What Do Emissions Markets Deliver and to Whom? Evidence from Southern California's NO_x Trading Program

Meredith Fowlie, Stephen P. Holland, and Erin T. Mansur*

March 28, 2011

Abstract

An advantage of cap-and-trade programs over more prescriptive environmental regulation is that compliance flexibility and cost effectiveness can make more stringent emissions reductions politically feasible. However, when markets (versus regulators) determine where emissions occur, it becomes more difficult to assure that mandated emissions reductions are equitably achieved. We investigate these issues in the context of Southern California's RECLAIM program by matching facilities in RECLAIM with similar California facilities also in non-attainment areas. Our results indicate that average emissions fell 20 percent at RECLAIM facilities relative to our counterfactual. Furthermore, observed changes in emissions do not vary significantly with neighborhood demographic characteristics.

JEL Classification: H23, Q52, D63, R20

Key Words: Environmental Regulation, Market Based Instruments, RECLAIM, Environmental Justice

---

* *Fowlie:* Department of Agricultural and Resource Economics, University of California, Berkeley 94720-3310 and NBER; email: fowlie@berkeley.edu. *Holland:* Department of Economics, University of North Carolina, Greensboro, NC 27402-6165, and NBER; email: sphollan@uncg.edu. *Mansur:* Department of Economics, Dartmouth College, 6106 Rockefeller Hall, Hanover, NH 03577, and NBER; email: erin.mansur@dartmouth.edu. We would like to thank Barbara Bamberger, Lucas Davis, John DiNardo, Justine Hastings, Matt Kahn, Patrick Kline, Justin McCrary, Mushfiq Mobarak, Manuel Pastor, Steven Redding, Randall Walsh, and seminar participants at the University of California Energy Institute, Yale University, University of North Carolina at Greensboro, Iowa State University, Harvard University, Dartmouth College, Texas A&M University, Camp Resources, POWER, University of Southern California, and the University of Pittsburgh for comments. We also thank Darryl Look at CA ARB and Danny Luong at SCAQMD for assistance with accessing data. Kate Foreman and Gray Kimbrough provided excellent research assistance. All authors thank the University of California Energy Institute for generous research support during this project.

1 Introduction

Policy makers have a variety of instruments at their disposal when pursuing emissions reduction objectives. Traditionally, regulators have relied upon "command and control" (CAC) approaches involving prescriptive emissions limits or pollution control technology standards. Increasingly, however, emissions trading programs are the preferred policy choice. In the United States, the Clean Air Act Amendments (CAAAs) of 1990 initiated a monumental shift away from CAC regulation towards more market-based alternatives such as emissions trading.[1] In parts of Europe, New Zealand, and regions of the United States, greenhouse gas regulations have helped to bring so called "cap-and-trade" to the fore.

Despite this prominence, questions remain about how emissions trading is working in practice. First, can these market-based programs reduce emissions beyond what could be achieved with more prescriptive CAC regulation? A perceived advantage of market-based approaches over CAC is that they can, in some circumstances, deliver more significant public health and environmental benefits because lower compliance costs and greater compliance flexibility make more stringent emissions reductions politically feasible (Keohane *et al.*, 1998; Ellerman, 2006; Tietenberg, 2006; US EPA, 1992). Although this hypothesis seems plausible, it has been difficult to test empirically (Ellerman, 2003; Harrington and Morgenstern, 2007; Stavins, 1998).

Second, some have expressed concern that a reliance on permit markets (versus pre-

[1] The CAAAs authorized the use of economic incentive regulation for the control of acid rain, the development of cleaner burning gasoline, the reduction of toxic air emissions, and for states to use in controlling carbon monoxide and urban ozone.

scriptive regulations and standards) to coordinate pollution abatement activity can lead to environmental injustice (Kaswan, 2008; Solomon and Lee, 2000; Vandenbergh and Ackerly, 2006).[2] If polluting facilities can achieve compliance by purchasing permits versus reducing emissions, there is the possibility that permitted pollution will flow into areas where poor or minority populations live. As detailed below, these environmental justice concerns are fueling heated opposition to emissions trading at the state and federal level.[3]

We examine these two issues in the context of a renowned emissions market: the REgional Clean Air Incentives Market (RECLAIM). Our primary objective is to identify the causal effects of this emissions trading program on facility-level emissions *vis a vis* the CAC regulations it replaced. Our essential challenge is to construct a credible benchmark; a precise and believable estimate of the emissions we would have observed in the absence of the program. Design features unique to RECLAIM facilitate the construction of this counterfactual. More specifically, we can exploit the fact that only a subset of industrial facilities located in non-attainment counties in California were removed from a CAC regime and required to participate in RECLAIM.

The RECLAIM program marked many firsts for emissions trading. It was the first mandatory trading program to supplant a pre-existing CAC regime that was, in theory, capable of achieving the same environmental objectives. It was the first program to include

[2] Although a broad literature examines environmental justice concerns with respect to plant siting, CAC regulation and neighborhood location choices (see, for example, Banzhaf and Walsh 2008), few papers assess the environmental justice effects of emissions trading.

[3] Among environmental justice advocates, concerns about greenhouse gas emissions trading pertain to co-pollutants. Whereas greenhouse gases are a global pollutant (damages do not depend on the spatial location of the source), co-pollutants, and associated damages, can be local.

a broad and diverse population of sources, making it particularly relevant to future trading programs which are likely to be more heterogeneous to achieve increasingly aggressive air quality and climate goals.[4] Illaudably, it was also the first emissions trading program to be challenged on the grounds of environmental injustice and noncompliance.

Our analysis of the RECLAIM program is motivated by three observations. First, a recent resurgence of interest in RECLAIM makes our study both timely and appropriate. Cap-and-trade programs have figured prominently in regional and federal proposals for addressing climate change, thus drawing increased attention to past experiences with market-based instruments in general, and RECLAIM in particular. Recent attempts to extract constructive insights from the RECLAIM experience have arrived at very different conclusions. Whereas some regard the program as a clear success (Stavins, 2007), others see a "spectacular" failure (Green *et al.*, 2007).[5]

Second, axiomatic questions about the effectiveness in reducing pollution of market-based programs relative to more traditional CAC regulations remain controversial and unresolved. Compared to the previous literature addressing these questions (see, for example, Harrington and Morgenstern, 2007), we take a fundamentally different approach.[6] We exploit the

[4] For example, under the auspices of A.B. 32, California is preparing to implement a greenhouse gas emissions trading program that covers large industrial sources of greenhouse gas emissions.

[5] Stavins (2007) summarizes domestic experience with emissions trading and reports that the RECLAIM program has generated significant environmental benefits "with NO_x emissions in the regulated area falling by 60 percent." Green *et al.* (2007) discuss the relative strengths and weaknesses of greenhouse gas emissions trading relative to a carbon tax. While reflecting upon past experiences with the former approach, they note that: "additional pitfalls and dilemmas of emissions trading can be seen through a review of the spectacular trading failure of the RECLAIM." They go on to argue that although "SCAQMD estimated that SO_2 and NO_x would be reduced by fourteen and eighty tons per day, respectively,...RECLAIM never came close to operating as predicted."

[6] Both Stavins (1998) and Ellerman (2003) note that, in the context of comprehensive cap-and-trade programs such as the Acid Rain Program, it has been difficult (if not impossible) to construct credible estimates of the emissions that would have been observed under a different regulatory regime. Harrington

participation requirements of the RECLAIM program in order to construct semi-parametric estimates of program impacts. Emissions trajectories at RECLAIM facilities are compared with those at similar California facilities that are exempt from RECLAIM. One important advantage of this approach is that it mitigates- or eliminates- the potentially confounding effects of changing economic conditions at the state-level, industry-wide production trends, and technological change.

Finally, our empirical framework facilitates an analysis of how RECLAIM-induced changes in emissions are distributed across communities with different socioeconomic characteristics. For a number of reasons, the RECLAIM market has been the most criticized of any emissions trading program with respect to environmental justice concerns. Some contend that RECLAIM has placed a disproportionate burden of the region's air pollution in low-income, minority communities (Drury *et al.*, 1999; Moore, 2004; Lejano and Hirose, 2005). We combine semi-parametric matching methods with parametric regression techniques. This allows us to examine correlations between RECLAIM-induced emissions changes and socioeconomic neighborhood characteristics with unprecedented precision.

Our results indicate that emissions at RECLAIM facilities have fallen by approximately 20 percent, on average, relative to control facilities (*i.e.*, similar California facilities that remained subject to command and control regulation over the duration of the study period). These results are robust to alternative estimation methods, functional form specifications,

et al. (2004) compare outcomes from controlling similar pollutants in the United States and Europe using different policy instruments. The limitation of this approach is that differences in outcomes across the two contexts likely reflect social, cultural, political, and economic differences, in addition to differences in regulatory regimes.

and different control group composition. Furthermore, we fail to reject the hypothesis that pollution reductions under RECLAIM were equally distributed across neighborhoods with different socio-economic characteristics.

The paper proceeds as follows. Section 2 provides background on Southern California's RECLAIM program, emphasizing past experiences with program evaluation and environmental justice issues in particular. Section 3 describes the research design and econometric approach. Section 4 summarizes the data. Section 5 presents the empirical findings. Section 6 concludes.

2 Background on RECLAIM

In this section, we introduce Southern California's RECLAIM program and provide some background on two areas of emphasis: the measurement of the emissions impacts of RECLAIM and related environmental justice concerns.

2.1 A Brief History of the Regional Clean Air Incentives Market

Los Angeles suffers from some of the worst air quality in the nation.[7] The South Coast Air Quality Management District (SCAQMD) is the government agency responsible for regulating air pollution in the Los Angeles basin. In 1989, SCAQMD introduced an aggressive set of rules and standards for stationary sources. Industry representatives fiercely opposed these

[7] Air pollution problems are due in part to meteorological and topographical conditions; the basin is sunny, warm, and poorly ventilated. The dense population, large number of vehicles, and high levels of industrial activity also contribute significantly to the problem. In 1988, ozone levels in the Los Angeles air basin exceeded state standards on 148 days (California Air Resources Board air quality data statistics accessed may 15, 2008. http://www.arb.ca.gov/adam/php_files/aqdphp/sc8start.php). Prior to the introduction of RECLAIM, estimates of health-related losses due to the poor environmental quality in the region were approaching $10 billion per year (Hall *et al.*, 1992).

rules on the grounds that compliance costs would prove excessive.

In 1990, Congress turned its attention to the widespread failure of US cities to attain health-based national ambient air quality standards (NAAQS). Under the 1990 CAAAs, Federal NO_x standards were significantly revised. Because SCAQMD was much further from attainment compared to other air basins, the district was given more time to comply. Although required reductions in ozone concentration levels were larger for the Los Angeles basin compared to other non-attainment areas in California, the required rates of concentration reductions over time were quite similar.[8]

The CAAAs also provided general authorization for states to use market-based regulatory programs to achieve federal standards. Market-based approaches to pollution regulation were endorsed on the grounds that CAC approaches were insufficient to address the worst of the nation's air quality problems, and that market-based approaches offered a "historic opportunity to help reconcile the nation's economic and environmental aspirations" (US EPA, 1992). While the use of economic incentives to achieve air quality standards was discretionary in most cases, it was required in extreme non-attainment areas, *i.e.*, Los Angeles.[9]

SCAQMD responded by replacing over 40 prescriptive rules, which had been so opposed by industry, with a market-based emissions trading program: RECLAIM.[10] This program was approved by state and federal regulators on the grounds that it would deliver emissions

[8] Section 5.3 characterizes the CAAA compliance requirements in more detail.

[9] Pursuant to Sections 182 and 187, the US EPA issued a final rule and guidance on Economic Incentive Programs (40, part 51, Subpart U) which outlined requirements for establishing EIPs. States or governing bodies in extreme ozone nonattainment areas were required to design and implement economic incentive programs $(51.492, 182(g)5)$.

[10] Although both NO_x and SO_2 emissions are capped under the program, the emphasis was on limiting NO_x emissions which are an important precursor to ozone formation.

reductions equivalent to–or greater than–what would have been achieved under the subsumed command-and-control provisions, and would help to bring the region into compliance with federal standards by the 2010 deadline.

At its inception in 1994, RECLAIM included 392 facilities whose combined NO_x emissions accounted for over 65% of the region's stationary NO_x emissions (Schubert and Zerlauth, 1999). Almost all facilities in the SCAQMD with annual NO_x or SO_2 emissions of four tons or more are included in the program.[11] Public facilities (such as police and fire fighting facilities) were categorically excluded. Sources emitting less than four tons per year remained subject to command-and-control programs.

A RECLAIM trading credit (RTC) confers the right to emit one pound of emissions within a twelve month period. At the outset of the program, facilities were informed of how many permits they would be allocated gratis each year through 2010.[12] These RTCs were distributed based on firms' historical fuel consumption and pre-determined production technology characteristics.[13] Figure 1 plots the aggregate allocation trajectory over time (the red line).[14] NO_x emissions permitted under RECLAIM were reduced by over 70 percent over the first ten years of the program. By the end of 2003, the aggregate permit allocation reached the level of emissions that the subsumed rules and control measures were intended

[11] Of these, 73% can be classified as manufacturing firms, 13% are involved in communication, transportation or utilities, 2% are involved in construction, 3% are operating in the service sector, 6% in wholesale trade, 2% are retail establishments, and the remaining 3% can be classified as government facilities.

[12] RTCs cannot be banked; a permit can only be used to certify emissions occurring within the twelve month period with which the permit is associated. For emissions in any quarter, firms can use either permits expiring in June or in December. See Holland and Moore (2008).

[13] The RTC allocation methodology is described in detail in SCAQMD Rule 2002 (SCAQMD, 1993).

[14] SCAQMD maintains a detailed database tracking all NO_x permits and quarterly, facility-level emissions. Section 4 includes a detailed description of these data.

to achieve by 2010.

Early on, most firms found they had an excess of credits (the blue line in Figure 1 represents aggregate tons of NO_x emissions). The aggregate cap did not start to bind until 1999 (SCAQMD, 2001). Figure 1 helps to illustrates this "cross-over" point. While it is clear that emissions permits were initially over-allocated, many believe that generous permit allocations in the early years of the program were necessary to engender political support for the program (US EPA, 2002).[15] Because permits cannot be banked, impacts of the initial over-allocation were confined to the early stages of RECLAIM.

Figure 1 also plots the trend in average RTC prices (the green line). In the first five years of the program, prices for NO_x RTCs remained relatively low, as expected.[16] However, the increase in prices following the cross-over was much larger than anticipated; the price of NO_x RTCs increased from approximately $2000 per ton in January of 2000 to over $120,000 per ton in March of 2001.

During the California electricity crisis, production levels at electricity generating facilities in the RECLAIM program increased significantly. Emissions at these facilities exceeded

[15] Nonetheless, RECLAIM may have changed firms' production and investment decisions in this early period. A firm making a long-lived investment may have abated early in anticipation of higher future prices. Furthermore, RECLAIM relaxed a vintage differentiated regulation, New Source Review, that has limited firms abilities to modify facilities. For example, only "BACT" is required and necessary offsets can be demonstrated with RTCs. RECLAIM annual reports show a very high rate of NSR activity. From 1994 to 2006, the reports show that on average forty-seven RECLAIM facilities had NSR activity per year. In contrast, Committee (2006) report that on average 125 NSR permits per year were issued for the entire country from 1997 to 2002 for NO_x.

[16] Before RECLAIM began, it was predicted that trading in the market would be slow at first because of the initial surplus of permits. In 1994, SCAQMD economists predicted that prices for NO_x RTCs would average around $577/ton in 1995 and rise to approximately $1,100/ton by 1999 (Miller, Michael.1994. "Firms Can Earn Credits for Keeping Emissions Down, Then Sell Them." *The San Francisco Examiner.* January 9, 1994: B1.

permit allocations, which in turn led to a sharp increase in RTC prices.[17] In May 2001, the RECLAIM rules were amended in an effort to stabilize the RTC market. The rule amendments (Rule 2009) removed fourteen power producers from the RECLAIM market. These facilities were required to pay a fee of $15,000 per ton of emissions in excess of their allocation. They were also required to install the "best available" control technologies on existing power generating units by the end of 2004.[18] In 2007, these large power producers re-entered the RECLAIM program as unrestricted market participants.

