau disclosure requirements. Those analyses are shown in Table 8.

Firms with multiple regulated facilities may be able to shift production, and thus emissions, across

Table 8: Within market heterogeneity estimates

	(1)	(2)	(3)
	Squared residual		
	Panel (a): RECLAIM		
RECLAIM X Post	-0.187^{**}	-0.207^{**}	
	(0.087)	(0.097)	
RECLAIM X Post X Multi-facility firm	-0.044		
	(0.080)		
RECLAIM X Post X Inland		-0.040	
		(0.081)	
	Panel (b): NBP		
NBP X Post			0.160
			(0.238)
NBP X Post X CAIR			-0.054
			(0.242)
Observations	11,500	11,500	32,500

Notes: Estimates of the effect of RECLAIM (Columns 1 and 2) or the NBP (Column 3) on the dispersion of distortions based on eq. 10″. Column (1) adds an interaction for RECLAIM facilities owned by a multi-facility firm. Column (2) adds an interaction term for RECLAIM facilities in inland counties. Column (3) adds an interaction term for NBP facilities that would later be covered under CAIR. Robust standard errors clustered at the zip code (Column (1) and (2)) or county level (Column (3)) in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).

their facilities in response to a market-based policy (Fowlie, Holland and Mansur, 2012; Gibson, 2019; Cui and Moschini, 2020). This added compliance flexibility for facilities owned by multi-facility firms suggests greater allocative efficiency gains for such facilities than for single-firm facilities under an environmental market. For RECLAIM, Column (1), Panel (a) of Table 8 augments our difference-in-differences effect in eq. (10″) with an interaction term for facilities owned by multi-facility firms.[39] We find these facilities experience a 2 percentage point higher improvement in allocative efficiency than facilities owned by single-facilities firms, though this difference is not statistically significant.[40] In Column (2) of Table A1, we find a similar effect for facilities in multi-plant firms when we restrict the sample to facilities owned by firms that operate facilities exclusively in either the treated or control group, which mitigates SUTVA violation concerns from firms that may shift production across treated and control facilities.

Trading frictions can also come from policy design. Under RECLAIM, pollution hotspot concerns led to greater trading restrictions for coastal than inland facilities. Column (2), Panel (a) of Table 8 shows that RECLAIM facilities in the inland zone experienced a 2 percentage point greater allocative efficiency improvement than coastal RECLAIM facilities, though this effect is not statistically significant.[41]

Finally, we consider policy durability. Policies believed by regulated facilities to be more durable may incentivize more abatement investments and thus allocative efficiency gains (Requate, 2005). This is consistent with our finding that allocative efficiency gains under RECLAIM improve over time; indeed, RECLAIM

[39] Unfortunately, all NBP-regulated manufacturing facilities are owned by multi-facility firms, and therefore this exercise cannot be conduct for NBP facilities.

[40] From Column (1), Panel (a) of Table 8, 2 percentage points is $[1 - e^{(-.187-.044)/2}] - [1 - e^{(-.187)/2}]$.

[41] From Column (2), Panel (a) of Table 8, 2 percentage points is $[1 - e^{(-.207-.04)/2}] - [1 - e^{(-.207)/2}]$.

continues to operate today. The NBP was not as durable. Over half of NBP facilities went on to be regulated by a new cap-and-trade market under the Clean Air Interstate Rule (CAIR), which was anticipated as early as 2005, while other facilities reverted solely to command-and-control regulation. In Column 3, Panel (b) of Table 8, we add an interaction term for NBP facilities that would later be covered under CAIR. These facilities experience an allocative efficiency gain compared to other facilities, though again, this effect is not statistically precise.

Taken together, these heterogeneity results – though all statistically imprecise – suggest there were greater allocative efficiency gains under fewer trading frictions and when the market-based policy is expected to be more durable.

6 Conclusion

Market-based interventions hold the promise of improving allocative efficiency. Pollution provides a classic example: the introduction of a pollution market can in theory efficiently allocate emissions across heterogeneous polluters, lowering the total cost of meeting an aggregate pollution target compared with more prescriptive regulations.

In this paper, we develop a framework for estimating the change in allocative efficiency across two arbitrary environmental policy regimes. We lean on a producer's first order condition to relate its observed average revenue per emissions to its marginal product of emissions. We then show how a quasi-experimental estimator can recover a lower bound in the change in allocative efficiency following the introduction of a market-based policy. In contrast to prior approaches, our framework does not assume that a market-based policy necessarily improves allocative efficiency and does not require explicit estimation of facility-level marginal abatement cost curves.

We study the introduction of two landmark U.S. air pollution markets aimed at reducing NO_x emissions: southern California's Regional Clean Air Incentives Market (RECLAIM) and the eastern U.S. NO_x Budget Trading Program (NBP). We combine manufacturing facility emissions data from regional and national environmental agencies with restricted-use revenue data from the U.S. Census of Manufactures and the Annual Survey of Manufactures. Both programs lowered emissions. We find that RECLAIM improved allocative efficiency consistently after its cap binds at an annual rate of 3.3 percentage points. By contrast, we do not find evidence of an allocative efficiency change under NBP. These contrasting results are consistent with RECLAIM facilities exhibit greater pre-program heterogeneity across multiple characteristics. Within policies, we find suggestive evidence that allocative efficiency improves when there are fewer emissions trading frictions and when the market-based policy is expected to be durable.

