WHAT DO EMISSIONS MARKETS DELIVER AND TO WHOM? EVIDENCE FROM
SOUTHERN CALIFORNIA'S NOX TRADING PROGRAM

Meredith Fowlie
Stephen P. Holland
Erin T. Mansur

What Do Emissions Markets Deliver and to Whom? Evidence from Southern California's
NOx Trading Program
Meredith Fowlie, Stephen P. Holland, and Erin T. Mansur
NBER Working Paper No. 15082
June 2009
JEL No. L5,Q52,Q53,R20

ABSTRACT

A perceived advantage of cap-and-trade programs over more prescriptive environmental regulation is that enhanced compliance flexibility and cost effectiveness can make more stringent emissions reductions politically feasible. However, increased compliance flexibility can also result in an inequitable distribution of pollution. We investigate these issues in the context of Southern California's RECLAIM program. We match facilities in RECLAIM with similar California facilities also located in non-attainment areas. Our results indicate that emissions fell approximately 24 percent, on average, at RECLAIM facilities relative to our counterfactual. Furthermore, we find that observed changes in emissions do not vary significantly with neighborhood demographic characteristics.

Meredith Fowlie
University of Michigan
Ford School of Public Policy
5224 Weill Hall
Ann Arbor, MI 48109-3091
and NBER
mfowlie@umich.edu

Stephen P. Holland
Bryan School of Business and Economics
University of North Carolina, Greensboro
P.O. Box 26165
Greensboro, NC 27402-6165
and NBER
sphollan@uncg.edu

Erin T. Mansur
Yale School of Management
135 Prospect Street
P.O. Box 208200
New Haven, CT 06520-8200
and NBER
erin.mansur@yale.edu

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 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 Europe, the European Union's greenhouse gas Emission Trading System has brought emissions trading 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? An important perceived advantage of market-based approaches over CAC is that they can, in some circumstances, deliver more significant public health and environmental benefits; 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; Tietenberg, 2006). Second, some have expressed concern that a reliance on permit markets (versus prescriptive regulations and standards) to coordinate pol-

---

[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.

lution abatement activity can lead to environmental injustice (Kaswan, 2008; Vandenbergh and Ackerly, 2007). 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. To the extent that there are pre-existing inequalities in the distribution of environmental risks, these inequalities can be exacerbated.[2]

We assess these questions 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 the 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. Specifically, 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 a broad and diverse population of sources, making it particularly relevant to future trading programs that will need to be more heterogeneous to achieve increasingly aggressive air quality and climate goals. Illaudably, it was also the first emissions trading program to

---

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

be challenged on the grounds of environmental injustice and noncompliance.

This 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 figure 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 arrive at very different conclusions. Whereas some regard the program as a clear success (Stavins, 2007), others see a "spectacular" failure (Green *et al.*, 2007).[3]

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.[4] We exploit the participation requirements of the RECLAIM program in order to construct semi-parametric estimates of program impacts. Emissions trajectories at RECLAIM facilities are compared

---

[3]Stavins (2007) summarizes domestic experience with emissions trading and reports that the RECLAIM program has generated significant environmental benefits "with NOx 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."

[4]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 *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.

with those at similar California facilities outside of RECLAIM. One important advantage of this approach is that it generates counterfactual emissions estimates that are free of 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 socio-economic 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). We combine semi-parametric matching methods with parametric regression techniques. This allows us to examine correlations between RECLAIM-induced emissions changes and socio-economic neighborhood characteristics with unprecedented precision.

Our results indicate that emissions at RECLAIM facilities have fallen by more than 20 percent, on average, relative to the control facilities (*i.e.* similar California facilities subject to command and control regulation). These results are generally robust to alternative estimation methods, functional form specifications, and different control group composition. 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 environ-

mental 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: 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.[5] The South Coast Air Quality Management District (SCAQMD) is the government agency responsible for regulating air pollution in the Los Angeles basin.[6] In 1989, SCAQMD introduced an aggressive set of rules and standards for stationary sources. Industry representatives fiercely opposed these 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 NOx standards were significantly revised. Because SCAQMD was much further from

---

[5]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). Estimates of health-related losses due to the poor environmental quality in the region were approaching $10 billion per year (Hall *et al.*, 1992).

[6]Figure 1 shows the district's boundaries.

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.[7]

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 so-called "command-and-control" 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.[8]

SCAQMD responded by replacing over 40 prescriptive rules, which had been so opposed by industry, with a market-based emissions trading program: RECLAIM.[9] This program was approved by state and federal regulators on the grounds that it would deliver emissions 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.

---

[7] Appendix A discusses the CAAA compliance requirements in more detail.

[8] 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 $(51.492, 182(g)5)$.

[9] 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.

At its inception, 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.[10] 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.[11]

A RECLAIM trading credit (RTC) confers the right to emit one pound of emissions within a twelve month period.[12] Figure 2 plots the aggregate allocation trajectory over time (the red line).[13] 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 to achieve by 2010.

Early on, most firms found they had an excess of credits (the blue line in Figure 2 represents aggregate tons of NO_x emissions).[14] The aggregate cap did not start to bind until

---

[10]Of these, 73% of these 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.

[11]Facilities could "opt-in" to the program. Facilities could not opt-out.

[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. Holland and Moore (2008) analyze this "overlapping cycles" feature of the permits.