By 2002, monthly average prices had fallen below $2000 per ton NO_x. Regulators were concerned that low permit prices were failing to provide sufficient incentives for facilities to install pollution control technologies that would be needed to bring the region into compliance with federal standards. In September of 2004, restrictions on power producers were made more stringent and the aggregate RTC allocation for compliance years 2007-2011 was reduced by an additional 20 percent.

2.2 RECLAIM Program Evaluation

Because RECLAIM represented such a major departure from the traditional regulatory approach, both federal and state agencies required extensive program evaluation and oversight. Emissions trading program evaluation is particularly challenging. Because industrial emissions are influenced by numerous factors, attributing changes in emissions patterns to specific policy interventions is difficult. These challenges notwithstanding, agencies in charge of over-

[17] Kolstad and Wolak (2003) provide evidence that some electricity producers in SCAQMD intentionally purchased NO_x RTCs at higher than competitive prices so as to be able to artificially increase electricity prices.

[18] For more information see SCAQMD (2007).

seeing RECLAIM remain committed to evaluating the emissions impacts of the program.

Unresolved disagreements about what constitutes an appropriate measure of counterfactual emissions have resulted in a plurality of opinions regarding RECLAIM's overall performance. After fifteen years of program evaluations, the emissions impacts of RECLAIM *vis a vis* the subsumed CAC rules remain controversial. Federal policy makers and other stakeholders have expressed frustration over the lack of consensus emerging from RECLAIM program evaluations, noting that the public is entitled to "real world information and practical comparisons in order to judge for itself whether the program is living up to their needs and expectations" (US EPA, 2002). Appendix A summarizes some of the contradictory evidence provided by past program evaluations and reports.

2.3 Environmental Justice and Emissions Trading

The term "environmental injustice" refers to any disproportionate human health or environmental impact on minority or low income populations (EO 12898, 1994). Empirical research conducted in the 1980s demonstrated significantly higher levels of exposure to environmental hazards in traditionally disadvantaged communities.[19] Subsequent work has brought more sophisticated empirical methods to bear on this issue (Banzhaf and Walsh, 2008).

Kaswan (2008) provides a detailed discussion of the perceived tensions between environmental justice and emissions trading. A dominant concern is that emissions trading programs fail to account for the distribution of pollution damages whereas permitting under the CAAAs can explicitly consider environmental justice concerns. If polluting facilities

[19] See, for example, Brown, 1995; UCC, 1987; GAO, 1983.

can purchase permits instead of reducing emissions, it is possible for permitted pollution to flow into areas where poor or minority populations live, thereby exacerbating pre-existing inequalities in the distribution of environmental risks. On the other hand, market-based programs could mitigate pre-existing environmental justice problems. If polluting facilities with relatively low marginal abatement costs are disproportionately located in traditionally disadvantaged neighborhoods, an efficient permit market should ensure that a larger share of the mandated emissions reductions will be achieved in these areas (Burtraw *et al.*, 2005).[20]

For a number of reasons, the RECLAIM market has been the most scrutinized of any emissions trading program with respect to environmental justice issues (Chinn, 1999; Drury *et al.*, 1999; Lejano and Hirose, 2005; Moore, 2004). First, the Los Angeles area is home to an exceptionally diverse population. Past studies have documented that race and ethnicity have historically played a "persistent explanatory role" in explaining the distribution of environmental health risks in Southern California (Morello-Frosch *et al.*, 2001). Second, NO_x is a non-uniformly mixed pollutant. This means that there is potential for significant spatial variation in damages from NO_x emissions, and thus potential for environmental injustice.[21] Third, the RECLAIM program was indirectly implicated in another highly controversial rule promulgated by SCAQMD that allowed stationary sources to offset their uncontrolled emissions of volatile organic compounds (VOCs) using mobile source emissions reduction

[20] Past studies looking at the distributional impacts of emissions trading have looked closely at the Acid Rain Program. These studies find no evidence of disproportionately high and adverse human health or environmental effects on minority, low-income, or other populations (EPA, 2005; Shadbegian *et al.*, 2007).

[21] In the interest of avoiding "hotspots", RECLAIM was designed as a zonal trading system. The SCAQMD was divided into two zones: the region along the coast, and an inland region. Facilities along the coast (where pollution problems tend to be more severe) may only purchase RTCs from other coastal facilities. Inland facilities can purchase permits from either inland or coastal facilities.

credits.[22] Although the links between RECLAIM and this controversial rule were indirect and inconsequential, the RECLAIM program has since been associated with environmental injustice allegations.[23]

Concerns about environmental justice have strongly influenced the debate surrounding California's greenhouse gas regulations (Hanemann, 2008; Sze *et al.*, 2009). Regulatory activities in California under AB 32 constitute the most ambitious and comprehensive effort to control GHG emissions currently underway in the United States. Prominent environmental justice advocates have come out in strong opposition to cap-and-trade in California due to concerns about mercury, benzene, and other co-pollutants. They cite RECLAIM as a "well documented" example of how emissions trading can disproportionately harm communities of color (Drury, 2009).[24] Citizens groups filing a lawsuit in 2011 to prevent greenhouse gas emissions trading in California alleged that "All the evidence show(s) that cap-and-trade programs have failed environmental justice communities" (Sweet, 2011).

3 Research design

Previous estimates of the emissions effects of RECLAIM are conditional on, and highly sensitive to, controversial assumptions about what emissions would have been in the absence of

[22] This rule was challenged by a coalition of environmental groups on the grounds that it violated Title VI of the Civil Rights Act; the rule allowed reductions in mobile source emissions (whose effects are arguably distributed widely across the region) to be substituted for VOC reductions at point sources located in minority communities. The lawsuit was withdrawn by the plaintiffs two weeks after the case was filed. See "CBE Sues SCAQMD Over Amendments to Car Scrapping Rule", California Environmental Insider: 12 (7), Sept. 15, 1998.

[23] The RECLAIM program, as it was originally designed, permitted the use of mobile source credits to achieve compliance. This mobile source credit compliance option was rarely used. Mobile source credits represented less than 0.02% of the total allocation of NO_x permits.

[24] See Appendix C for a more detailed discussion of these arguments.

the program. In this study, we exploit some unique design features of the RECLAIM program in order to construct more tenable and transparent estimates of counterfactual emissions. Rather than rely on ex ante expectations about what aggregate emissions trajectories would have been absent RECLAIM, we use econometrically adjusted ex post observed emissions at facilities that were subject to CAC regulation over the same time period. In what follows, we introduce our empirical framework and identification strategy.

3.1 Empirical framework

Building on the potential outcome framework that is now standard in the program evaluation literature (see Holland (1986) for a survey), we postulate that there are two regulatory states to which California's industrial NO_x emitters could have been assigned: the market-based RECLAIM program or the CAC regime that prevails in non-attainment counties outside of SCAQMD (and which the SCAQMD continues to use to regulate smaller emitters). Let $D_i = 1$ if the i^{th} facility is included in RECLAIM (*i.e.*, the facility is "treated"). Let $D_i = 0$ if facility i remains subject to CAC regulation throughout the duration of our study. Potential outcomes $Y_{it}(1)$ and $Y_{it}(0)$ denote annual emissions at facility i at time t conditional on participation and non-participation, respectively.

We are primarily interested in estimating the sample average treatment effect on the treated (SATT):

$$\alpha_{TT} = E[Y_{it'}(1) - Y_{it'}(0) \mid D_i = 1], \tag{1}$$

where t' represents a year following the introduction of the RECLAIM program and α_{TT} measures the average effect of the RECLAIM program on annual facility level NO_x emissions.[25]

Emissions at both treated and untreated facilities are observed prior to the RECLAIM program (*i.e.*, when all facilities in California's non-attainment areas were subject to CAC regulation) and over several years following the introduction of the program. Facility-level emissions data collected from RECLAIM participants during years following the introduction of the program can be used to identify $E[Y_{it'}(1)|D_i = 1]$. However, $[Y_{it'}(0)|D_i = 1]$ is not observed. We will construct estimates of these counterfactual outcomes using emissions observed at control facilities subject to CAC regulation for the duration of the time period.

Incomplete program participation requirements provide us with two potential comparison groups. First, the RECLAIM program applies only to major sources located within SCAQMD. Thousands of California facilities located outside the Los Angeles air basin are subject to more traditional CAC. Second, hundreds of smaller emitters within SCAQMD remain subject to more traditional CAC rules.

The simplest and most naive estimate of α_{TT} is obtained by computing an unconditional differences-in-differences. This estimator will be biased if factors that are related to facility-level emissions dynamics vary significantly across the treatment and comparison groups. In order to reduce the bias potentially introduced by observable differences across RECLAIM participants and non-participants, we employ two strategies that condition on observable covariates.

[25] We will also evaluate program impacts in percentage terms, although we will emphasize [1] as a more informative measure of the average effect of RECLAIM on industry emissions.

3.1.1 Regression-based conditioning strategies

Ordinary least squares (OLS) can be used to control for factors other than regulatory regime that affect facility-level emissions trajectories. We estimate the following simple specification:

$$Y_{it'} - Y_{it^0} = \beta' X_i + \alpha D_i + \varepsilon_i, \tag{2}$$

where X_i is a vector of observable covariates and t^0 denotes the time period prior to the introduction of RECLAIM. This approach implicitly assumes that the variables in X are exogenous to treatment status. In our case, these variables will include facility level emissions before RECLAIM was introduced, four-digit industry classification, county-level attainment status, and pre-treatment, facility specific economic and demographic measures. The parameter α captures the average effect of the RECLAIM program on changes in facility-level emissions over time conditional on variables in X. The error term ε_i is assumed to be independent of the covariates in X_i and the treatment indicator D_i.

There are several potential problems with this approach. First, if there is only limited overlap in the distributions of X across the treatment and control groups and functional form assumptions are incorrect, missing outcomes will be incorrectly imputed. Estimates of average treatment effects can also be biased if control observations are not appropriately reweighted to control for differences in the distribution of the X variables over regions common to the control and treatment groups. In the interest of mitigating these potential biases, we turn to semi-parametric matching estimators.

3.1.2 Semi-parametric conditioning strategies

Matching estimators are an extension of standard regression approaches. One clear advantage is that parametric assumptions about the relationship between the outcome variable and the covariates in X can be avoided. Our general approach to matching follows Heckman *et al.* (1997, 1998) who introduce the following generalized DID matching estimator:

$$\widehat{\alpha_{DID}} = \frac{1}{N_1} \sum_{j \in I_1} \{(Y_{jt'}(1) - Y_{jt^0}(0)) - \sum_{k \in I_0} w_{jk}(Y_{kt'}(0) - Y_{kt^0}(0))\}. \quad (3)$$

Here, I_1 denotes the set of program participants, I_0 denotes the set of nonparticipants, and N_1 is the number of facilities in the treatment group. The participants are indexed by j; the non-participants are indexed by k. The weight placed on facility k when constructing the counterfactual estimate for treated facility j is w_{jk}. Our nearest neighbor matching estimator weights control facilities according to their similarity to treated facilities where similarity is based on X.

3.2 Identifying assumptions

Our most important identifying assumption is that the biases in the unconditional DID estimates can be removed by adjusting for differences in observable covariates. More formally, we assume that the distribution of the control outcome $Y_{it'}(0)$, conditional on observable facility and neighborhood characteristics (such as historic emissions, industry classification, county attainment status), is the same among participating and non-participating facilities. If this conditional unconfoundedness assumption is satisfied, once we adjust for observable differences, we can interpret differences in observed outcomes as the effect of RECLAIM versus the CAC regimes of other California air basins.

In our context, we also invoke a stronger variant of the unconfoundedness assumption. In order to interpret [3] as an estimate of the effect of RECLAIM on emissions vis a vis what would have been observed under the status quo, it must be that trends in the stringency of the control treatment (*i.e.*, the CAC regulations to which the control facilities are subjected) follow the trajectory that the SCAQMD CAC regime would have taken absent RECLAIM.

Our estimation strategy also requires that the support of the distribution of the conditioning covariates in the treatment group overlaps the support of the distribution of these covariates in the comparison group.

Finally, in order to rule out spillovers and general equilibrium effects, it must also be the case that potential outcomes at one facility are independent of the treatment status of other facilities. We refer to this subsequently as the stable unit treatment value assumption (SUTVA).

Some of these assumptions can be directly tested. For instance, it is straightforward to demonstrate that the overlap condition is satisfied by simply looking at the joint distributions of the covariates in the treated and control groups. Other assumptions, including unconfoundedness and SUTVA, are not testable in principle. However, we will conduct indirect tests in order to evaluate the plausibility of these assumptions.

3.3 Treatment effect heterogeneity

Thus far, we have been exclusively concerned with estimating the average effect of RECLAIM on facility-level emissions. We are also interested in investigating whether treatment effects vary systematically across facilities located in neighborhoods with different socio-economic characteristics. We estimate the following weighted regression:

$$Y_{it'} - Y_{it^0} = \delta_j + \beta' X_i + \theta' X_i D_i + \alpha D_i + \varepsilon_i, \tag{4}$$

where the δ_j are group specific fixed effects and group j is comprised of treated facility j and its m_j closest matches. What distinguishes this approach from more standard regression-based strategies is that observations are weighted as in matching. To investigate the extent to which emissions trading has exacerbated (or mitigated) environmental injustice vis a vis CAC regulations, socio-economic and demographic variables are included in X_i.

4 Data

About 10,000 polluting facilities in California report emissions of criteria pollutants to the California Air Resources Board (ARB). All polluting facilities are required to report to their local Air Quality Management District. The ARB maintains a database of emissions reports from these local districts. Our primary data comes from this database which also includes information on industry classification. We use addresses, geocodes, and industry classifications to ensure a consistent coding of facilities across our panel.[26] We also use separate emissions data from RECLAIM to verify the emissions reported to the ARB database.[27]

We obtain demographics data from 1990 and 2000 Censuses at the block group level.[28]

[26] To ensure consistent coding over time, we identify facilities with different ID's but the same address and SIC. If the facilities do not report emissions in more than one overlapping year, then we code the facilities with the same ID. To ensure consistent coding within a year, we combine facilities with different ID's but the same geocodes and SIC.

[27] Details available upon request. The data from RECLAIM were obtained under a public records request and included information on allocations and quarterly emissions.

[28] See http://factfinder.census.gov decennial census data sets. Demographic data is summarized in the

The data include median household income, in 1989 and in 1999, and population by ethnicity and race. We construct a measure of percent minority as the percent of the population that is either non-Hispanic black or Hispanic.[29] To account for the possibility that households can sort based on pollution exposure, we emphasize the 1990 data, versus the more recent 2000 data that may be endogenous to emissions due to sorting (Banzhaf and Walsh, 2008).

To prepare these demographics data for use in our analysis, we construct radii of differing lengths surrounding each facility. We determine the percent of a block group's geographic area that is within a half, one, and two miles of each facility and use these percentages to characterize the neighborhood surrounding each facility. For example, for the one mile radius, we calculate the percent of each block group that is within a mile of a facility. We multiply that percentage by the corresponding census block group populations (separately for each demographic group). We then aggregate over block groups to get the total number of affected individuals for that facility. We replicate this procedure for a 1/2, 1, and 2 mile radius, and for each population sub-group. Note that this assumes a uniform geographic distribution of population within a block group.

Trends in facility-level NO_x emissions

Figure 2 shows the declining trends in total NO_x emissions at California facilities between 1990 and 2005. The figure illustrates that, in the aggregate, NO_x emissions from both facilities in RECLAIM and those in comparison groups were declining at similar rates prior

appendix table A1.

[29] Figure A1 helps to illustrate the spatial distribution of this measure. This figure was generated using zip code level demographics data. In the econometric analysis we use more disaggregated (i.e. census block) data which should be less susceptible to aggregation issues such as the ecological fallacy and the modifiable areal unit problem.

to the introduction of RECLAIM. In the early years of the RECLAIM program (*i.e.*, when the aggregate cap was not binding) emissions of RECLAIM facilities appear to increase slightly relative to facilities outside the program. After the cross-over point in 2000, however, the average rate of emissions decrease among RECLAIM facilities exceeds that of non-RECLAIM facilities. Overall, emissions among RECLAIM facilities have dropped 72 percent relative to pre-1993 levels, whereas emissions among non-participating facilities have dropped only 62 percent over the same period.