Finally, our analysis adds to an emerging literature using quasi-experimental approaches to quantify the aggregate consequences of input misallocation. Here, our contribution is to develop a framework that can be applied to the introduction of environmental markets. Potential applications extend beyond environmental markets and include evaluating market-based approaches in other domains such as education, healthcare, and food provision.

References

Agarwal, Nikhil, Itai Ashlagi, Eduardo Azevedo, Clayton R. Featherstone, and Ömer Karaduman. 2019. "Market Failure in Kidney Exchange." *American Economic Review*, 109(11): 4026–70.

Anderson, Soren T., and James M. Sallee. 2011. "Using Loopholes to Reveal the Marginal Cost of Regulation: The Case of Fuel-Economy Standards." *American Economic Review*, 101(4): 1375–1409.

Aronoff, Daniel, and Will Rafey. 2023. "Conservation priorities and environmental offsets: Markets for Florida wetlands." National Bureau of Economic Research.

Arrow, Kenneth J. 1969. "THE ORGANIZATION OF ECONOMIC ACTIVITY: ISSUES." *The Analysis and Evaluation of Public Expenditures: the PPB System: pt. 1. The appropriate functions of government in an enterprise system. pt. 2. Institutional factors affecting efficient public expenditure policy. pt. 3. Some problems of analysis in evaluating public expenditure alternatives*, 1: 47.

Asker, John, Allan Collard-Wexler, and Jan De Loecker. 2014. "Dynamic inputs and resource (mis) allocation." *Journal of Political Economy*, 122(5): 1013–1063.

Ayres, Andrew B, Kyle C Meng, and Andrew J Plantinga. 2021. "Do environmental markets improve on open access? Evidence from California groundwater rights." *Journal of Political Economy*, 129(10): 2817–2860.

Banzhaf, H Spencer. 2020. "A history of pricing pollution (or, why Pigouvian taxes are not necessarily Pigouvian)." National Bureau of Economic Research.

Banzhaf, H Spencer. 2023. *Pricing the priceless: a history of environmental economics.* Cambridge University Press.

Baumol, William J, and Wallace E Oates. 1971. "The Use of Standards and Prices for Protection of the Environment." *Swedish Journal of Economics*, 73: 42–54.

Bau, Natalie, and Adrien Matray. 2023. "Misallocation and capital market integration: Evidence from India." *Econometrica*, 91(1): 67–106.

Berta, Nathalie. 2017. "On the definition of externality as a missing market." *The European Journal of the History of Economic Thought*, 24(2): 287–318.

Bils, Mark, Peter J Klenow, and Cian Ruane. 2021. "Misallocation or mismeasurement?" *Journal of Monetary Economics*, 124: S39–S56.

Bohi, Douglas R, and Dallas Burtraw. 1992. "Utility investment behavior and the emission trading market." *Resources and Energy*, 14(1-2): 129–153.

Calel, Raphael. 2020. "Adopt or Innovate: Understanding Technological Responses to Cap-and-Trade." *American Economic Journal: Economic Policy*, 12(3): 170–201.

Calel, Raphael, and Antoine Dechezleprêtre. 2016. "Environmental policy and directed technological change: evidence from the European carbon market." *Review of economics and statistics*, 98(1): 173–191.

Carlin, Alan. 1992. *United States Experience with Economic Incentives to Control Environmental Pollution.* US EPA, Office of Policy, Planning and Evaluation, Office of Policy Analysis.

Carlson, Curtis, Dallas Burtraw, Maureen Cropper, and Karen L Palmer. 2000. "Sulfur dioxide control by electric utilities: What are the gains from trade?" *Journal of Political Economy*, 108(6): 1292–1326.

Carter, Andrew V, Kevin T Schnepel, and Douglas G Steigerwald. 2017. "Asymptotic behavior of at-test robust to cluster heterogeneity." *Review of Economics and Statistics*, 99(4): 698–709.

Chan, H Ron, B Andrew Chupp, Maureen L Cropper, and Nicholas Z Muller. 2018. "The impact of trading on the costs and benefits of the Acid Rain Program." *Journal of Environmental Economics and Management*, 88: 180–209.

Cherniwchan, Jevan. 2017. "Trade liberalization and the environment: Evidence from NAFTA and US manufacturing." *Journal of International Economics*, 105: 130–149.

Chow, Melissa C, Teresa C Fort, Christopher Goetz, Nathan Goldschlag, James Lawrence, Elisabeth Ruth Perlman, Martha Stinson, and T Kirk White. 2021. "Redesigning the Longitudinal Business Database." National Bureau of Economic Research.

Coase, R. H. 1960. "The Problem of Social Cost." *Journal of Law and Economics*, 3: pp. 1–44.

Coggins, Jay S, and John R Swinton. 1996. "The price of pollution: a dual approach to valuing SO2 allowances." *Journal of environmental economics and management*, 30(1): 58–72.

Colmer, Jonathan, Ralf Martin, Mirabelle Muûls, and Ulrich J Wagner. 2022. "Does pricing carbon mitigate climate change? Firm-level evidence from the European Union emissions trading scheme."

Costello, Christopher, and Daniel Ovando. 2019. "Status, institutions, and prospects for global capture fisheries." *Annual Review of Environment and Resources*, 44(1): 177–200.