Table 1 summarizes a balanced sample of these same data. To construct this table, the data are partitioned into four non-overlapping periods. Period 1 encompasses years prior to the introduction of the RECLAIM program (*i.e.*, 1990-1993). Period 2 covers the early years of the RECLAIM program when the emissions cap exceeded aggregate emissions (1997-98). Period 3 includes years immediately following the "cross-over" point (2001-02). Period four includes the most recent years (2004-2005). The sample includes all facilities reporting positive emissions in each period. Overall, annual facility-level emissions are significantly larger among RECLAIM facilities *vis a vis* the comparison group. Average emissions among RECLAIM facilities fell 70 percent between period 1 and period 4.[30] This table also illustrates that annual emissions are distributed differently across RECLAIM facilities and others in all periods.

The full panel of facility-level data is unbalanced. Between periods 1 and 4, 32 percent of RECLAIM facilities and close to 54 percent of non-RECLAIM facilities fail to report

[30] When the sample omits the fourteen power producers removed from RECLAIM in Period 3, average emissions fell from 72.2 to 31.5 (a similar percentage reduction).

emissions in one or more years. Facility-level emissions data in a given period may be missing for a number of reasons, including errors in the data, a facility's failure to report emissions in a given period, or the exit of a facility. On average, treated facilities reporting emissions in all periods were larger emitters in period 1, although not significantly so.[31] Section 5.5 discusses sample selection issues in more detail.

Industrial composition of the treatment and control groups

Table 2 examines the distributions of historic, facility-level NO_x emissions among treated and control facilities, respectively. We focus on the twelve industries which accounted for the largest shares of NO_x emissions in Period 1. While refining and electricity generation are the largest polluters, about 40 percent of emissions are from firms in other four-digit SIC codes. The final column of this table reports the proportion of RECLAIM facilities with historic emissions within the 2.5^{th} and 97.5^{th} percentiles of the empirical distribution of historic emissions among control facilities in the same industry. In most cases, the support of the distribution of emissions in RECLAIM is completely overlapped by the support of the distribution in the control group. These summary statistics help to highlight a limitation of our matching strategy. Ideally, we would like to match each treated facility with a large number of control facilities in the same industry to average out idiosyncratic shocks in our estimate of counterfactual emissions. However, in some industries, the number of control facilities with very similar historic emissions will be limited. This could have implications

[31] Among RECLAIM participants, average period 1 emissions are 101.8 tons and 95.0 tons for "balanced" facilities (*i.e.*, those facilities reporting emissions in all four periods) and unbalanced facilities, respectively. For this sample, a simple regression of emissions on an indicator variable of being in the balanced sample has a standard error of 35.3 tons. Among the control group, these averages are 102.8 tons and 57.5 tons, respectively. These are statistically different as the standard error of this sample is 13.3.

for match quality. We revisit this issue below.

Emissions changes across neighborhoods

Although average NO_x emissions at RECLAIM facilities fell by 70 percent between periods 1 and 4, this average could hide increases in emissions exposure in certain neighborhoods. Table 3 investigates how changes in emissions vary with demographics, as measured by the 1990 Census. We calculate the number of individuals N_{jd} in demographic group d living near RECLAIM facility j where near is defined by the fraction of the block group within a given distance from facility j. Let Δ_j represent facility j's observed change in emissions from period 1 to period 4. For each demographic group, we measure the average change in local emissions weighted by that group's population:

$$\frac{\sum_{j \in I_1} \Delta_j N_{jd}}{\sum_{j \in I_1} N_{jd}}. \tag{5}$$

These group-specific changes are reported in the left panel of Table 3. For all three distance measures (1/2, 1, and 2 miles), we find that all groups experienced a reduction of emissions.[32] Within a 1/2 mile, high-income whites saw the largest actual reductions, while the group that saw the smallest reductions was low-income blacks. Over all races and ethnicities, high income households experienced the largest reductions. Across all incomes, whites experienced the largest reductions in emissions. The exact magnitude of the results change depending on the distance from facilities, but the findings are qualitatively similar.

In section 5, we seek to isolate only those changes in emissions that are attributable to

[32] Standard errors are computed by assuming the facility-level changes in emissions are IID. One group did not have a statistically significant drop in emissions: the standard errors for high-income blacks are very large.

RECLAIM (vis a vis CAC regulation). The right panel of Table 3 previews these results. These adjusted changes are smaller for all groups because the control group also experienced a reduction in emissions over the study period. Most importantly, the relative emissions comparisons suggest that no group was exposed to more emissions due to emissions trading. It is still the case that the reduction in emissions experienced by some groups was smaller than for others. We will examine these results more closely in section 5.4.

5 Results

In this section, we present our treatment effects estimates and conduct a series of robustness checks and falsification tests. We then test for heterogeneity in our treatment effect estimates and discuss selection issues.

Our outcome of interest is the change in facility-level annual NO_x emissions across different time periods. We report results generated using both levels and log transformed data. In the latter case, the SATT can be interpreted as our estimate of the average effect in percentage terms. Throughout, the control group is restricted to facilities located in counties that, like the RECLAIM counties, were not in attainment with the 1-hour ozone NAAQS standards in 1990 and 1993.

Recall that fourteen power producers were removed from the RECLAIM market in 2001 (Period 3) but later reentered the market.[33] For a long-term view of the overall effectiveness of RECLAIM, we analyze changes in facility level emissions between Period 1 and Period 4.

[33] One of these, Riverside Canal Power Company, is not in our complete dataset since it was decommissioned shortly after the electricity crisis due to the lack of environmental controls. (see http://www.energy.ca.gov/sitingcases/highgrove/documents/applicant/AFC_CD-ROM/Volume_01_AES_Highgrove_Project_AFC/8.13%20Waste%20Management.pdf)

Our preferred approach uses data from all RECLAIM facilities to estimate this model. To estimate the effect of trading in RECLAIM during the crucial window surrounding the price spike, we analyze changes in facility level emissions between Period 2 and Period 3. Here, our preferred specification excludes the fourteen power producers since they were not part of the market during that time. We discuss below how including or excluding these facilities changes our results and their interpretation.

5.1 Differences-in-differences estimates

We first use a simple linear regression framework to generate conditional DID estimates. Facility-level emissions changes are regressed on industry fixed effects and NO_x emissions in period 1.

Panel A of Table 4 shows that the DID estimates of the RECLAIM program's impacts on long run emissions changes are statistically significant at the five percent level. The levels estimate (-32.58 tons per year) is approximately 33 percent of the average annual emissions at RECLAIM facilities in period 1. Using the log-transformed data, the estimate is -0.30. Looking at changes in emissions over the cross-over period, estimated OLS treatment effects are also negative (Panel B). However, in levels, we cannot reject the null hypothesis of zero effect. In all of these regressions, the period 1 NO_x coefficient (not reported) is statistically significant and negative in all specifications, indicating that historic emissions are a good predictor of emissions in later years.

5.2 Semi-parametric matching

The non-parametric nearest neighbor (NN) matching estimator constructs the counterfactual estimate for each treatment case using the control cases that most closely resemble the treatment cases.[34] If m nearest neighbors are selected for each program participant, the w_{jk} are set equal to $1/m$ for the selected neighbors and zero for all other members of the comparison group.[35] We impose a strict overlap condition; only those control facilities in the same industries as RECLAIM facilities are included in the pool of potential controls. We also require that all facilities be located in ozone non-attainment areas.

Following Abadie and Imbens (2006), we augment this non-parametric matching estimation with a regression-based bias adjustment so as to mitigate any bias introduced by poor match quality. After matching the treated facilities with m nearest neighbors, within pair differences are adjusted using a parametric regression of the control outcome on X.[36]

In all of our matching, we require an exact match on the four digit standard industrial classification code. We prioritize industrial classification because these industry indicators are likely to be correlated with unobserved determinants of facility-level emissions including

[34] Within the class of matching estimators, there are a variety of matching algorithms to choose from. Asymptotically, all matching estimators produce the same estimate. However, in finite samples, different matching estimators can yield very different treatment effect estimates, particularly if one or more of the identifying assumptions is violated. Alternative matching estimators are presented in Appendix B.

[35] Although a larger m reduces the expected variance of the estimate because more information is used to construct the counterfactual for each participant, a large m also increases the bias of the estimate as the probability of making poorer matches increases. One drawback of this estimator is that all "neighbors" are equally weighted, regardless of their distance from the treated facility.

[36] More specifically, using data from matched control facilities, we regress the dependent variable (*i.e.*, differences in emissions) on the covariates. We then use this regression model to impute counterfactual estimates for all treated facilities. Note that these estimates are not likely to be sensitive to our parametric assumptions because regression techniques are only used to impute differences in outcomes among very similar facilities. These bias adjustments are discussed in more detail in Appendix B.

production technology characteristics, firm size, and demand for the products produced by the facility.

Our primary continuous matching variable is pre-treatment (*i.e.*, period 1) NO_x emissions. As we note above, historic NO_x emissions are a good predictor of emissions in subsequent periods. Our most parsimonious specification matches on attainment status, SIC code, and historic NO_x only. We refer to this as the base specification. We also experiment with matching on other observable factors that could conceivably be correlated with facility-level emissions trajectories such as the demographic and racial characteristics of the neighborhoods surrounding the facility in 1990 and the size of the facility (as measured by number of employees). The larger the number of variables we use for matching, the less accurately we are to match on those variables for which we do not require exact matching. When we include additional matching covariates, we add an industry-specific emissions quartile indicator to the list of exact match variables.

Table 4 reports results for the base NN specification and one alternative specification that includes race and demographic matching variables in addition to historic emissions and industry classification.[37] Standard error estimates are constructed using the variance formula of Abadie and Imbens (2006). In each case, RECLAIM facilities are matched to their three nearest neighbors. Appendix tables A2 and A3 demonstrate that our results are not overly sensitive to the choice of m or the bias adjustment.

For the overall change in emissions (Panel A of Table 4), the NN estimate, -20.59 tons

[37] Appendix B also summarizes results from additional matching exercises. Figure A2 reports the cumulative effects of the program for each year from 1995 through 2005.

per year, is statistically significant at the five percent level. This represents a 20 percent of the average annual emissions at RECLAIM facilities in period 1. Using log-transformed emissions data, the estimated coefficient is -0.25, implying that emissions reductions declined (in percentage terms) by approximately 25 percent more, on average, among RECLAIM facilities versus matched control facilities.[38] These estimated treatment effects are smaller (in absolute value) as compared with the OLS results. This suggests that differences in the distribution of covariates across the treatment and control group bias treatment effect estimates. When the thirteen facilities that were removed from RECLAIM in 2001 are removed from the data set, our SATT estimates remain statistically significant, although the point estimates are smaller in absolute value.

Making the period 2 period 3 comparison (Panel B), the NN estimate is -8.29 and statistically significant at the five percent level. This represents a 12 percent reduction in the average period 2 emissions at RECLAIM facilities. The SATT estimate using log transformed data is 0.26. Notably, when we include in our sample the thirteen, major polluting facilities that were removed from the RECLAIM program in 2001, estimated level impacts fall and cease to be statistically significant. However, in the log specification, these large emitters are relatively less of an outlier: here the estimates are not significantly affected.

To summarize, these results indicate that emissions reported by facilities in the RECLAIM program fell by significantly more over the fifteen year study period (*i.e.*, 1990-2005) as compared to emissions reported by a group of California facilities located in non-

[38] The estimated average annual reduction in the log specification is somewhat larger than the average reduction expressed as a percentage of period 1 emissions. This is consistent with percentage reductions being larger at smaller facilities.

attainment counties, operating in the same industries, with similar pre-RECLAIM emissions levels. When we narrow our focus to the window of time surrounding the cross-over point (*i.e.*, the point at which the aggregate cap began to bind), we continue to find that emissions reductions among RECLAIM facilities are significantly greater on average as compared to the matched controls without the thirteen power producers. When all facilities are included in the sample and the model is estimated using untransformed data, we can no longer reject the null hypothesis of zero difference in emissions trajectories across RECLAIM and control facilities during this cross-over period. However, using log-transformed data, the treatment effect estimates remain highly significant over this cross-over period, with or without the thirteen power producers.

5.3 Evaluating the underlying assumptions

In order to interpret these estimates as an unbiased measure of RECLAIM program impacts, some important assumptions must hold, in particular: conditional unconfoundedness and stable unit treatment values. Although these assumptions are not directly testable in principle, there are steps we can take to assess their plausibility.

Assessing unconfoundedness

First, our analysis assumes that the emissions trajectories of facilities in the control group are representative of the emissions trajectories that would have been observed at similar RECLAIM facilities had RECLAIM not been implemented. The weaker unconfoundedness assumption implies that $Y_{it'}(0)$ will be distributed similarly within sub-populations that are homogeneous in observable covariates. As we have two different control groups (*i.e.*, facilities located within SCAQMD exempt from RECLAIM, and similar facilities located outside the SCAQMD), we can test whether the assumption holds across these two groups. The key to this test is that these two control groups are likely to have different biases. The emissions trends at facilities outside of SCAQMD may differ from the counterfactual trends of matched treated facilities because they are operating in different counties and are regulated by different regional agencies. In contrast, the emissions trends at smaller facilities in the SCAQMD may differ from the counterfactual trends of matched treated facilities because they have lower baseline emissions.

To conduct the test, we redefine our "treated" group to be facilities in the SCAQMD but not regulated by RECLAIM. Our pool of control facilities consists of facilities located in non-attainment areas other than the South Coast. If unconfoundedness holds for these two groups, the estimated "pseudo" treatment effects should not be statistically distinguishable from zero.

Table 5 summarizes the results from this experiment. We find that the change in the average emissions (in levels or logs) among these facilities located in SCAQMD that remained subject to more prescriptive forms of emissions regulation is not statistically different from that of the control group. Put differently, the emissions trajectories among smaller SCAQMD facilities exempt from RECLAIM and the emissions at observably similar facilities located in other California air basins follow similar paths. These results are consistent with the weak unconfoundedness condition upon which our estimation is predicated.

The stronger unconfoundedness assumption requires that the control regulations mimic the changes in the stringency of regulations that RECLAIM facilities would have been subjected to had RECLAIM not been introduced. To assess the plausibility of this assumption (albeit crudely) we look at the ozone concentrations reductions mandated in SCAQMD vis a vis other California air basins over the study period.

Figure 3 illustrates the compliance requirements required under the CAAA for five air basins in California. The dotted lines connect one hour ozone concentration values in 1990 (when the CAAAs were passed) with the Federal one hour standard (0.12 ppm) in the year in which the air basin was required, under the auspices of the CAAA, to come into compliance.[39] The broken lines represent the more recently required ozone concentration reduction trajectories that pertain to the federal 8-hour ozone standard.[40] The black lines (associated with the highest ozone concentrations) correspond to the SCAQMD. Because SCAQMD was much further from attainment as compared to other air basins, the district was given more time to comply. Although ozone concentrations (and thus the extent of non-attainment) in the South Coast significantly exceed that of other California non-attainment areas, mandated reductions follow similar—if not parallel—trajectories over time. This figure helps to illustrate how mandated ozone concentration reduction trajectories were similar across California's non-attainment counties. This is consistent with our assumption that changes in the stringency of regulations affecting industrial sources of NO_x emissions in

[39]Under Title I of the 1990 CAAAs, requirements for the 96 metropolitan areas failing to attain federal ozone standards were significantly revised. Nonattainment areas were reclassified according to the extent to which they exceeded federal standards. Each classification was subject to a different deadline for achieving compliance.

[40]In 1997, the EPA concluded that the 1-hour standard was inadequate for protecting public health. The Agency issued a Federal 8-hour standard of 0.08 ppm which was officially upheld by the courts in 2001. Deadlines for compliance with the 8-hour standard can be found at http://www.epa.gov/ozonedesignations/regions/region9desig.htm.

SCAQMD and other non-attainment areas would have followed similar paths had RECLAIM not been introduced.

Assessing the stability of unit treatment values

Our analysis also assumes that the treatment received by one facility does not affect emissions at other facilities. If the introduction of the RECLAIM program caused production and associated emissions to shift from RECLAIM facilities to those exempt from the program, this would bias our counterfactual emissions estimates and exaggerate our estimates of program impacts.

Violations of this assumption are empirically intractable unless we generate some specific hypotheses regarding how these violations would manifest. We test three such hypotheses using different subsets of the control group to identify the sample average treatment effect. First, if the introduction of RECLAIM caused production to shift to control facilities, and if this shift disproportionately affected control facilities in close proximity, we would expect to find larger treatment effects when the control group is restricted to nearby facilities. The first row of results in Table 6 shows that dropping the closest facilities in the control group (*i.e.*, those located within the SCAQMD) does not significantly affect the results. The second row excludes the facilities farthest away (*i.e.*, Northern California facilities) from the control group. This also has no significant impact on the results.

Second, if RECLAIM induced shifts in production were more likely to occur in relatively less stringently regulated regions where the limits imposed by CAC regulation are more lax, we would expect to find smaller treatment effects when the control group is restricted in this way. We restrict the control group to those facilities located in *severe* non-attainment counties. The third row of results in Table 6 report SATT estimates obtained using only data from facilities in severe (versus moderate) non-attainment areas as controls. Estimated program effects are not significantly impacted.

Finally, if moving production (and thus emissions) from one facility to another is more easily coordinated within a firm versus across firms, RECLAIM induced shifts in production will be more likely to occur within a parent company with facilities inside and outside of RECLAIM (versus across facilities that do not share a common owner). In this case, we would expect to find smaller treatment effects when the control group is restricted to single plant firms. The final row of results in Table 6 shows that our results are robust to this restriction.

5.4 Heterogeneous Treatment Effects

Next, we ask whether the reduction in emissions that occurred under RECLAIM, in comparison to those in the control group, are correlated with demographics. In particular, we ask whether traditionally disadvantaged neighborhoods in the SCAQMD experienced similar emission reductions as compared with other neighborhoods.

Table 7 summarizes the results of estimating equation (4).[41] Estimation of the θ parameters in (4) facilitates a test of whether the treatment effect is heterogeneous with respect to historic emissions, income, and percent minority. We estimate each effect separately as well as jointly. Panel A of Table 7 presents the results where the dependent variable is

[41] Results using the log transformed values are reported in Table A6.

the change in the level of emissions from period 1 to 4 for the full sample. In Panel B, the dependent variable is the change in the level of emissions from period 2 to 3 for the restricted data set that focuses on those facilities that were participating in (and complying with) the cap-and-trade program during this period. We do find that RECLAIM facilities polluting more in period 1 reduced emissions more so during this time period. However, we do not find evidence of 1990 demographics being a significant determinant of which facilities reduced emissions.[42]

In all specifications, the *Period 1 NO_x* coefficient is statistically significant. Ideally, our within group matching on historic NO_x emissions would be perfect and the *Period 1 NO_x* coefficient would not be identified. In fact, our data are not sufficiently rich to facilitate perfect matching; historic emissions do vary within a group of matched facilities. Moreover, we find that this within-group variation in historic emissions is significantly correlated with the dependent variable. These results serve to highlight our concerns about the bias potentially introduced by poor match quality. All of our matching estimation incorporates a parametric adjustment to mitigate this bias (Abadie and Imbens, 2006).[43]

In Panel B of Table 7, the variable *Treat * Period 1 NO_x* is statistically significant, indicating larger emissions reductions at larger facilities. Appendix Figure A3 helps to illustrate this relationship between changes in emissions and historic emissions both for RECLAIM and other facilities in more detail.[44] We smooth the observations, separately for

[42] We have also estimated these models using the restricted sample for the change in emissions from period 1 to 4, and for the full sample from period 2 to 3. The models have also been estimated using 2000 census data, as well as using a log specification. See Appendix B for a discussion of these results.

[43] Appendix table A2 shows that our results are not highly sensitive to this bias adjustment.

[44] For each treated observation, we construct a measure of what the change in emissions would have been

RECLAIM and for other facilities, using a k-Nearest Neighbor estimator. We see that the relationship between historic emissions and change in emissions is decreasing over the range of zero to 80 tons per year of historic emissions. In contrast, the control group is relatively flat at zero for most of the range: from zero to 55 tons that accounts for over 80% of the sample.

Thus far, our analysis has focused on average correlations between the relative impacts of RECLAIM on facility-level emissions trajectories and neighborhood characteristics. We might also be interested in the distribution around the mean, and in particular, investigating whether *any* neighborhoods were exposed to more emissions under RECLAIM vis a vis the CAC counterfactual. Figure 4 illustrates the geographic distribution of emissions under RECLAIM and the CAC counterfactual, respectively. We compute the fraction of each block group that is within two miles of each RECLAIM facility and then use these fractions to assign emissions to each block group. Panel A of Figure 4 shows the RECLAIM emissions assigned to each block group by this procedure and Panel B shows the counterfactual emissions assigned to each block group. Note that if two facilities are located within two miles of a block group, emissions from both facilities are assigned to the block group.[45]

Panel A of Figure 4 shows that there is spatial clustering of the emissions permitted under RECLAIM. However, Panel B illustrates similar spatial patterns of emissions implied by the CAC counterfactual. The preceding analysis has demonstrated that, on average,

for the control group if the control group had the same historic emissions as the treated observation. This is done by using bias adjustments developed by Abadie and Imbens (2006) to mitigate bias introduced by poor match quality. We use a quadratic fit (see Appendix Table A2).

[45]This procedure is equivalent to a crude pollution transport model with transfer coefficients equal to the fraction of the block group area located within two miles of the facility.

facility-level emissions are lower under RECLAIM as compared to the CAC counterfactual. Figure 4 shows that these relative reductions are distributed across the entire SCAQMD jurisdiction. This evidence suggests that RECLAIM did not contribute to hotspots.

Our results suggest that some neighborhoods were exposed to higher levels of emissions under RECLAIM. Figure A4 in the appendix identifies these neighborhoods explicitly. Using a similar approach, we construct changes in NO_x emissions (*i.e.*, observed emissions less the CAC counterfactual emissions) by block group. A very small subset of affected block groups did see a relative increase in emissions at facilities within two miles. Almost all affected block groups had a net reduction in emissions from RECLAIM.[46]

5.5 Selection issues

Section 4 describes the unbalanced nature of our panel. Non-random selection into and out of our balanced panel could introduce selection bias. The direction of this bias, were it present, is unclear. One might be concerned that facilities with relatively high abatement costs would be more likely to exit a CAC regime that offers less compliance flexibility. This would result in inflated estimates of RECLAIM program impacts vis a vis the CAC counterfactual. On the other hand, if a market-based approach makes more stringent emissions reductions politically feasible, RECLAIM facilities with relatively high abatement costs might exit with higher frequency, thus biasing our results in the opposite direction.

Appendix B attempts to assess selection bias by estimating a Heckman selection model,

[46]The small subset of block groups that are exposed to higher emissions levels under the RECLAIM regime as compared to the CAC counterfactual is comprised of fewer minority and low income households as compared to the average block. Overall, these households are 34 percent white (versus an average of 30 percent); average household income is $52,000 versus the average $47,000.

analyzing patterns of entry and exit, and imputing missing emissions.

6 Conclusions

In this paper, we exploit some unique design features of the RECLAIM program in order to bring new evidence to bear on two important questions. First, did emissions reductions at facilities subject to Southern California's RECLAIM program exceed emissions reductions achieved at very similar facilities subject to CAC regulation over the same time period? Second, has the compliance flexibility afforded by market-based environmental regulation resulted in more (or less) pollution in traditionally disadvantaged communities?

Our results indicate that emissions at RECLAIM facilities fell approximately 20 percent, on average, relative to the control facilities over the first ten years of the program. These results are robust to alternative matching strategies. During the period of great permit price volatility, the results are more nuanced. During this period, fourteen power producers were removed from the program. When these facilities are excluded from the analysis, we find strong evidence that emissions among RECLAIM facilities fell relative to very similar control facilities. However, when all facilities are included in our analysis of emissions trends during this volatile time, the evidence is weaker.

We find no evidence that the estimated relative effects of RECLAIM on facility-level emissions vary systematically with neighborhood demographic characteristics. In particular, we find no correlation between our estimated effects and neighborhood measures of income or percent minority. We conclude that no racial or income group experienced a significant increase in emissions due to RECLAIM.

References

[1] Abadie, Alberto and Guido W. Imbens. 2006. "Large Sample Properties of Matching Estimators for Average Treatment Effects." *Econometrica*, 74(1): 235-267.

[2] Banzhaf, H. Spencer and Randall P. Walsh. 2008. "Do People Vote with their Feet? An Empirical Test of Tiebout's Mechanism," *American Economic Review*, 98(3): 843-63.

[3] Brown, Phillip. 1995. "Race, Class and Environmental Health: A Review and Systematization of the Literature." *Environmental Research*, 69(1): 15-30.

[4] Burtraw, Dallas, David A. Evans, Alan Krupnick, Karen Palmer, and Russell Toth. 2005. "Economics of Pollution Trading for SO_2 and NO_x," *Annual Review of Environment and Resources*, 30: 253-289.

[5] Chinn, Lily. 1999. "Can the Market be Fair and Efficient? An Environmental Justice Critique of Emissions Trading." *Ecology Law Quarterly*, 26(1): 89-125.

[6] Committee on Changes in New Source Review Programs for Stationary Sources of Air Pollution. 2006. *New Source Review for Stationary Sources of Air Pollution* The National Academies Press.

[7] Drury, Richard Toshiyuki. 2009. Letter to Professor Larry Goulder, Chair of the Economic and Allocation Advisory Committee, California Air Resources Board, December 3, 2009.

[8] Drury, Richard Toshiyuki, Michael E. Belliveau, J. Scott Kuhn and Shipra Bansal. 1999. "Pollution Trading and Environmental Injustice: Los Angeles' Failed Experiment in Air Quality Policy." *Duke Environmental Law Policy Forum*, 9(2): 233-289.

[9] Ellerman, A. Denny. 2006. "Are Cap-and-Trade Programs More Environmentally Effective than Conventional Regulation?" *Moving to Markets in Environmental Regulation: Lessons from Twenty Years of Experience.* Jody Freeman and Charles Kolstad, Eds. Oxford University Press.

[10] Ellerman, A. Denny. 2003. "The U.S. SO_2 Cap-and-Trade Program." Proceedings of the OECD Workshop on Ex Post Evaluation of Tradable Permits: Policy Evaluation and Reform" Paris: Organization for Economic Co-operation and Development.

[11] Ellerman, A. Denny, Paul L. Joskow, and David Harrison, Jr. 2003. "Emissions Trading in the U.S.: Experience, Lessons, and Considerations for Greenhouse Gases." Pew Center on Global Climate Change.

[12] GAO. 1983. "Siting of Hazardous Waste Landfills and Their Correlation with Racial and Economic Status of Surrounding Communities." Washington, D.C.

[13] Green, Kenneth P., Steven F. Hayward and Kevin A. Hassett. 2007. "Climate Change: Caps vs. Taxes." *Environmental Policy Outlook.* American Enterprise Institute for Public Policy Research, June.

[14] Hall, Jane V., Arthur M. Winer, Michael T. Kleinman, Frederick W. Lurmann, Victor Brajer, and Steven D. Colome. 1992. "Valuing the Health Benefits of Clean Air." *Science*, 255(5046): 812-817.

[15] Hanemann, Michael. 2008. "California's New Greenhouse Gas Laws." *Review of Environmental Economics and Policy*, 2(1): 114-129.

[16] Harrington, Winston, Richard D. Morgenstern, Thomas Sterner, and J. Clarence (Terry) Davies. 2004. "Lessons from the Case Studies" in *Choosing Environmental Policy: Comparing Instruments and Outcomes in the United States and Europe.* Winston Harrington, Richard D. Morgenstern, and Thomas Sterner, Eds., RFF Press, Washington, DC.

[17] Harrington, Winston and Richard D. Morgenstern. 2007. "Economic Incentives Versus Command and Control: What's the Best Approach for Solving Environmental Problems?" *Acid in the Environment: Lessons Learned and Future Prospects.* Gerald R. Visgilio and Diana M. Whitelaw, Eds. Springer. US.

[18] Heckman, J. 1979. "Sample selection bias as a specification error." *Econometrica*, 47: 153–61.

[19] Heckman, James, Hidehiko Ichimura and Petra Todd. 1997. "Matching As An Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program," *Review of Economic Studies,* 64(4): 605-654.

[20] Heckman, James, Hidehiko Ichimura, Jeffrey Smith, and Petra Todd. 1998. "Characterizing Selection Bias Using Experimental Data." *Econometrica*, 66(5): 1017-1098.

[21] Holland, P. 1986. "Statistics and causal inference." *Journal of the American Statistical Association*, 81: 945–70.

[22] Holland, Stephen P. and Michael R. Moore. 2008. "When to Pollute, When to Abate? Intertemporal Permit Use in the Los Angeles NOx Market." NBER WP no. 14254.

[23] Kaswan, A. 2008. "Environmental Justice and Domestic Climate Change Policy." *Environmental Law Reporter.*

[24] Klier, Thomas H., Richard H. Mattoon, and Michael A. Prager. 1997. "A Mixed Bag: Assessment of Market Performance and Firm Trading Behavior in the NO_x RECLAIM Programme." *Journal of Environmental Planning and Management*, 40(6): 751-774.

[25] Keohane, Nathaniel O., Erin T. Mansur, and Andrey Voynov. 2009. "Averting Regulatory Enforcement: Evidence from New Source Review," *Journal of Economics and Management Strategy*, 18(1): 75-104.

[26] Keohane, Nathaniel, Richard Revesz, and Robert Stavins. 1998. "The Choice of Regulatory Instruments in Environmental Policy," *Harvard Environmental Law Review*, 22(2): 313-367.

[27] Kolstad, Jonathan and Frank Wolak. 2003. "Using Environmental Emissions Permit Prices to Raise Electricity Prices: Evidence from the California Electricity Market," *CSEM Working Paper* 113.

[28] Lejano, Raul P. and Rei Hirose. 2005. "Testing the assumptions behind emissions trading in non-market goods: the RECLAIM program in Southern California." *Environmental Science and Policy*, 8: 367-377.

[29] Morello-Frosch, Rachel, Manuel Pastor, and James Sadd. 2001. "Environmental Justice and Southern California's 'Riskscape': The Distribution of Air Toxics Exposures and Health Risks among Diverse Communities." *Urban Affairs Review* 36: 551-578.

[30] SCAQMD. 1998. *Three Year Audit and Progress Report*, Diamond Bar, California.

[31] SCAQMD. 2000. *Review of RECLAIM Findings*. Diamond Bar, California.

[32] SCAQMD. 2001. *White Paper on Stabilization of NO_x RTC Prices*. Diamond Bar, California.

[33] SCAQMD. 2002. "Comments on Draft Report 'An Evaluation of the South Coast Air Quality Management District's Regional Clean Air Incentives Market- Lessons in Environmental Market and Innovation'"Diamond Bar, California.

[34] SCAQMD. Various years. *Annual RECLAIM Audit Report.* Published annually beginning with compliance year 1994. Diamond Bar, California.

[35] SCAQMD. 2007. *Over a Dozen Years of RECLAIM Implementation: Key Lessons Learned in California's First Air Pollution Cap-and-Trade Program.* Diamond Bar, California.

[36] Schubert, U. and A. Zerlauth. 1999. "Air quality management systems in urban regions: The case of the emission trading programme RECLAIM in Los Angeles and its transferability to Vienna," *Environment and Health.*

[37] Shadbegian, Ronald, Wayne Gray and Cynthia Morgan. 2007. "Benefits and Costs From Sulfur Dioxide Trading: A Distributional Analysis." *Acid in the Environment : Lessons Learned and Future Prospects.* Gerald R. Visgilio and Diana M. Whitelaw, Eds. Springer.

[38] Smith, Jeffrey and Petra Todd. 2005. "Does Matching Overcome Lalonde's Critique of Nonexperimental Estimators?" *Journal of Econometrics*, 125 (1-2): 305-353.

[39] Solomon, B.D. and R. Lee R. 2000. "Emissions trading systems and environmental justice." *Environment*, 42:32–45.

[40] Stavins, Robert N. 1998. "What Can We Learn from the Grand Policy Experiment? Positive and Normative Lessons from SO2 Allowance Trading." *Journal of Economic Perspectives*, 12(3): 69-88.