Crocker, T. 1966. "The structuring of atmospheric pollution control systems. The economics of air pollution." *The economics of air pollution. New York, WW Norton & Co*, 61–86.

Cuffe, John, and Nathan Goldschlag. 2018. "Squeezing more out of your data: Business record linkage with Python." Center for Economic Studies, US Census Bureau.

Cui, Jingbo, and GianCarlo Moschini. 2020. "Firm internal network, environmental regulation, and plant death." *Journal of Environmental Economics and Management*, 101: 102319.

Cui, Jingbo, Harvey Lapan, and GianCarlo Moschini. 2016. "Productivity, export, and environmental performance: air pollutants in the United States." *American Journal of Agricultural Economics*, 98(2): 447–467.

Curtis, E Mark. 2018. "Who loses under cap-and-trade programs? The labor market effects of the NOx budget trading program." *Review of Economics and Statistics*, 100(1): 151–166.

Dales, John H. 1968. *Pollution, Property and Prices: An Essay in Policy.* Toronto: University of Toronto Press.

Demsetz, Harold. 1969. "Information and efficiency: another viewpoint." *The journal of law and economics*, 12(1): 1–22.

DeSalvo, Bethany, Frank Limehouse, and Shawn D Klimek. 2016. "Documenting the business register and related economic business data." *US Census Bureau Center for Economic Studies Paper No. CES-WP-16-17*.

Deschenes, Olivier, Michael Greenstone, and Joseph S Shapiro. 2017. "Defensive investments and the demand for air quality: Evidence from the NOx budget program." *American Economic Review*, 107(10): 2958–89.

Epple, Dennis, Richard E. Romano, and Miguel Urquiola. 2017. "School Vouchers: A Survey of the Economics Literature." *Journal of Economic Literature*, 55(2): 441–92.

Färe, Rolf, Shawna Grosskopf, CA Knox Lovell, and Carl Pasurka. 1989. "Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach." *The review of economics and statistics*, 90–98.

Färe, Rolf, Shawna Grosskopf, CA Knox Lovell, and Suthathip Yaisawarng. 1993. "Derivation of shadow prices for undesirable outputs: a distance function approach." *The review of economics and statistics*, 374–380.

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

Fowlie, Meredith, and Jeffrey M Perloff. 2013. "Distributing pollution rights in cap-and-trade programs: are outcomes independent of allocation?" *Review of Economics and Statistics*, 95(5): 1640–1652.

Fowlie, Meredith, Christopher R Knittel, and Catherine Wolfram. 2012. "Sacred cars? Cost-effective regulation of stationary and nonstationary pollution sources." *American Economic Journal: Economic Policy*, 4(1): 98–126.

Fowlie, Meredith, Stephen P. Holland, and Erin T. Mansur. 2012. "What Do Emissions Markets Deliver and to Whom? Evidence from Southern California's NO_x Trading Program." *American Economic Review*, 102(2): 965–93.

Fullerton, Don, and Gilbert E Metcalf. 2001. "Environmental controls, scarcity rents, and pre-existing distortions." *Journal of public economics*, 80(2): 249–267.

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

Godby, Robert. 2002. "Market power in laboratory emission permit markets." *Environmental and Resource Economics*, 23: 279–318.

Gollop, Frank M, and Mark J Roberts. 1983. "Environmental regulations and productivity growth: The case of fossil-fueled electric power generation." *Journal of political Economy*, 91(4): 654–674.

Gollop, Frank M, and Mark J Roberts. 1985. "Cost-minimizing regulation of sulfur emissions: Regional gains in electric power." *The Review of Economics and Statistics*, 81–90.

Goulder, Lawrence H, Ian WH Parry, Roberton C Williams Iii, and Dallas Burtraw. 1999. "The cost-effectiveness of alternative instruments for environmental protection in a second-best setting." *Journal of public Economics*, 72(3): 329–360.

Grainger, Corbett, and Thanicha Ruangmas. 2018. "Who wins from emissions trading? Evidence from California." *Environmental and resource economics*, 71(3): 703–727.

Greenstone, Michael, Rohini Pande, Nicholas Ryan, and Anant Sudarshan. 2025. "Can pollution markets work in developing countries? Experimental evidence from India." *The Quarterly Journal of Economics*, 140(2): 1003–1060.

Hagerty, Nick. 2024. "Transaction Costs and the Gains from Trade in Water Markets." *Unpublished, Working Paper*.

Hahn, Robert W. 1984. "Market Power and Transferable Property Rights." *The Quarterly Journal of Economics*, 99(4): 753–765.

Haltiwanger, John, Robert Kulick, and Chad Syverson. 2018. "Misallocation Measures: The Distortion That Ate the Residual." National Bureau of Economic Research Working Paper 24199.

Harstad, Bård, and Gunnar S Eskeland. 2010. "Trading for the future: Signaling in permit markets." *Journal of public economics*, 94(9-10): 749–760.

Hernandez-Cortes, Danae, and Kyle C Meng. 2022. "Do Environmental Markets Cause Environmental Injustice? Evidence from California's Carbon Market." National Bureau of Economic Research Working Paper 27205.

Hottman, Colin J, Stephen J Redding, and David E Weinstein. 2016. "Quantifying the sources of firm heterogeneity." *The Quarterly Journal of Economics*, 131(3): 1291–1364.