[41] Stavins, Robert N. 2007. "A U.S. Cap-and-Trade System to Address Global Climate Change" *KSG Working Paper* No. RWP07-052.

[42] Sweet, Cassandra. 2011. "California Cap-and-Trade Faces Potential Hurdle." The *Wall Street Journal*. March 3. Retrieved from http://online.wsj.com/article/SB10001424052748703300904576178431416877032.html on March 9, 2011.

[43] Sze, Julie, Gerardo Gambirazzio, Alex Karner, Dana Rowan, Jonathan London, and Deb Niemeier. 2009. "Best in Show? Climate and Environmental Justice Policy in California." *Environmental Justice*. 2(4).

[44] Tietenberg, Tom H. 2006. *Emissions Trading Principles and Practice*. Resources for the Future. Washington D.C.

[45] UCC. 1987. "Toxic Wastes and Race in the United States." Report of the Commission for Racial Justice. United Church of Christ. 1987.

[46] US EPA. 1992. "The United States Experience with Economic Incentives to Control Environmental Pollution." Washington, D.C.

[47] US EPA. 2001. "The United States Experience with Economic Incentives for Protecting the Environment." Washington, D.C.

[48] US EPA. 2002. *An Evaluation of the South Coast Air Quality Management District's Regional Clean Air Incentives Market- Lessons in Environmental Market and Innovation.* Washington, D.C.

[49] US EPA. 2005. "The Acid Rain Program and Environmental Justice: Staff Analysis." Washington, D.C.

[50] Vandenbergh, M.P. and B.A. Ackerly. 2006. "Climate Change: The Equity Problem." *Virginia Environmental Law Journal*, 26.

Figures and Tables

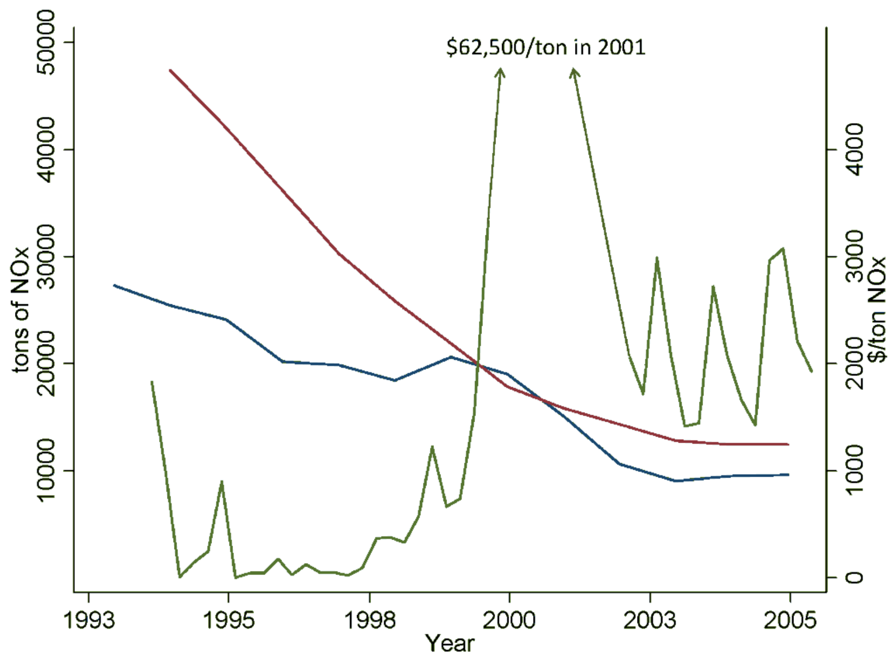


Figure 1: Trends in Nitrogen Oxides Emissions (blue), Allocations (red), and Permit Price (green).

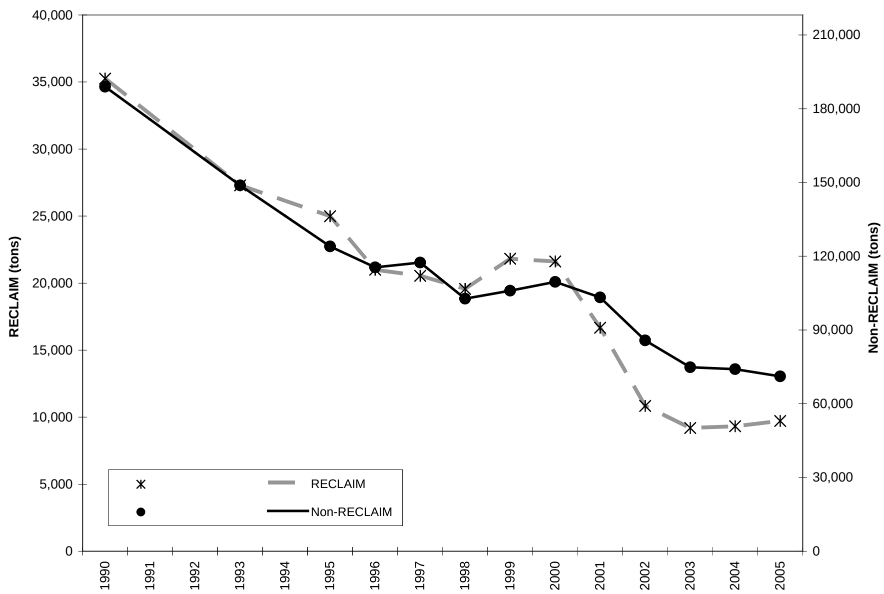


Figure 2: Total NO_x Emissions in RECLAIM and in the rest of California.

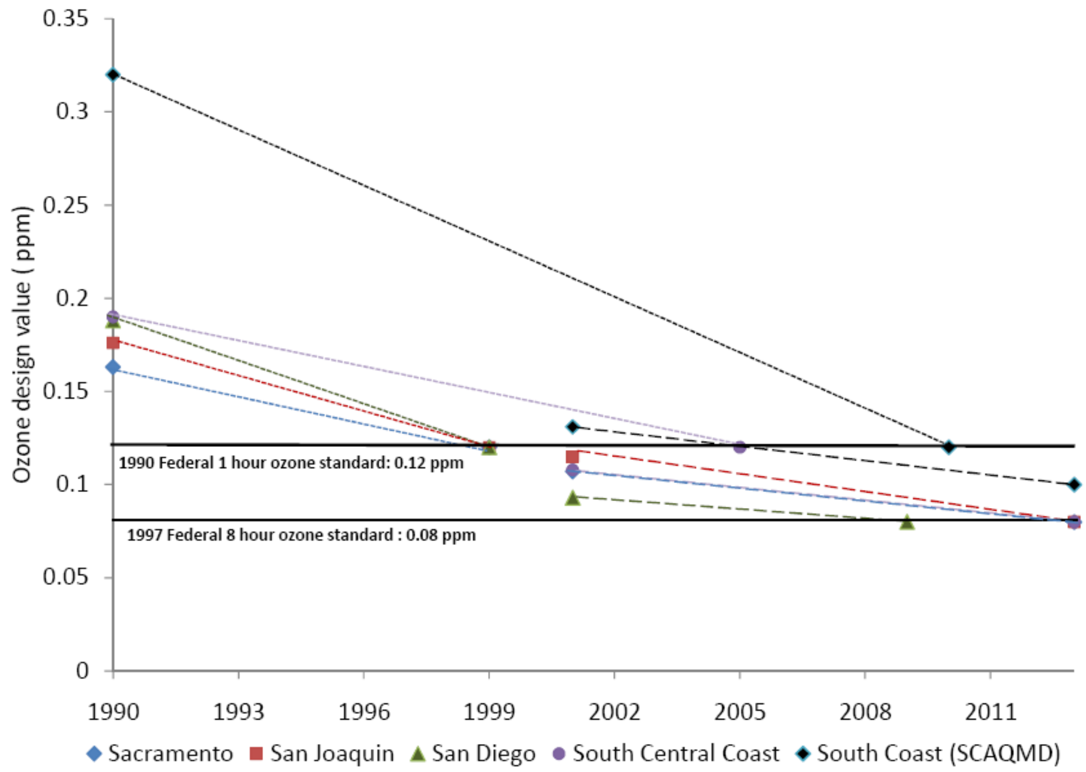
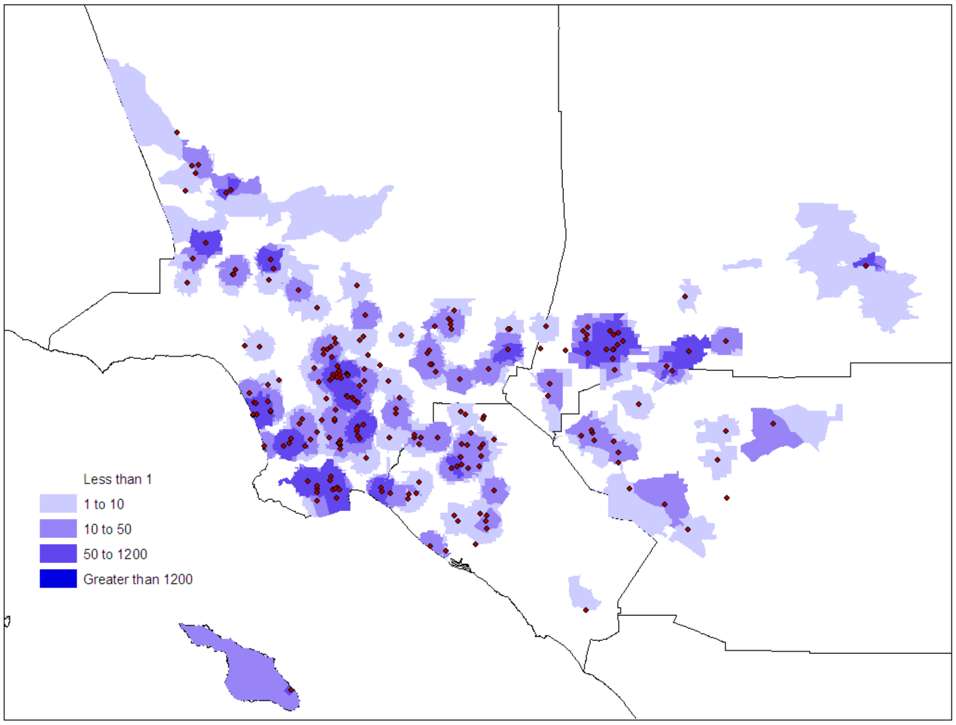


Figure 3: Required Ozone Concentration Reductions for Five Californian Air Basins.

Notes: This figure illustrates the ozone concentration reductions required of the five California air basins with the most severe air quality problems. Dotted lines connect an area's 1990 "design value" with the Federal 1- hour ozone standard in the year the basin is required to achieve compliance. A design value is an air quality measurement that is used to determine an area's air quality status (in reference to a National Ambient Air Quality Standard). Areas that had relatively high ozone concentrations in 1990 (and high design values) were given more time to come into attainment with the Federal standard. Compliance deadlines were established under the CAAA 1990. In 1997, the EPA issued a Federal 8-hour standard of 0.08 ppm. This standard was officially upheld by the courts in 2001. The broken lines connect an area's 8-hour design standard (measured in 2001) and the Federal 8-hour standard in the year the area must comply with this standard. Deadlines for compliance with the 8-hour standard can be found at http://www.epa.gov/ozonedesignations/regions/region9desig.htm. Historical data on ozone design values are available from California Air Resources Board: http://www.arb.ca.gov/adam/php_files/aqdphp/sc8start.php.

Panel A: Actual Emissions under RECLAIM



Panel B: Counterfactual Emissions under Command-and-Control (CAC)

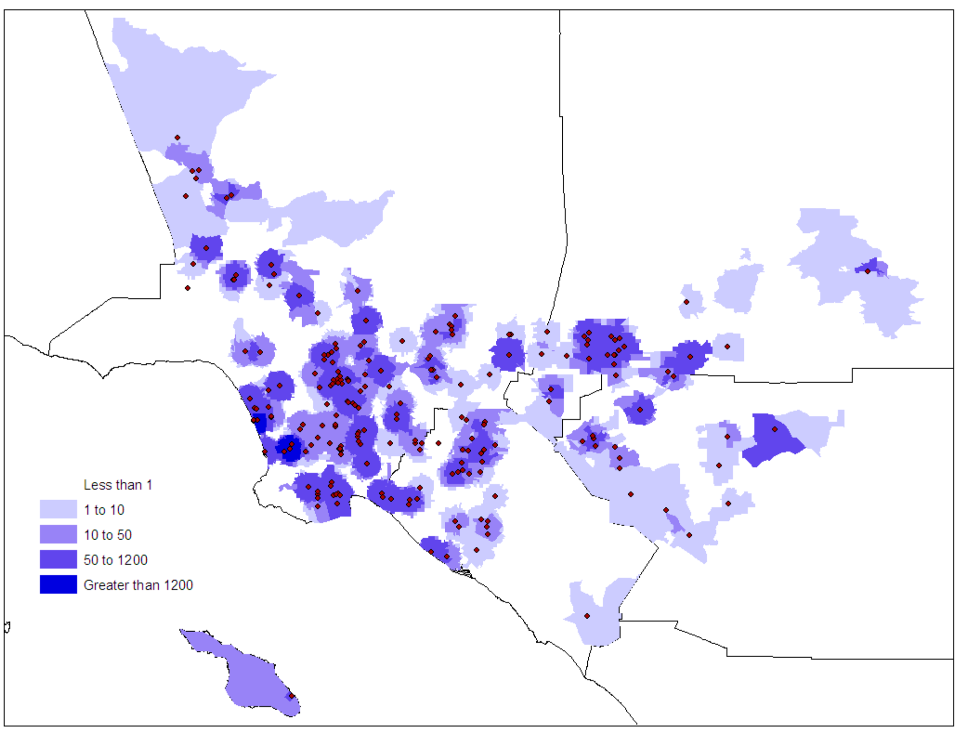


Figure 4: Actual Emissions under RECLAIM and Counterfactual, Command-and-Control Emissions in tons of Nitrogen Oxides in Period 4.

Table 1: Summary Statistics of NO_x Emissions.

Period	RECLAIM	Control	Total
Period 1 (1990-1993)	101.8 (304.4)	102.8 (430.5)	102.6 (411.9)
Period 2 (1997-1998)	62.7 (179.8)	80.0 (371.0)	77.1 (346.3)
Period 3 (2001-2002)	43.8 (125.4)	67.9 (339.6)	63.8 (314.0)
Period 4 2004-2005)	30.8 (117.1)	53.0 (290.8)	49.2 (269.6)

Notes: We report the summary statistics on the balanced sample of facilities with positive emissions in all four periods. We include the 13 RECLAIM facilities temporarily removed from the program. We report the mean tons of NO_x emissions per facility (e.g., 101.8) as well as the standard deviation (304.4). There are 213 facilities in RECLAIM and 1052 in the control group. The control group is restricted to facilities in the same two-digit SIC codes as RECLAIM facilities and that were located in counties that, during 1990 and 1993, were not in attainment with the 1-hour ozone NAAQS standards.

Table 2: Summary Statistics for Major Industries.

	RECLAIM	Treatment			Control			95 percentile
Industry	Share	obs	mean	sd	obs	mean	sd	Overlap
Petroleum Refining	37.5%	10	880	978	18	988	1570	1
Electric Services	23.9%	21	378	408	85	393	981	1
Crude Petroleum/ Natural Gas	7.1%	10	116	124	191	68	190	1
Cement	4.1%	2	699	909	9	1885	951	1
Glass Containers	3.8%	1	611		5	856	341	1
Natural Gas Trans. and Distribution	2.3%	8	85	83	4	474	612	0.88
Paper Mills	1.8%	6	82	166	5	121	170	0.83
Electric and Other Services Combined	1.6%	4	107	83	65	330	854	1
Industrial Inorganic Chemicals,	0.9%	5	31	30	10	223	683	1
Steel Works, Blast Furnaces	0.9%	3	103	120	4	20	36	0.66
Steam and Air-Conditioning Supply	0.9%	7	39	37	2	55	55	0.57
Products of Petroleum and Coal, NEC	0.8%	1	260		1	580		1
Total for Major Industries	87%	78	288	498	399	282	768	0.96

Notes: "RECLAIM Share" is the 4-digit SIC industry share of initial, period 1 NO_x emissions. We report summary statistics of tons of facility level NO_x emissions during period 1 for both treated and the control facilities. The final column reports the proportion of the treatment group that falls within the 2.5th and 97.5th percentiles of the empirical distribution of period 1 NO_x emissions in the corresponding SIC code class of controls.

Table 3: Change in Emissions (tons) Weighted by Demographic Group from the 1990 Census

	Actual change			Relative change		
Group	0.5 miles	1 mile	2 miles	0.5 miles	1 mile	2 miles
White, Low Income	-23.5 (7.4) ***	-56.0 (22.1) **	-58.4 (17.6) ***	-8.7 (3.4) **	-12.9 (5.7) **	-14.0 (5.3) ***
White, Middle Income	-94.9 (42.7) **	-69.6 (21.0) ***	-64.6 (21.3) ***	-37.2 (19.6) *	-24.1 (10.1) **	-19.3 (8.4) **
White, High Income	-170.3 (68.4) **	-163.5 (56.0) ***	-135.3 (44.1) ***	-58.5 (21.9) ***	-53.5 (18.7) ***	-38.9 (13.9) ***
Black, Low Income	-14.5 (5.3) ***	-16.9 (5.1) ***	-29.8 (10.2) ***	-2.9 (2.5)	-3.6 (2.5)	-11.7 (5.7) **
Black, Middle Income	-48.8 (20.7) **	-47.2 (22.2) **	-43.0 (22.5) *	-19.3 (10.9) *	-17.3 (11.9)	-16.0 (12.5)
Black, High Income	-110.0 (74.7)	-108.3 (71)	-67.8 (36.4) *	-55.4 (41.7)	-53.5 (39.7)	-25.8 (20.3)
Asian, Low Income	-16.2 (5.7) ***	-23.1 (8.8) **	-29.7 (8.7) ***	-4.4 (2.8)	-5.4 (5.3)	-9.0 (5.1) *
Asian, Middle Income	-36.7 (9.5) ***	-38.8 (11.5) ***	-46.8 (21.3) **	-13.9 (5.2) ***	-12.2 (5.9) **	-13.9 (8.4) *
Asian, High Income	-131.9 (55.7) **	-116.6 (45.4) **	-95.6 (39.8) **	-62.6 (34.0) *	-42.2 (17.7) **	-28.4 (14.2) **
Hispanic, Low Income	-20.3 (5.7) ***	-28.5 (9.1) ***	-33.8 (12.4) ***	-4.3 (2.4) *	-6.7 (5.2)	-10.8 (7.6)
Hispanic, Middle Income	-35.3 (10.7) ***	-34.3 (10.0) ***	-33.8 (8.5) ***	-12.0 (3.6) ***	-7.1 (4.8)	-8.6 (4.6) *
Hispanic, High Income	-108.9 (35.6) ***	-90.9 (25.5) ***	-66.7 (17.6) ***	-48.1 (19.8) **	-35.1 (11.0) ***	-19.0 (6.9) ***
All Whites	-109.8 (35.4) ***	-105.6 (30.6) ***	-94.5 (27.3) ***	-39.5 (13.1) ***	-33.8 (10.9) ***	-26.9 (9.0) ***
All Blacks	-37.8 (16.9) **	-36.3 (15.8) **	-37.8 (15.7) **	-15.2 (9.3)	-13.5 (8.7)	-14.5 (8.8)
All Asians	-55.2 (17.4) ***	-53.9 (16.1) ***	-56.2 (20.3) ***	-23.7 (10.3) **	-17.8 (6.9) **	-16.8 (7.8) **
All Hispanics	-31.3 (6.9) ***	-34.6 (7.9) ***	-36.3 (10.0) ***	-9.8 (2.9) ***	-8.8 (4.5) *	-10.9 (6.1) *
All Low Income	-19.9 (5.3) ***	-30.9 (7.9) ***	-36.2 (10.3) ***	-4.9 (2.2) **	-7.2 (4.1) *	-11.2 (6.1) *
All Middle Income	-59.4 (20.1) ***	-49.2 (12.5) ***	-47.8 (13.8) ***	-22.5 (9.1) **	-14.9 (6.2) **	-13.9 (5.9) **
All High Income	-151.7 (54.8) ***	-142.8 (45.5) ***	-115.0 (35.6) ***	-57.1 (19.3) ***	-49.2 (15.8) ***	-33.5 (11.6) ***
Total Population	-61.2 (15.4) ***	-60.0 (13.4) ***	-56.9 (13.0) ***	-21.8 (6.0) ***	-18.4 (5.5) ***	-16.9 (5.6) ***
Unweighted	-71.6 (15.1) ***	-71.6 (15.1) ***	-71.6 (15.1) ***	-20.6 (7.6) ***	-20.6 (7.6) ***	-20.6 (7.6) ***

Notes: Change in emissions from Period 1 to Period 4. Electric facilities are included. We denote significance with *** at the 1-percent level, ** at the 5-percent level, and * at the 10-percent level. The number of observations ranges from 131 to 211.

Table 4: Average Treatment Effect using Nearest Neighbors Matching.

Panel A: Change in NO_x Emissions between Periods 1 and 4.

	Levels	Logs	RECLAIM facilities	Controls
OLS	-32.58**	-0.30***	212	1,222
	(13.77)	(0.10)		
Nearest neighbor matching (base specification)	-20.59***	-0.25***	212	1,222
	(7.63)	(0.09)		
Nearest neighbor matching (alternative specification)	-18.12	-0.11	211	1,191
	(11.51)	(0.08)		
Nearest neighbor matching (restricted sample)	-14.16**	-0.20**	199	1,222
	(6.86)	(0.09)		

Panel B: Change in NO_x Emissions between Periods 2 and 3.

	Levels	Logs	RECLAIM facilities	Controls
OLS	-6.84	-0.22***	255	1,577
	(6.65)	(0.04)		
Nearest neighbor matching (base specification)	-8.29**	-0.26***	255	1,577
	(3.85)	(0.06)		
Nearest neighbor matching (alternative specification)	-6.18	-0.16***	252	1,493
	(5.06)	(0.06)		
Nearest neighbor matching (unrestricted sample)	-6.37	-0.23***	268	1,577
	(4.57)	(0.06)		

Notes: We define periods as averages of positive emissions in two years: 1990 and 1993 (period 1); 1997-98 (period 2); 2001-02 (period 3); and 2004-05 (period 4). All observations are from historic non-attainment counties. The OLS estimates control for average NO_x emissions during Period 1 and four-digit SIC code indicator variables, with standard errors clustered by air basin. For all semi-parametric matching, we match on the three closest neighbors with linear bias adjustment in levels and quadratic bias adjustment in logs. The baseline nearest neighbor matching model matches on historic emissions and exactly on four-digit SIC codes. In the alternative specification, industry-specific emissions quartile indicators are added to the exact matching variables, pre-determined demographic characteristics (race and income) are added to the matching variables. Panel A's restricted sample omits 13 facilities removed from the program in 2001. Panel B's unrestricted sample includes these facilities. For the log specifications, emissions differences are defined as ln(EmitX+1)-ln(EmitY+1) and all matching is on ln(Emit1+1). Standard errors are reported in parentheses. We denote significance with *** at the 1-percent level, ** at the 5-percent level, and * at the 10-percent level.

Table 5: Indirect Test of Unconfoundedness.

Panel A: Change in NO_x Emissions between Periods 1 and 4.

	Levels	Logs	Treated facilities	Controls
Nearest neighbor matching (base specification)	-0.96 (2.13)	-0.07 (0.06)	265	554
Nearest neighbor matching (alternative specification)	3.01 (2.49)	-0.05 (0.07)	249	520

Panel B: Change in NO_x Emissions between Periods 2 and 3.

	Levels	Logs	Treated facilities	Controls
Nearest neighbor matching (base specification)	-0.35 (1.98)	0.08 (0.06)	434	642
Nearest neighbor matching (alternative specification)	0.02 (1.17)	0.01 (0.06)	394	547

Notes: The treated facilities are redefined to be facilities in the South Coast Air Quality Management District who remained subject to CAC regulation on account of their low levels of emissions. See Table 4 for additional notes.

Table 6: Robustness to Control Group using Nearest Neighbor Matching.

Panel A: Change in NO_x Emissions between Periods 1 and 4.

Control Group	Levels	Logs	RECLAIM facilities	Controls
Base Specification	-20.59***	-0.25***	212	1,222
	(7.63)	(0.09)		
Exclude L.A. Facilities	-23.50***	-0.34***	210	778
	(7.96)	(0.09)		
Exclude Northern CA	-26.60***	-0.23**	210	767
	(7.58)	(0.11)		
Severe Non-Attainment Only	-21.65**	-0.29**	208	475
	(7.89)	(0.11)		
Single Facility Only	-19.92**	-0.23**	210	781
	(7.60)	(0.10)		

Panel B: Change in NO_x between Periods 2 and 3.

Control Group	Levels	Logs	RECLAIM facilities	Controls
Base Specification	-8.29**	-0.26***	255	1,577
	(3.85)	(0.06)		
Exclude L.A. Facilities	-8.49*	-0.21***	247	877
	(4.40)	(0.07)		
Exclude Northern CA	-14.24***	-0.28***	255	1090
	(3.90)	(0.07)		
Severe Non-Attainment Only	-13.14***	-0.17**	244	541
	(4.01)	(0.07)		
Single Facility Only	-14.99***	-0.21***	253	1027
	(4.67)	(0.06)		

Notes: Panels report results for the base specifications. See Table 4 for notes.

Table 7: Environmental Justice Results

Panel A: Change in NO_x Emissions between Periods 1 and 4.

	1	2	3	4	5	6	7
Treatment	-20.64**	-20.38*	-17.49**	-20.46**	-18.52**	-15.26***	-17.71**
	(7.81)	(8.85)	(6.17)	(7.41)	(7.04)	(4.36)	(5.29)
Treat * Period 1 NO_x	-0.19			-0.19	-0.19		-0.18
	(0.11)			(0.11)	(0.11)		(0.11)
Treat * Income		-1.27		-0.65		0.42	-0.02
		(0.96)		(1.09)		(1.95)	(1.53)
Treat * %Minority			0.94		0.43	1.04	0.41
			(0.60)		(0.36)	(0.96)	(0.51)
Period 1 NO_x	-0.48***	-0.49**	-0.49**	-0.48***	-0.48***	-0.49**	-0.48***
	(0.11)	(0.15)	(0.15)	(0.11)	(0.11)	(0.14)	(0.11)
Income		0.10		0.16		-0.66	-0.24
		(0.80)		(0.74)		(1.47)	(1.04)
%Minority			-0.35		-0.22	-0.52	-0.28
			(0.31)		(0.26)	(0.56)	(0.37)
R^2	0.87	0.85	0.85	0.87	0.87	0.85	0.87

Panel B: Change in NO_x between Periods 2 and 3.

	1	2	3	4	5	6	7
Treatment	-6.70***	-7.19**	-6.29***	-7.16***	-6.62***	-6.45***	-7.05***
	(1.43)	(2.22)	(1.35)	(1.45)	(1.25)	(1.85)	(1.23)
Treat * Period 1 NO_x	-0.06***			-0.07***	-0.07***		-0.07**
	(0.02)			(0.02)	(0.02)		(0.02)
Treat * Income		-0.16		-0.09		-0.12	-0.22
		(0.24)		(0.17)		(0.36)	(0.35)
Treat * %Minority			0.09*		-0.004	0.05	-0.07
			(0.04)		(0.045)	(0.11)	(0.14)
Period 1 NO_x	-0.35***	-0.34***	-0.34***	-0.34***	-0.34***	-0.34***	-0.34***
	(0.08)	(0.05)	(0.05)	(0.08)	(0.08)	(0.06)	(0.08)
Income		0.19		0.16		0.05	0.15
		(0.36)		(0.33)		(0.47)	(0.46)
%Minority			-0.11		-0.05	-0.10	-0.02
			(0.07)		(0.06)	(0.11)	(0.11)
R^2	0.52	0.47	0.47	0.49	0.49	0.47	0.49

Notes: Panels report results for the base specifications. For regressions with 1990 demographic data, there are 875 and 1043 observations in Panels A and B, respectively. Group fixed effects are not shown. Treated observations receive a weight of one and control observations receive a weight of $1/m_j$, where m_j is the size of the control group for treated facility j. %Minority is percent of population that is black or Hispanic. See Table 4 for additional notes.

Appendices: Not for Publication

Appendix A: Ex post evaluation of the RECLAIM program

Evaluations of the RECLAIM program have been carried out by SCAQMD staff (SCAQMD, various years), the United States Environmental Protection Agency (US EPA 2002; US EPA 2006), and academic researchers (Gangadharan, 2000; Schubert and Zerlauth, 1999). Although these studies and reports arrive at different conclusions, there is consensus that a RECLAIM program evaluation is an important exercise:

> How have actual emissions reductions [in RECLAIM] compared to those that would have occurred under the subsumed CAC system? While there can be no definitive answer, this question is so central to the affected public in any area contemplating converting from CAC to a trading based program that we are obligated to try to answer it. (US EPA, 2002)

In the periodic program evaluations carried out by SCAQMD, the aggregate RTC permit allocation serves as a proxy for counterfactual emissions. The authors maintain that this is a reasonable, and potentially conservative estimate of counterfactual emissions because the aggregate permit allocation was designed to track *ex ante* expected endpoint mass emissions under the subsumed suite of CAC rules that were being fiercely opposed by industry. These periodic evaluations routinely conclude that RECLAIM is achieving emissions reductions equivalent to, and possibly greater than, what would have been achieved under the subsumed CAC measures.

A comprehensive EPA study (US EPA, 2002) argues that assumptions made during initial projections for the RECLAIM program were "not valid predictors of real world behavior," nor were they substantiated with actual data (US EPA, 2002). Consequently, initial RTC allocations are dismissed as invalid measures of counterfactual emissions. The authors allege that RECLAIM has "produced far less emissions reductions than could have been expected from the subsumed CAC system" (US EPA, 2002).[47]

Unresolved disagreements about what constitutes an appropriate measure of counterfactual emissions have resulted in a plurality of opinions regarding RECLAIM's overall performance. Whereas the Deputy Executive Officer for the California Air Resources Board has stated publicly that RECLAIM "hasn't done as well as the regulations it replaced" (US EPA, 2006), a Pew Center report concludes that "the [RECLAIM] program's ten-year phase-in design and trading provided the flexibility that led to the achievement of environmental goals that had been previously elusive." (Ellerman *et al.*, 2003).

Appendix B : Additional robustness tests

This appendix provides additional evidence on the robustness of the main results to alternative specifications and matching estimators.

Number of neighbors

We use a leave-one-out validation approach to choose among the nearest neighbor estimators (Black and Smith, 2004). Our objective is to estimate the counterfactual emissions for RECLAIM facilities, $Y_{it'}(0)$. Although we do not observe this for any RECLAIM facility, we do observe this at the control facilities. Leave-one-out validation uses these control obser-

[47] SCAQMD was quick to respond to allegations that their counterfactual emissions significantly exceeded that which could realistically have been expected under the subsumed CAC rules. This dispute was never resolved. A more recent, retrospective overview of the RECLAIM program published by the US EPA concludes: "RECLAIM shows the critical nature of baseline credibility in a program's perceived success or failure"(US EPA, 2006).

vations to determine which of the competing models best fit the data. The basic approach is as follows. We drop observation j in the control group and use the remaining control observations to estimate $\widehat{Y}_{jt'}(0)$. The associated forecast error is given by $e_j = \widehat{Y}_{jt'}(0) - Y_{jt'}(0)$. This process is repeated for all facilities in the control group. We select the estimator that minimizes the mean squared error of the forecasts.

Table A2 reports the robustness of the main nearest neighbors matching results to the number of neighbors. For the overall effect (period 1 to 4), the results are significant and qualitatively similar for 1, 2, or 3 neighbors. With more neighbors, the estimates are only weakly significant. For the trading effects (period 2 to 3), the results are quite similar for 2, 3, 4, or 5 neighbors.