Hsieh, Chang-Tai, and Peter J Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly journal of economics*, 124(4): 1403–1448.

Keohane, Nathaniel O. 2006. "Cost savings from allowance trading in the 1990 Clean Air Act: Estimates from a choice-based model." *Moving to markets in environmental regulation: Lessons from twenty years of experience*, 19: 23.

Kim, Seho. 2025. "Optimal Carbon Taxes and Misallocation across Heterogeneous Firms."

Klenow, Peter J., Ernesto Pasten, and Cian Ruane. 2024. "Carbon Taxes and Misallocation in Chile."

Kneese, Allen V. 1964. *The economics of regional water quality management.* Johns Hopkins Press for Resources for the Future.

Ladd, Helen F. 2002. "School vouchers: A critical view." *Journal of economic perspectives*, 16(4): 3–24.

Lange, Ian, and Peter Maniloff. 2021. "Updating allowance allocations in cap-and-trade: Evidence from the NOx Budget Program." *Journal of Environmental Economics and Management*, 105: 102380.

Lawley, Chad, and Vincent Thivierge. 2018. "Refining the evidence: British Columbia's carbon tax and household gasoline consumption." *The Energy Journal*, 39(2).

Linn, Joshua. 2008. "Technological modifications in the nitrogen oxides tradable permit program." *The Energy Journal*, 29(3).

Malik, Arun S. 1990. “Markets for pollution control when firms are noncompliant.” *Journal of Environmental Economics and management*, 18(2): 97–106.

Malueg, David A. 1990. “Welfare consequences of emission credit trading programs.” *Journal of Environmental Economics and management*, 18(1): 66–77.

Mansur, Erin T, and Glenn Sheriff. 2021. “On the measurement of environmental inequality: Ranking emissions distributions generated by different policy instruments.” *Journal of the Association of Environmental and Resource Economists*, 8(4): 721–758.

Massey, Catherine G, and Amy O’Hara. 2014. “Person Matching in Historical Files using the Census Bureau’s Person Validation System.” Center for Economic Studies, US Census Bureau.

Meng, Kyle C. 2017. “Using a Free Permit Rule to Forecast the Marginal Abatement Cost of Proposed Climate Policy.” *American Economic Review*, 107(3): 748–84.

Milgrom, Paul, and Ilya Segal. 2020. “Clock auctions and radio spectrum reallocation.” *Journal of Political Economy*, 128(1): 1–31.

Montgomery, W. David. 1972. “Markets in Licenses and Efficient Pollution Control Programs.” *Journal of Economic Theory*, 5(3): 395 – 418.

Murray, Brian, and Nicholas Rivers. 2015. “British Columbia’s revenue-neutral carbon tax: A review of the latest “grand experiment” in environmental policy.” *Energy Policy*, 86: 674–683.

Nevo, Aviv. 2001. “Measuring market power in the ready-to-eat cereal industry.” *Econometrica*, 69(2): 307–342.

Newell, Richard G, and Robert N Stavins. 2003. “Cost heterogeneity and the potential savings from market-based policies.” *Journal of Regulatory Economics*, 23: 43–59.

Petrick, Sebastian, and Ulrich J Wagner. 2014. “The impact of carbon trading on industry: Evidence from German manufacturing firms.” *Available at SSRN 2389800*.

Popp, David. 2010. “Exploring links between innovation and diffusion: adoption of NOx control technologies at US coal-fired power plants.” *Environmental and Resource Economics*, 45(3): 319–352.

Prendergast, Canice. 2022. “The allocation of food to food banks.” *Journal of Political Economy*, 130(8): 000–000.

Rafey, Will. 2023. “Droughts, deluges, and (river) diversions: Valuing market-based water reallocation.” *American Economic Review*, 113(2): 430–471.

Requate, Till. 2005. “Timing and commitment of environmental policy, adoption of new technology, and repercussions on R&D.” *Environmental and resource Economics*, 31: 175–199.

Restuccia, Diego, and Richard Rogerson. 2008. “Policy distortions and aggregate productivity with heterogeneous establishments.” *Review of Economic dynamics*, 11(4): 707–720.

Restuccia, Diego, and Richard Rogerson. 2013. “Misallocation and productivity.”

Restuccia, Diego, and Richard Rogerson. 2017. "The Causes and Costs of Misallocation." *Journal of Economic Perspectives*, 31(3): 151–74.

Roth, Alvin E, Tayfun Sönmez, and M Utku Ünver. 2007. "Efficient kidney exchange: Coincidence of wants in markets with compatibility-based preferences." *American Economic Review*, 97(3): 828–851.

Salzman, James, Genevieve Bennett, Nathaniel Carroll, Allie Goldstein, and Michael Jenkins. 2018. "The Global Status and Trends of Payments for Ecosystem Services." *Nature Sustainability*, 1(3): 136–144.

Samuelson, Paul A. 1954. "The pure theory of public expenditure." *The review of economics and statistics*, 387–389.

Shapiro, Joseph S, and Reed Walker. 2018. "Why is pollution from US manufacturing declining? The roles of environmental regulation, productivity, and trade." *American Economic Review*, 108(12): 3814–54.

Shapiro, Joseph S, and Reed Walker. 2024. "Is Air Pollution Regulation Too Lenient? Evidence from US Offset Markets."