Bias adjustments

We also examine the robustness of our nearest neighbor results to the bias adjustments we make. When matches between treatment facilities and the closest controls are inexact, our semi-parametric matching estimator adjusts the difference in the predicted counterfactual outcome. This adjustment is based on an estimated regression of the emissions differences in the control group on historic emissions and industry fixed effects. This regression function, estimated using least squares, is best approximated with a linear function when the data are in levels. A more flexible specification is warranted when we use log-transformed data. We use a quadratic bias adjustment.

Table A3 reports results using no correction, a linear correction, and a quadratic correction. Patterns of coefficient significance are not significantly affected by these bias adjustments. Results are also robust to using additional covariates in the bias adjustment function (not shown).

Matching covariates

The baseline specification, which we emphasize in the paper, matches on pre-period attainment status, 4 digit standardized industry classification (sic), and historic NO_x emissions. Our approach assumes that, conditional on these variables, NO_x emissions at facilities operating under a CAC regime would follow parallel paths over the study period. This is a strong assumption, and we would ideally control for additional factors that could affect emissions (such as production technologies, firm size, or characteristics of the markets served by these facilities). Unfortunately, for many of the facilities in our analysis, there is a paucity of data available.

In an effort to augment our matching with additional facility level information, we gained access to an extract of the National Establishment Time-Series (NETS) Database that includes all business establishments in California over the sample period 1992–2004. These data are derived from Dun & Bradstreet data, and include detailed information about the location of establishments, establishment-level standard industrial classification, ownership structure, employment, and credit rating information. Merging these data with our database was not straightforward due to differences in how facilities report their facility names, locations, and primary industry classifications. We used a truple matching algorithm that searches for common three letter combinations in the names and locations of facilities appearing in our database and the NETS data, respectively. This allowed us to successfully merge approximately 40 percent of the facilities in our data with the NETS data.

Table A4 reports results from matching exercises not summarized in the paper. Note that the larger the number of variables we use, the less accurately we match on those variables for which we do not require exact matching. When we include additional matching covariates, we add an industry-specific emissions quartile indicator to the list of exact match variables.

Alternative matching estimator

Asymptotically, all matching estimators produce the same estimate. However, in finite samples, different matching estimators can yield very different treatment effect estimates, particularly if one or more of the identifying assumptions is violated. Since the seminal work of Rosenbaum and Rubin (1983), there has been considerable interest in methods that avoid adjusting directly for observable covariates and instead adjust for differences in the propensity score (*i.e.*, the conditional probability of treatment). An important result in the literature is that, if unconfoundedness holds, conditioning only on the propensity score assures independence of D_i and $Y_i(0)$. Recent work has demonstrated that, when there is good overlap in the distribution of propensity scores for treated and control facilities, reweighting estimators outperform nearest neighbor or kernel matching in finite samples (Busso *et al.*, 2009). We implement a propensity score matching estimator. All treated observations receive a weight of one, whereas control observations receive a weight $\frac{\widehat{p}}{1-\widehat{p}}$ (where $\widehat{p}$ is the estimated propensity score).

The propensity score equation describes the process by which the data are filtered or selected to produce the observed sample. We estimate the propensity scores using a reduced form probit model. Explanatory variables include industry affiliation, historic emissions, and squared historic emissions. We enforce a common support. Balance is achieved and there is significant overlap in the propensity scores of the treatment and comparison groups.

Although matching on propensity scores balances treatment and controls across the set of covariates, facilities with very similar propensity scores may have different combinations of observable characteristics. In our case, we find that matching on p-scores does not always imply a close match on observables (even after adding higher order terms to the selection equation). This poor match quality can introduce bias. Consequently, we use a propensity score based refinement of weighted regression: the so-called "double robust" (DR) estimator (Robins and Rotnitzky, 1995; Robins and Ritov, 1997). By combining propensity score matching with regression, we can reduce bias introduced by poor match quality.[48] Table A5 summarizes the main results. These SATT estimates are larger in absolute value as compared to the NN estimates and somewhat noisier.

Heterogeneous treatment effects

Table A6 reports the results from estimating equation (4) using the log transformed data. From period 1 to 4, we find that larger historic polluters reduced emissions by a greater percentage. In specifications (3) and (6), we find weak evidence (*i.e.*, the coefficients are significant at the 10 percent level) that neighborhoods with greater percent minority experienced more emissions, all else equal. However, these results are no longer significant when one controls for the heterogeneous treatment effect of historic emissions, as in specifications

[48] This double robust estimator will not always constitute an improvement upon the more standard parametric regression approach. Reweighting of observations will only add noise if the parametric regression model is correctly specified (Freedman and Berk, 2008).

(5) and (7). Panel B does not find evidence of heterogeneous treatment effects.

We test the robustness of these results as well as the results of Table 7 to using 2000 demographic data. We also examine these models using the restricted sample for the change in emissions from period 1 to 4, and for the full sample from period 2 to 3. In testing the robustness of Table 7, none of the 28 estimates of income and only one of the 28 estimates of percent minority is significant at the 5 percent level (percent minority is significant using the 2000 demographics data when looking at the period 2 to 3 trading for non-electricity facilities. Table A6 results are slightly more sensitive. None of the 28 estimates of income is significant at the 5 percent level. However, of the 28 estimates of percent minority, nine are significant. In particular, using 2000 demographics data, the coefficient on percent minority is significant in the regressions of changes in log emissions from period 1 to 4 for all firms as well as for non-electricity firms. Note that we use 1990 demographics in our main specification because 2000 demographics are potentially endogenous. Finally, even with 1990 demographics, percent minority is significant at the 5 percent level with an implied elasticity of 0.63 when evaluating the change in log emissions from period 1 to 4 for the non-electricity facilities but not controlling for historic emissions (the equivalent of Column 3 of Table A6, Panel A).

Selection

Since the analysis matches on emissions levels before RECLAIM began, facilities which entering during the time frame of the study cannot be included in the analysis. Also, facilities exiting prior to the post-treatment period are excluded. Non-random entry and exit might introduce selection bias into our results. We first discuss the Heckman test for sample selection bias, then analyze patterns of entry and exit and explore whether our measures of entry and exit might simply arise from misreporting. We then test the robustness of our main result to imputing missing emissions.

To credibly identify a Heckman selection model, we need a variable that significantly determines selection into our sample, but can be credibly excluded from the outcome equation. We went to great lengths to find such a variable, but we were ultimately unsuccessful. Absent a credible exclusion restriction, the standard Heckman selection correction is technically possible to implement, but not very informative.

These identification issues notwithstanding, we do conduct a Heckman test for selection bias. When we include the inverse Mills ratio as an additional explanatory variable in our parametric regressions, it is not statistically significant. For reasons we have articulated, this result is not very meaningful.

To develop a better sense of the patterns of entry and exit in the data, we define variables to indicate whether a facility entered or exited across two periods. For example, the variable Exit14 takes a value of one if the facility reported positive emissions in period 1, but no emissions in period 4. For the panel of 535 RECLAIM facilities and 10,447 non-RECLAIM (*i.e.*, control) facilities, 30% of the RECLAIM facilities exited and 40% of the non-RECLAIM facilities exited between periods 1 and 4. When we regress the entry and exit indicators on a RECLAIM dummy, emissions in the observed period (*e.g.*, period 1 emissions when Exit14 is the dependent variable), an interaction of these two variables and SIC fixed effects, we find that entry and exit are less likely in RECLAIM than in the rest of California. One possible explanation for the differential entry and exit rates might be missing data reports. We construct a missing indicator to show facilities which did not report emissions data without entering or exiting. For example, "Missing14" is one if emissions are reported in periods 1 and 4, but are missing in either period 2 or 3. Only about 1-2% of the facilities have missing emissions reports by this measure.

While differential entry and exit rates are not necessarily indicative of selection bias, they do warrant concern insofar as these differences could be indicative of non-random. In order to gain any traction on the selection issue, we need to make additional assumptions about the nature of the selection process. One approach to investigating the potential for selection bias involves imputing emissions for the facilities that drop out of our sample and examining whether the results change when we re-estimate our model using this "completed" data set. If we assume that the missing emissions observations can be reasonably imputed using the emissions at similar facilities that remain in the data set, equivalent to assuming that selection is random conditional on observables, this approach can shed light upon how our results are affected by non-random exit. For those facilities that drop out of the data, we construct an imputed estimate of the missing emissions observations using data from similar facilities in the same emissions regulation regime. More precisely, we match the attrited facilities with similar facilities in the same industry and with similar period 1 emissions. We perform this matching separately for the treatment and control groups, respectively. Results are reported in Table A7. In all cases, the estimates using the completed sample are somewhat smaller in absolute value, but highly statistically significant.

Appendix C: Further evidence on the environmental justice implications of emissions trading

Concerns about the environmental justice implications of emissions trading have strongly influenced the debate surrounding California's greenhouse gas regulations (Hanemann, 2008; Sze et al., 2009). Lejano and Hirose (2005) show that, in the very early years of RECLAIM, the purchase of RTCs by sources in one low income community in particular (Wilmington California) led to NO_x concentrations that exceeded what would have been observed under autarky. These findings have since been interpreted as evidence that traditionally disadvantaged communities were more harmed under RECLAIM than they would have been under CAC (see, for example, Drury 2009).

Although these findings certainly warrant concern, there are two potential problems with interpreting this as conclusive evidence that emissions trading disproportionately harmed low income communities in Southern California. First, using the initial RTC allocation as a proxy for the emissions that would have occurred under CAC is problematic. Emissions limits imposed by RECLAIM are allegedly much more stringent than what would have been politically feasible under CAC. Consequently, the initial RTC allocations likely provide a biased estimate of emissions absent RECLAIM. Second, in order to conclude that RECLAIM disproportionately harms traditionally disadvantaged communities, it is important to investigate the source of the permits that flowed into Wilmington, to consider emissions outcomes beyond 1996, and to look at the overall pattern of emissions trading under RECLAIM. The authors are careful to emphasize that a more comprehensive analysis is required in order to conclude that emissions trading in RECLAIM disproportionately harmed poor and minority communities.

In this appendix, we extend our analysis in order to revisit this relationship between permit allocations and emissions. More precisely, we investigate whether the relationship between facility-specific permit allocation trajectories and facility-level emissions trends over the study period vary systematically with the demographic characteristics of the neighborhood in which the facility is located. On average, we would expect trends in allocations and emissions to be strongly positively correlated. If emissions exactly equal allocations, this coefficient is exactly one. When facility-level changes in emissions over the study period are regressed on the corresponding change in facility-specific permit allocations the coefficient is 0.65 and precisely estimated (standard error 0.02). Because permits were initially overallocated, permit allocations fell more precipitously than emissions on average.

If, as alleged, permits flowed disproportionately into low income communities, we would expect this positive correlation to be decreasing with income. If permits flowed disproportionately into minority communities, we would expect the positive correlation to be increasing with percent minority with income. When we interact the facility-specific allocation changes with our measures of neighborhood demographic variables, the coefficient on the income interaction is positive and statistically significant at the five percent level; the coefficient on the percent minority interaction is negative and statistically significant at the five percent level. These results suggest that, relative to the number of permits allocated, emissions fell relatively more sharply in low income and minority neighborhoods. These findings are not consistent with the claim that emissions permits flowed disproportionately into traditionally disadvantaged neighborhoods. Of course, additional research is warranted in order to definitively resolve this issue.

Appendix References: Not for Publication

[1] Black, Dan A., Jeffrey A. Smith. 2004. "How robust is the evidence on the effects of college quality? Evidence from Matching." *Journal of Econometrics*, 121(1-2): 99-124, Higher education (Annals issue).

[2] Busso, Matias, DiNardo, John E. and McCrary, Justin. 2009. "New Evidence on the Finite Sample Properties of Propensity Score Matching and Reweighting Estimators." *IZA Discussion Paper No. 3998.*

[3] Freedman, David A. and Richard A. Berk. 2008. "On Weighting Regressions by Propensity Scores." *Evaluation Review*, 32: 392-409.

[4] Gangadharan, Lata. 2000. "Transaction Costs in Pollution Markets: An Empirical Study." *Land Economics*, 76(4): 601-614.

[5] Robins, J.M. and Ritov, Y. 1997. "Towards a Curse of Dimensionality Appropriate (CODA) Asymptotic Theory for Semi-Parametric Models." *Statistics in Medicine*, 16: 285-310.

[6] Robins J.M., Rotnitzky A and Zhao L.P. 1995. "Analysis of Semiparametric Regression-Models for Repeated Outcomes in the Presence of Missing Data." *Journal of the American Statistical Association*, 90(429): 106-121.

[7] Rosenbaum, P. R. and D. B. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika*, 70(1): 41-55.

[8] US EPA. 2006. "An Overview of the Regional Clean Air Incentives Market." Washington, D.C.

Appendix Figures and Tables

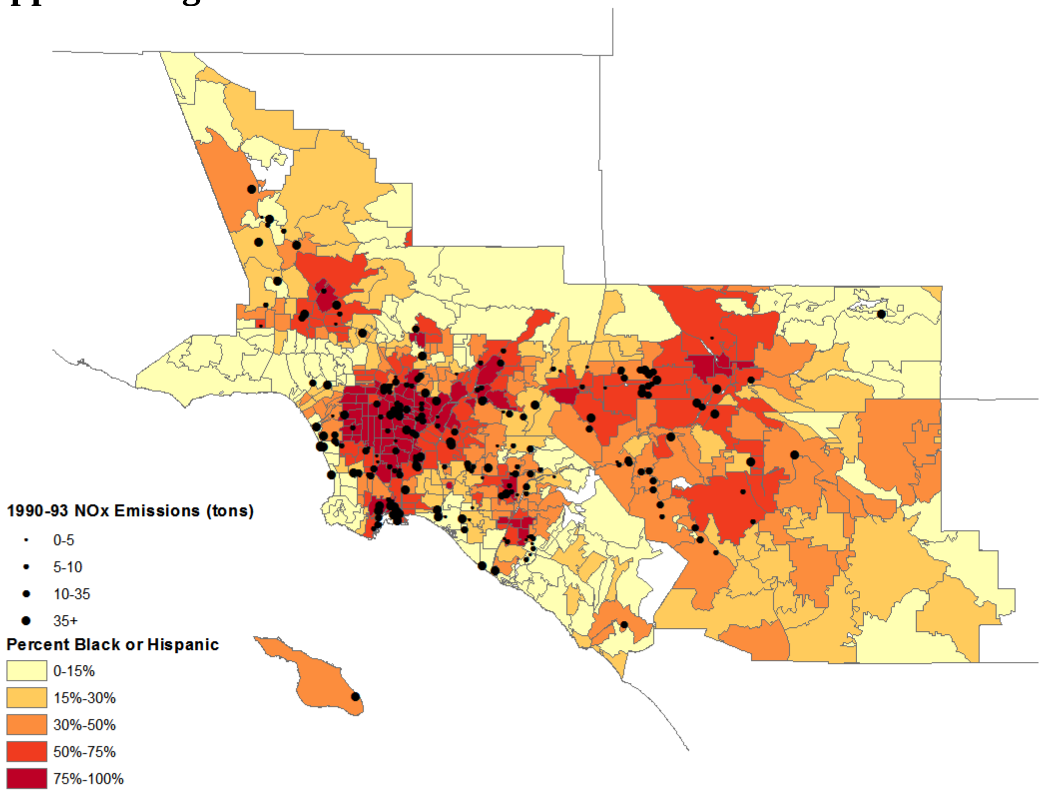


Figure A1: The South Coast Air Quality Management District.

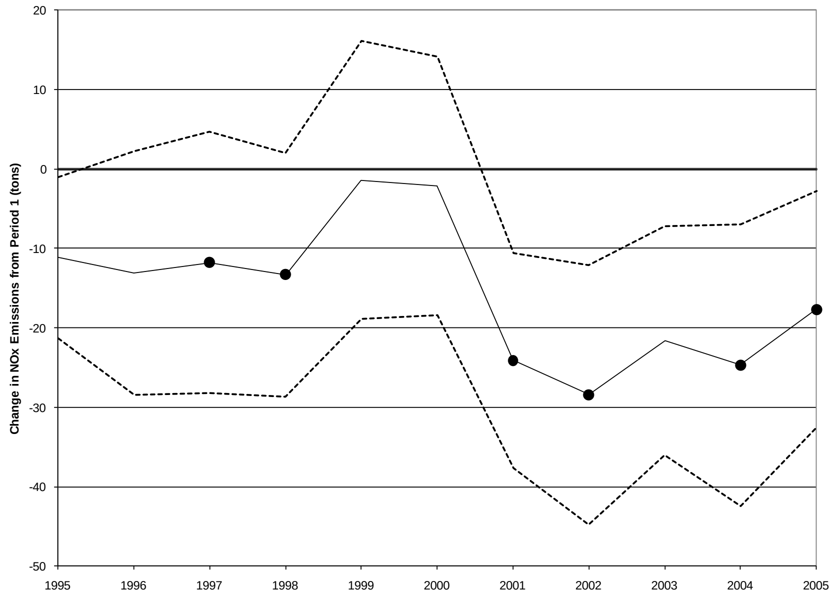


Figure A2: Average Cumulative Treatment Effect by Year (relative to Period 1 emissions), matching m=3.