South Coast Air Quality Management District. 2005. "Rule 2012 - Protocol for Monitoring, Reporting, and Recordkeeping for Oxides of Nitrogen (NOx) Emissions: APPENDIX A."

South Coast Air Quality Management District. 2006. "Annual RECLAIM Audit Report for 2004 Compliance Year."

Sraer, David, and David Thesmar. 2023. "How to Use Natural Experiments to Estimate Misallocation." *American Economic Review*, 113(4): 906–38.

Stavins, Robert N. 1995. "Transaction Costs and Tradeable Permits." *Journal of Environmental Economics and Management*, 29(2): 133–148.

Swinton, John R. 2002. "The Potential for Cost Savings in the Sulfur Dioxide Allowance Market: Empirical Evidence from Florida." *Land Economics*, 78(3): 390–404.

Swinton, John R. 2004. "Phase I completed: an empirical assessment of the 1990 CAAA." *Environmental and Resource Economics*, 27(3): 227–246.

Tietenberg, Thomas. 2010*a*. *Emissions trading: principles and practice.* Routledge.

Tietenberg, Tom. 2010*b*. "Cap-and-trade: the evolution of an economic idea." *Agricultural and Resource Economics Review*, 39(3): 359–367.

U.S. Environmental Protection Agency. 2002. "An Evaluation of the South Coast Air Quality Management District's Regional Clean Air Incentives Market-Lessons in Environmental Markets and Innovation."

U.S. Environmental Protection Agency. 2006. "NOx Budget Trading Program: 2005 Program Compliance and Environmental Results."

U.S. Environmental Protection Agency. 2007. "Identification and Evaluation of Candidate Control Measures: Final Technical Support Document." Ozone Transport Commission.

U.S. Environmental Protection Agency. 2009. "The NOx Budget Trading Program: 2008 Emission, Compliance, and Market Analyses."

World Bank. 2021. "State and Trends of Carbon Pricing."

A Theory appendix

A.1 Proposition 1

It is obvious from eq. (5) that if there is allocative efficiency with $\phi_{is} = 1 \; \forall i$ then $\theta_s = 1$. To establish that $\theta_s = 1$ implies efficiency, we rewrite θ_s as

$$\begin{aligned} \theta_s &= NE[a_{is}\phi_{is}] \\ &= NE[a_{is}]E[\phi_{is}] + Ncov(a_{is}(\phi_{is}), \phi_{is}) \\ &= e^{\sigma_s^2/2} + Ncov(a_{is}(\phi_{is}), \phi_{is}) \end{aligned}$$

where the second line applies the covariance definition and the third line uses $\sum_i a_{is} = 1$ and $E[\phi_{is}] = e^{\sigma_s^2/2}$ when $\phi_{is} \sim \mathcal{LN}(0, \sigma_s^2)$ from Assumption 1. Observe that because $e^{\sigma_s^2/2} \geq 1$, $\theta_s = 1$ or $\frac{1-e^{\sigma_s^2/2}}{N} = cov(a_{is}(\phi_{is}), \phi_{is})$ is satisfied under two cases: (i) when $cov(a_{is}(\phi_{is}), \phi_{is}) < 0$ and $e^{\sigma_s^2/2} > 1$ and (ii) when $cov(a_{is}(\phi_{is}), \phi_{is}) = 0$ and $e^{\sigma_s^2/2} = 1$. Note that

$$\begin{aligned} cov(a_{is}(\phi_{is}), \phi_{is}) =& E[\phi_{is}a_{is}(\phi_{is})] - E[\phi_{is}]E[a_{is}(\phi_{is})] \\ =& E[\phi_{is}a_{is}(\phi_{is}) - e^{\sigma_s^2} * a_{is}(\phi_{is})] \\ =& E[(\phi_{is} - e^{\sigma_s^2})a_{is}(\phi_{is})] \\ =& E[(\phi_{is} - e^{\sigma_s^2})(a_{is}(\phi_{is}) - a_{is}(e^{\sigma_s^2})] \end{aligned}$$

where the final line follows from $E[(\phi_{is} - 1)a_{is}(\phi_{is})c]$=0 for any constant c, in this case $e^{\sigma_s^2}$. By Assumption 2a, a increasing in ϕ implies $(\phi_{is} - e^{\sigma_s^2})(a_{is}(\phi_{is}) - a_{is}(e^{\sigma_s^2})) \geq 0$ and thus $cov(a_{is}(\phi_{is}), \phi_{is}) \geq 0$. This rules out case (i). Next observe that $(\phi_{is} - e^{\sigma_s^2})(a_{is}(\phi_{is}) - a_{is}(e^{\sigma_s^2})) = 0$ only when $\phi_{is} = 1, \forall i$, in which case $\sigma_s^2 = 0$ and $e^{\sigma_s^2/2}$=1. Thus, we have case (ii) in which $cov(a_{is}(\phi_{is}), \phi_{is}) = 0$ and $e^{\sigma_s^2/2} = 1$, which implies $\phi_{is} = 1$ $\forall i$. That is, under Assumptions 1 and 2a, $\theta_s = 1$ implies allocative efficiency, or $\phi_{is} = 1 \; \forall i$, establishing Proposition 1a.

To show that θ_s is increasing in the dispersion of distortions, σ_s^2, we must establish that

$$\frac{d\theta_s}{d\sigma_s^2} = \frac{de^{\sigma_s^2/2}}{d\sigma_s^2} + N\frac{dcov(a_{is}(\phi_{is}), \phi_{is})}{d\sigma_s^2} > 0$$

where the first right hand side term is positive. Turning to the second right hand side term and dropping subscripts, it is sufficient to establish $\frac{dcov(a(\phi),\phi)}{dE[\phi]} > 0$. Observe that we can define $h(\phi) = m + u$ where $h()$ is increasing u is functionally independent of m, or $\phi = h^{-1}(m + u)$, allowing us to write

$$\frac{dcov(a(\phi), \phi)}{dE[\phi]} = \frac{dcov(a(\phi), \phi)}{dm}\frac{dm}{dE[\phi]}$$

where because $h()$ is increasing u is functionally independent of m, $E[\phi]$ is increasing in m and vice-versa. Thus establishing $\frac{dcov(a(\phi),\phi)}{dE[\phi]} > 0$ requires $\frac{dcov(a(\phi),\phi)}{dm} > 0$. Applying $cov(a(\phi), \phi) = E[\phi a(\phi)] - E[\phi]E[a(\phi)]$

and taking a derivative, we have

$$\begin{aligned}
\frac{dcov(a(\phi),\phi)}{dm} &= E\left[\frac{a(\phi)+\phi a'(\phi)}{h'(\phi)}\right] - E[a(\phi)]E\left[\frac{1}{h'(\phi)}\right] - E[\phi]E\left[\frac{a'(\phi)}{h'(\phi)}\right] \\
&= cov\left(a(\phi), \frac{1}{h'(\phi)}\right) + E[a(\phi)]E\left[\frac{1}{h'(\phi)}\right] + cov\left(\phi, \frac{a'(\phi)}{h'(\phi)}\right) \\
&+ E[\phi]E\left[\frac{a'(\phi)}{h'(\phi)}\right] - E[a(\phi)]E\left[\frac{1}{h'(\phi)}\right] - E[\phi]E\left[\frac{a'(\phi)}{h'(\phi)}\right] \\
&= cov\left(a(\phi), \frac{1}{h'(\phi)}\right) + cov\left(\phi, \frac{a'(\phi)}{h'(\phi)}\right) \\
&= cov(a(\phi),\phi) + cov(\phi, a'(\phi)\phi)
\end{aligned}$$

where ones applies the Leibniz integral rule and the inverse function theorem after the first equality and the covariance definition after the second equality. The last equality follows under Assumption 1, whereby a lognormal distribution implies $h() = ln()$ such that $h'(\phi) = 1/\phi$. By Assumption 2a, a increasing in ϕ implies $cov(a(\phi),\phi) \geq 0$. By Assumption 2b, $a'(\phi)\phi$ increasing in ϕ implies $cov(\phi, a'(\phi)\phi) \geq 0$. This implies $\frac{d\theta_s}{d\sigma_s^2} > 0$, establishing Proposition 1b.

A.2 Proposition 2

We expand $\frac{\theta_m}{\theta_b}$ from eq. (6)

$$\begin{aligned}
\frac{\theta_m}{\theta_b} =& \frac{N(E[a_{im}]E[\phi_{im}] + cov(\phi_{im}, a_{im}))}{N(E[a_{ib}]E[\phi_{ib}] + cov(\phi_{ib}, a_{ib}))} \\
=& \frac{E[\phi_{im}]}{E[\phi_{ib}]}\left[1 + N\left(\frac{cov(\phi_{im}, a_{im})}{E[\phi_{im}]} - \frac{cov(\phi_{ib}, a_{ib})}{E[\phi_{ib}]}\right)\right] + \mathcal{O}^2 \\
\approx& \frac{E[\phi_{im}]}{E[\phi_{ib}]}\left[1 + N\left(\underbrace{\frac{cov(\phi_{im}, a_{im})}{E[\phi_{im}]}}_{Z_m} - \underbrace{\frac{cov(\phi_{ib}, a_{ib})}{E[\phi_{ib}]}}_{Z_b}\right)\right]
\end{aligned} \tag{A.1}$$

where the first line applies the definition of a covariance; second line applies a Taylor expansion around $E[\phi_{im}]$ and $E[\phi_{ib}]$ and uses $\sum_i a_{is} = 1$; and the third line retains first order term of the Taylor series as an approximation.

To establish Proposition 2, we must demonstrate $dZ_s/dE[\phi_{is}] > 0$. When $\frac{E[\phi_{im}]}{E[\phi_{ib}]} < 1$ or $E[\phi_{im}] - E[\phi_{ib}] < 0$, having $Z_m - Z_b < 0$ implies $\frac{\theta_m}{\theta_b} < \frac{E[\phi_{im}]}{E[\phi_{ib}]}$ and thus $\frac{E[\phi_{im}]}{E[\phi_{ib}]}$ is a lower bound on the true allocative efficiency gain $\frac{\theta_m}{\theta_b}$. Conversely, when $\frac{E[\phi_{im}]}{E[\phi_{ib}]} > 1$ or $E[\phi_{im}] - E[\phi_{ib}] > 0$, having $Z_m - Z_b > 0$ implies $\frac{\theta_m}{\theta_b} > \frac{E[\phi_{im}]}{E[\phi_{ib}]}$ and so $\frac{E[\phi_{im}]}{E[\phi_{ib}]}$ is a lower bound on the true allocative efficiency loss $\frac{\theta_m}{\theta_b}$.