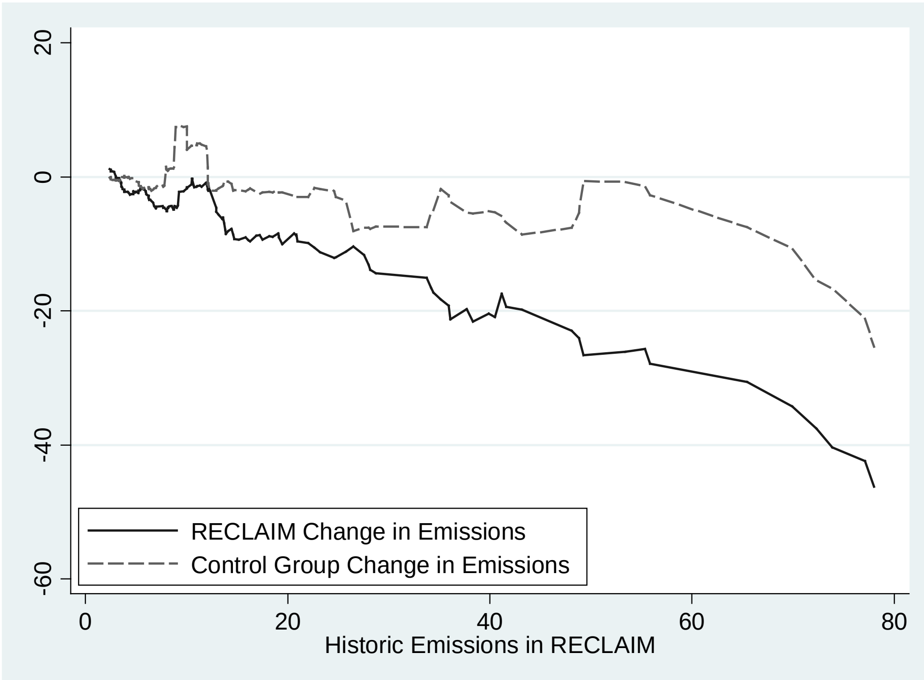


Figure A3: k-Nearest Neighbor Regression of Changes in Emissions from Period 1 to Period 4 in the RECLAIM and Control Groups on Period 1 Emissions. The sample is from the main results shown in Table 4.

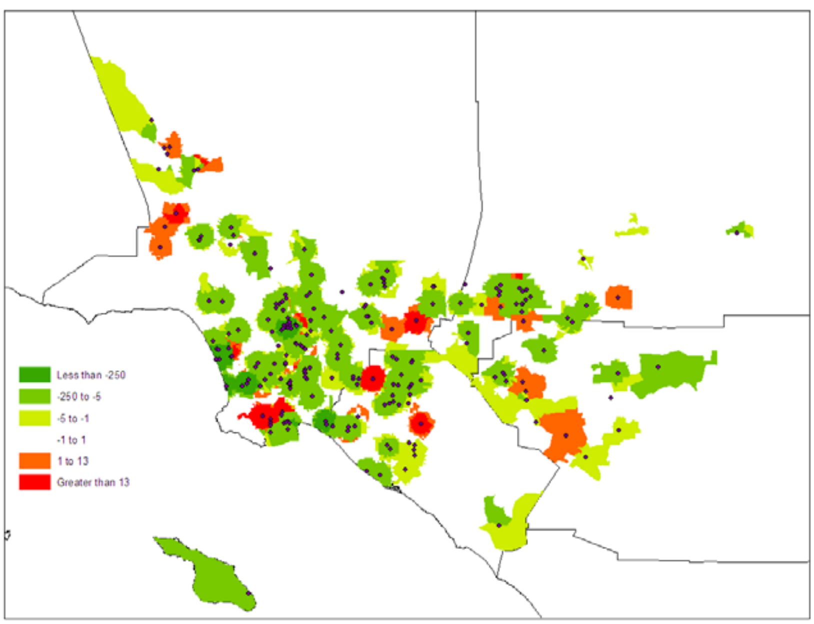


Figure A4: Difference between Actual and Counterfactual Command-and-Control Emissions in Period 4 in tons.

Table A1: Demographic Summary Statistics

Variables for RECLAIM facilities	n	mean	std dev	min	max
median income within 1 mile of a facility (in $1000s) in 1989	211	35.7	11.7	10.2	80.8
percent black or Hispanic within 1 mile of facility (in 1990)	211	49%	27%	0%	99%
indicator of whether toxics were measured on site	211	28%	45%	0%	100%
% change in total county employment from period 1 to 4	211	2276%	1837%	1112%	6909%
% change in total county payroll from period 1 to 4	211	7617%	3008%	5626%	15300%
% change in total county establishments from period 1 to 4	211	1703%	667%	1287%	3549%
Indicator of coastal permits	211	70%			
Petroleum Refining	211	7%			
Stone, Clay, and Glass Products	211	9%			
Primary Metal Industries	211	9%			
Electric and Gas Services	211	17%			

Variables for control facilities	n	mean	std dev	min	max
median income within 1 mile of a facility (in $1000s) in 1989	664	34.4	11.5	13.8	85.2
percent black or Hispanic within 1 mile of facility (in 1990)	664	41%	27%	0%	99%
indicator of whether toxics were measured on site	664	20%	40%	0%	100%
% change in total county employment from period 1 to 4	664	2689%	1689%	587%	12254%
% change in total county payroll from period 1 to 4	664	9343%	3249%	5626%	27731%
% change in total county establishments from period 1 to 4	664	1439%	827%	-161%	6038%
Indicator of coastal permits	664	70%			
Petroleum Refining	664	8%			
Stone, Clay, and Glass Products	664	9%			
Primary Metal Industries	664	8%			
Electric and Gas Services	664	16%			

Table A2: Robustness to Number of Neighbors

Panel A: Change in NO_x Emissions between Periods 1 and 4.

Dependent variable	m(1)	m(2)	m(3)	m(4)	m(5)
Levels	-27.19***	-23.98***	-20.59***	-18.52**	-17.96**
	(8.74)	(8.07)	(7.63)	(7.81)	(8.10)
Logs	-0.38***	-0.26**	-0.25***	-0.26***	-0.24***
	(0.11)	(0.10)	(0.09)	(0.09)	(0.08)

Panel B: Change in NO_x Emissions between Periods 2 and 3.

Dependent variable	m(1)	m(2)	m(3)	m(4)	m(5)
Levels	-4.94	-9.18**	-8.29**	-10.18**	-10.32**
	(7.08)	(4.13)	(3.85)	(4.64)	(4.06)
Logs	-0.29***	-0.26***	-0.26***	-0.23***	-0.22***
	(0.07)	(0.06)	(0.06)	(0.06)	(0.05)

Notes: m(*n*) denotes the *n* neighbors matched. Panels report results for the base specifications. See Table 4 for notes.

Table A3: Robustness to Bias Adjustment

Panel A: Change in NO_x Emissions between Periods 1 and 4.

Dependent variable	No bias adjustment	linear bias adjustment	quadratic bias adjustment
levels	-25.02***	-20.59***	-17.79**
	(7.63)	(7.63)	(7.63)
logs	-0.32***	-0.27***	-0.25***
	(0.09)	(0.09)	(0.09)

Panel B: Change in NO_x Emissions between Periods 2 and 3.

Dependent variable	No bias adjustment	linear bias adjustment	quadratic bias adjustment
levels	-9.60**	-8.29**	-9.42**
	(3.85)	(3.85)	(3.85)
logs	-0.26***	-0.26***	-0.26***
	(0.06)	(0.06)	(0.06)

Notes: Panels report results for the base specifications. See Table 4 for notes.

Table A4: Robustness to Alternative Matching Specifications

Panel A: Change in NO_x Emissions between Periods 1 and 4.

Description	Levels	Logs	RECLAIM facilities	Controls
Base specification	-20.59**	-0.25***	212	1222
	(7.63)	(0.09)		
Base specification + %minority	-25.00**	-0.16*	211	1191
	(10.55)	(0.08)		
Base specification + % income	-15.43	-0.16**	211	1191
	(11.04)	(0.08)		
Base specification + 90employment	-7.34	0.11	80	332
	(25.66)	(0.17)		

Panel B: Change in NO_x between Periods 2 and 3 for Non-Electricity Facilities.

Description	Levels	Logs	RECLAIM facilities	Controls
Base specification	-8.29**	-0.26***	255	1577
	(3.85)	(0.06)		
Base specification + %minority	-6.76	-0.15*	252	1493
	(4.71)	(0.08)		
Base specification + % income	-6.91	-0.17***	252	1493
	(4.83)	(0.05)		
Base specification + 90employment	6.79	0.15	94	431
	(21.16)	(0.11)		

Table A5: Average Treatment Effect using Propensity Score Matching

Panel A: Change in NO_x Emissions between Periods 1 and 4.

Dependent variable	No bias adjustment
levels	-24.81*
	(13.86)
logs	-0.27**
	(0.12)

Panel B: Change in NO_x Emissions between Periods 2 and 3.

Dependent variable	No bias adjustment
levels	-14.78***
	(2.22)
logs	-0.28***
	(0.02)

Table A6: Environmental Justice Results in Logs

Panel A: Change in log NO_x Emissions between Periods 1 and 4.

	1	2	3	4	5	6	7
Treatment	-0.25**	-0.21*	-0.20**	-0.21**	-0.20**	-0.14*	-0.15**
	(0.10)	(0.10)	(0.08)	(0.09)	(0.07)	(0.07)	(0.06)
Treat * Period 1 NO_x	-0.13**			-0.11**	-0.11**		-0.09**
	(0.05)			(0.05)	(0.04)		(0.03)
Treat * Income		-0.21		-0.18		0.26	0.21
		(0.42)		(0.42)		(0.58)	(0.55)
Treat * %Minority			0.82*		0.67	0.96*	0.80
			(0.41)		(0.40)	(0.51)	(0.45)
Period 1 NO_x	-0.35**	-0.34**	-0.35**	-0.36**	-0.36***	-0.37***	-0.38***
	(0.11)	(0.11)	(0.11)	(0.11)	(0.10)	(0.10)	(0.10)
Income		-0.25		-0.22		-0.48	-0.41
		(0.26)		(0.25)		(0.41)	(0.39)
%Minority			-0.20		-0.14	-0.55	-0.45
			(0.31)		(0.30)	(0.46)	(0.43)
R^2	0.32	0.34	0.34	0.34	0.34	0.34	0.35

Panel B: Change in log NO_x between Periods 2 and 3 for Non-Electricity Facilities.

	1	2	3	4	5	6	7
Treatment	-0.25***	-0.23***	-0.21***	-0.23***	-0.21***	-0.19***	-0.19***
	(0.04)	(0.03)	(0.04)	(0.03)	(0.05)	(0.05)	(0.05)
Treat * Period 1 NO_x	-0.02*			-0.02	-0.02		-0.01
	(0.01)			(0.01)	(0.02)		(0.02)
Treat * Income		-0.03		-0.03		-0.16	-0.17
		(0.18)		(0.18)		(0.15)	(0.16)
Treat * %Minority			0.03		0.01	-0.11	-0.13
			(0.25)		(0.25)	(0.18)	(0.20)
Period 1 NO_x	-0.06	-0.07	-0.09*	-0.07	-0.09*	-0.09**	-0.09**
	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)	(0.03)	(0.04)
Income		-0.001		-0.001		-0.13	-0.13
		(0.096)		(0.097)		(0.09)	(0.10)
%Minority			-0.20		-0.19	-0.32	-0.31
			(0.20)		(0.21)	(0.24)	(0.25)
R^2	0.12	0.14	0.14	0.14	0.14	0.15	0.14

Notes: See notes in Table 7. Here the sample size is 838 and 1005 in Panels A and B, respectively.

Table A7: Average Treatment Effect using Nearest Neighbors Matching and Imputed Emissions Observations

Panel A: Change in NO_x Emissions between Periods 1 and 4.

Description	Levels	Logs	RECLAIM facilities	Controls
Base specification	-20.59**	-0.25***	212	1222
	(7.63)	(0.09)		
Specification w/ imputed emissions	-11.76**	-0.13**	373	5,324
	(4.76)	(0.06)		

Panel B: Change in NO_x between Periods 2 and 3.

Description	Levels	Logs	RECLAIM facilities	Controls
Base specification	-8.29**	-0.26***	255	1577
	(3.85)	(0.06)		
Specification w/ imputed emissions	-10.68*	-0.21***	359	5, 324
	(6.41)	(0.06)		

Notes: Panels report results for the base specifications. See Table 4 for notes.

Table A8: Deviations from Initial Permit Allocation

Panel A: Change in NO_x Emissions between Periods 1 and 4.

Variable	(1)	(2)	(3)
Change in permit allocation	0.71***	0.69***	0.03
	(0.07)	(0.07)	(0.13)
Income		0.88	0.53
		(1.72)	(1.07)
% minority		1.02	0.48
		(0.93)	(0.51)
Period 1 NO_x			-0.69***
			(0.10)
Constant	2.40	-81.13	-38.52
	(7.32)	(100.83)	(61.38)

Panel B: Change in NO_x between Periods 2 and 3.

Variable	(1)	(2)	(3)
Change in permit allocation	0.51***	0.49***	-0.11
	(0.08)	(0.08)	(0.19)
Income		-0.46	-0.31
		(0.59)	(0.47)
% minority		-0.06	-0.10
		(0.34)	(0.26)
Period 1 NO_x			-0.27***
			(0.08)
Constant	-3.24	16.09	15.19
	(2.15)	(36.27)	(28.96)

Notes: Panels report results for the base specifications. See Table 4 for notes.

Table A9: Period 2 to 3 Results when Including Electric Facilities.

Panel A: Robustness to Control Group (Table 6).

Control Group	Levels	Logs	RECLAIM facilities	Controls
Table 4 Results	-6.18 (5.06)	-0.16*** (0.06)	252	1,493
Exclude L.A. Facilities	-7.15 (5.47)	-0.19*** (0.07)	260	877
Exclude Northern CA	-12.75** (5.20)	-0.26*** (0.07)	268	1090
Severe Non-Attainment Only	-11.94** (5.34)	-0.15** (0.07)	257	541
Single Facility Only	-12.87** (5.56)	-0.20*** (0.06)	266	1027

Panel B: Environmental Justice Results (Table 7).

	1	2	3	4	5	6	7
Treatment	-4.86	-5.79	-4.03	-5.82*	-4.19	-4.48	-4.94
	(3.10)	(3.27)	(3.28)	(3.15)	(3.72)	(2.67)	(3.21)
Treat * Period 1 NO_x	-0.04			-0.04	-0.04		-0.04
	(0.02)			(0.03)	(0.03)		(0.03)
Treat * Income		-0.51*		-0.41*		-0.60	-0.69*
		(0.23)		(0.19)		(0.37)	(0.32)
Treat * %Minority			0.18		0.08	0.01	-0.12
			(0.16)		(0.13)	(0.22)	(0.18)
Period 1 NO_x	-0.32***	-0.32***	-0.31***	-0.32***	-0.31***	-0.32***	-0.31***
	(0.07)	(0.06)	(0.06)	(0.08)	(0.07)	(0.07)	(0.08)
Income		0.39		0.40		0.15	0.25
		(0.27)		(0.25)		(0.36)	(0.33)
%Minority			-0.22		-0.18	-0.19	-0.12
			(0.12)		(0.12)	(0.15)	(0.15)
R^2	0.40	0.36	0.37	0.37	0.37	0.37	0.37