First, observe that we can define $h(\phi) = m + u$ where $h()$ is increasing u is functionally independent of m, or $\phi = h^{-1}(m + u)$, allowing us to write

$$\frac{dZ}{dE[\phi]} = \frac{dZ}{dm}\frac{dm}{dE[\phi]}$$

where because $h()$ is increasing u is functionally independent of m, $E[\phi]$ is increasing in m and vice-versa. Thus establishing $\frac{dZ}{dE[\phi]} > 0$ requires $\frac{dZ}{dm} > 0$. Rearranging Z_s into $Z_s E[\phi_{is}] = cov(\phi_{is}, a_{is}(\phi_i s)) = E[\phi_{is} a_{is}(\phi_{is})] - E[\phi_{is}]E[a_{is}(\phi_{is})]$, dropping subscripts, and taking a derivative, we have

$$\begin{aligned}
\frac{dZ}{dm}E[\phi] + ZE\left[\frac{1}{h'(\phi)}\right] &= E\left[\frac{a(\phi)+\phi a'(\phi)}{h'(\phi)}\right] - E[a(\phi)]E\left[\frac{1}{h'(\phi)}\right] - E[\phi]E\left[\frac{a'(\phi)}{h'(\phi)}\right] \\
&= cov\left(a(\phi), \frac{1}{h'(\phi)}\right) + E[a(\phi)]E\left[\frac{1}{h'(\phi)}\right] + cov\left(\phi, \frac{a'(\phi)}{h'(\phi)}\right) \\
&\quad + E[\phi]E\left[\frac{a'(\phi)}{h'(\phi)}\right] - E[a(\phi)]E\left[\frac{1}{h'(\phi)}\right] - E[\phi]E\left[\frac{a'(\phi)}{h'(\phi)}\right] \\
&= cov\left(a(\phi), \frac{1}{h'(\phi)}\right) + cov\left(\phi, \frac{a'(\phi)}{h'(\phi)}\right) \\
\Rightarrow \frac{dZ}{dm} &= \frac{1}{E[\phi]}\left(cov\left(a(\phi), \frac{1}{h'(\phi)}\right) + cov\left(\phi, \frac{a'(\phi)}{h'(\phi)}\right) - \frac{cov\left(\phi, a(\phi)\right)}{E[\phi]}E\left[\frac{1}{h'(\phi)}\right]\right)
\end{aligned}$$

where ones applies the Leibniz integral rule and the inverse function theorem after the first equality and the covariance definition after the second equality. The last line follows by rearranging terms and dividing by $E[\phi]$. Under Assumption 1, a lognormal distribution implies $h() = ln()$ such that $h'(\phi) = 1/\phi$. This implies

$$\frac{dZ}{dm} = \frac{cov(\phi, a'(\phi)\phi)}{E[\phi]}$$

Under Assumption 1, $E[\phi] > 0$, thus the sign of $\frac{dZ}{dm}$ has the sign of $cov(\phi, a'(\phi)\phi)$, which is positive if $a'(\phi)\phi$ is increasing in ϕ, or when $\frac{-\phi a''(\phi)}{a'(\phi)} < 1$ as required by Assumption 2b. This establishes Proposition 2. Observe that in a consumption setting, the condition $\frac{-\phi a''(\phi)}{a'(\phi)} < 1$ occurs with a constant relative risk aversion utility function with weakly positive utility values, which mirrors our setting in which abatement shares must also be weakly positive.

B Data appendix

Record linkage procedure

To match facilities over time across U.S. Census Bureau and pollution datasets, we use different combinations of non-unique identifiers consisting of facility name, facility address, industry classifiers, zip code, and FIPS county codes.

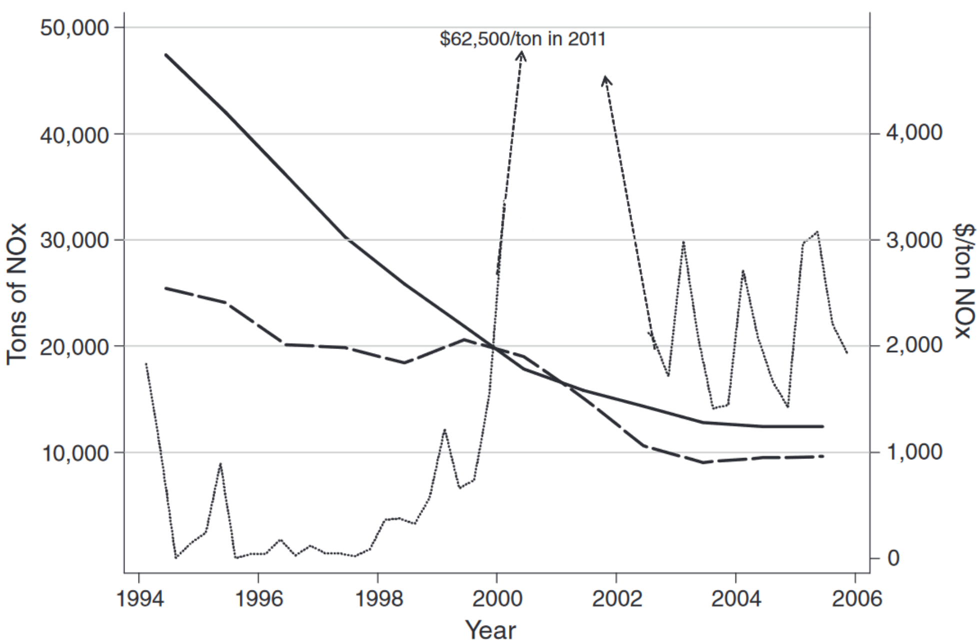
We first standardize facility names and addresses in both pollution and ASMCM data. For example, for facility names we clean and remove company suffixes such as CO and INC, and common expressions in facility names like USA, international, manufacturing, or industries. For addresses, we clean and remove street abbreviations and suffixes such as road, rd., st., avenue, ave., or boulevard. For both facility names and addresses, we also drop special characters.

In the second step, we iteratively block match our standardized data using different combinations of non-unique identifiers. Block matching reduces the numbers of comparisons made. For example, if we block on county FIPS code and 6-digit NAICS, then the names and addresses of refineries in Santa Barbara County in the CARB data are only matched to name and addresses of refineries in Santa Barbara County in the ASMCM data. Importantly, we do not block on matches on years. This allows us to account for variation in facility names, addresses, or other identifiers over time between facilities. Changes in facility name could reflect typographical error, but it could also reflect changes in ownership. Similarly, changes in industry classifier could be a consequence of spurious industry switching in the data, or could be legitimate industry switching documented as establishments respond to economic shocks (Chow et al., 2021).

After each matching iteration, we remove the uniquely matched facilities before moving to the next matching iteration. In the first iteration, we use the most stringent matching requirement by matching exactly on name, address, industry and geography. More than half of our matches come from this first iteration. In subsequent iterations, we block the data on different combinations of industry and geographic identifiers, and then exact or fuzzy match on facility name or facility address. To further ensure the quality of the matches, we manually review a subset of matches at each step.

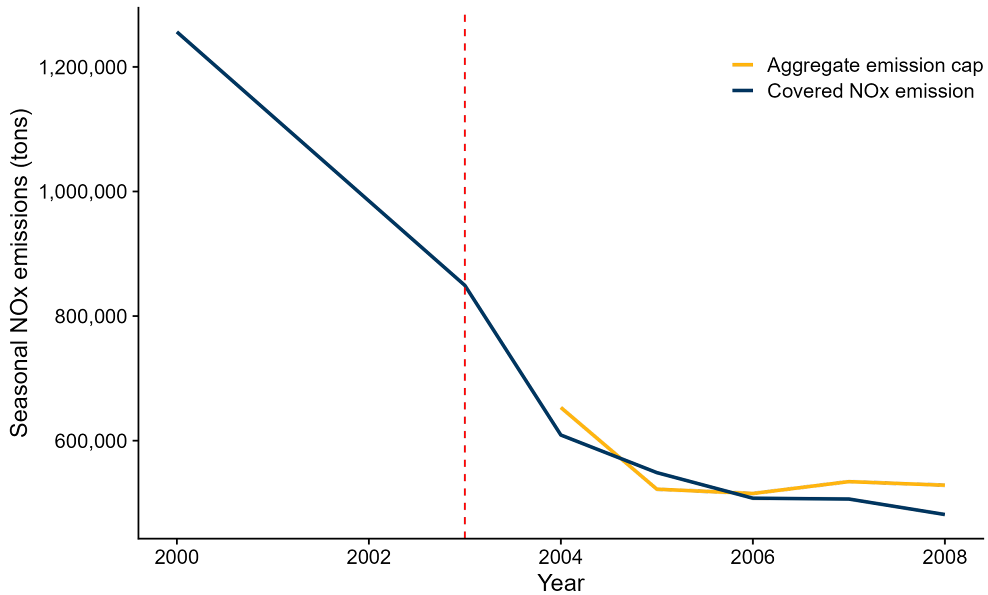
C Figure appendix

Figure A1: RECLAIM covered NO_x emissions, cap, and permit prices



Notes: Time series of RECLAIM NO_x aggregate covered emissions (dashed), aggregate emissions cap (solid) and permit prices (dotted), reproduced from Fowlie, Holland and Mansur (2012). There is a typographic error in the original Fowlie, Holland and Mansur (2012) figure. It should read ”in 2001”, and not ”in 2011”.

Figure A2: NBP covered NO_x emissions and cap



Notes: Time series of NBP NO_x aggregate covered emissions (blue) and aggregate emissions cap (yellow). The year 2003 cap is omitted from the graph since all final states joined the NBP in 2004.

D Table appendix

Table A1: RECLAIM effect by firm ownership

	(1)	(2)
	Squared residual	
RECLAIM X Post	-0.181^{**}	-0.128
	(0.081)	(0.080)
RECLAIM X Post X Multi-facility firm		-0.163
		(0.146)
Sample	Single group firms	Single group firms
Observations	9,500	9,500

Notes: Column (1) shows estimates of the effect of RECLAIM on the dispersion of distortions using eq. 10" for the sample of facilities owned by firms that operate only treated or control facilities. Column (2) further interacts the treatment variable with a dummy equal to one if regulated facility is owned by a multi-facility firm. All models include facility and year fixed effects. Robust standard errors clustered at the zip code level in parentheses (* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$).