[13]SCAQMD maintains a detailed database tracking all NOx permits and quarterly, facility-level emissions. RTC price data were obtained from Evolution Markets LLC. Section 4 includes a detailed description of these data.

[14]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 (Keohane *et al.* (forthcoming) discuss this literature). Under CAC, operating permits are generally valid for several years unless modifications trigger New Source Review (NSR), which can be contentious and costly. Under RECLAIM, traditional NSR regulations were replaced. There

1999 (SCAQMD, 2001). The figure 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). Because permits cannot be banked, impacts of the initial over-allocation were confined to the early stages of RECLAIM.

Figure 2 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.[15] 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, many generators without retrofitted NOx reduction equipment were required to run and hence exceeded their allocations significantly.[16] In May 2001, the RECLAIM rules were amended to reduce the RTC demand from power producing facilities and to stabilize the sharply increasing NOx RTC prices. The rule amendments (Rule 2009) isolated fourteen power producers from the rest of the RECLAIM market. These facilities were not allowed to trade RTC's. Instead,

---

is some evidence that RECLAIM's internal NSR procedure may have facilitated NSR for the RECLAIM facilities. For example, only "BACT" is required and necessary offsets can be demonstrated with RTCs. Furthermore, the 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.

[15]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).

[16]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.

they were required to offset excess emissions at a price of $14,000 per ton. In addition, these facilities were required to submit plans to install the "best available" control technologies on all existing power generating units by the end of 2004.[17]

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 have required an unprecedented level of program evaluation and oversight. In general, emissions trading program evaluation has proven to be particularly challenging (Ellerman, 2003; Stavins, 1998; Tietenberg, 2006). 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 overseeing RECLAIM remain committed to evaluating the emissions impacts of the program.

Unresolved disagreements about what constitutes an appropriate measure of counterfactual emissions has resulted in a plurality of opinions regarding RECLAIM's overall per-

---

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

formance. Appendix B summarizes some of the contradictory evidence provided by past program evaluations and reports. After fifteen years of program evaluations, the emissions impacts of RECLAIM *vis a vis* the subsumed CAC rules remain controversial.[18] 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). Disagreements surrounding counterfactual emissions beget disagreements about the overall success of the RECLAIM program. The proverbial jury is still out.

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). A recent review of the literature concludes that the evidence is mixed (Shapiro, 2005).

Environmental justice advocates have historically opposed market-based approaches to environmental protection (Bansal and Davis, 1998; Drury *et al.*, 1999). Kaswan (2008)

---

[18]Some studies have constructed counterfactual estimates using *ex ante* expectations about economic trends and the control factors for subsumed rules (SCAQMD, Various years). Critics contend that this fails to account for unanticipated economic conditions or technological changes that affect emissions patterns with or without emissions trading. Other studies have tried to adjust *ex ante* predictions using *ex post* observed emissions trends (US EPA, 2001). This approach likely confounds exogenous and endogenous changes in emissions. See Appendix B for further discussion of these regulatory studies.

[19]See, for example, Brown, 1995; GAO, 1983; US EPA, 1992.

provides a detailed discussion of the perceived tensions between environmental justice and emissions trading. The most common criticism 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 can purchase permits instead of reducing emissions, it is possible for pollution concentrations to flow into areas where poor or minority populations live, thereby exacerbating any pre-existing inequalities in the distribution of environmental risks. On the other hand, market-based programs could mitigate pre-existing environmental justice problems. If relatively dirty facilities with low marginal abatement costs are disproportionately located in traditionally disadvantaged neighborhoods, a well functioning 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 criticized of any emissions trading program with respect to environmental justice issues (Chinn, 1999; Drury *et al.*, 1999; 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. Damages from NO_x emissions can vary significantly across space.[21] Finally, the

---

[20] An EPA analysis of outcomes under other the Acid Rain Program finds no evidence of disproportionately adverse impacts on minority or low-income populations (EPA, 2005). A more recent paper finds that minority groups receive a disproportionate share of the net benefits from the program (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

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 credits.[22] Although RECLAIM was only indirectly involved in this high profile controversy, the program has since been associated with environmental injustice.[23]

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 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.

---

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.

[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.

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 assume that there are two regulatory states to which the California's industrial NOx 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 ith facility is 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 average treatment effect on the treated (ATT):

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

where t' represents any year following the introduction of the RECLAIM program and α_{TT} measures the average effect of the RECLAIM program on annual facility level NOx emissions.[24]

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

---

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

of the program can be used to identify $E[Y_{it'}(1)|D_i = 1]$. Note that $[Y_{it'}(0)|D_i = 1]$ is unobservable.

The RECLAIM program applies only to major sources located within SCAQMD (see Figure 1). Thousands of California facilities located outside the Los Angeles air basin are subject to more traditional CAC. Furthermore, hundreds of smaller emitters within SCAQMD remain subject to more traditional CAC rules. These incomplete program participation requirements provide us with a potential comparison group. We use econometrically adjusted emissions outcomes of these non-participants to estimate the unobserved potential outcomes.

The simplest and most naive estimate of α_{TT} is obtained using an unconditional differences-in-differences estimator. This estimator may be biased if variables 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 will condition on observable covariates.

3.1.1 Regression-based conditioning strategies

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

$$Y_{i1} - Y_{i0} = \beta' X_i + \alpha D_i + \varepsilon_i, \tag{2}$$

where X_i is a vector of observable covariates that vary across observations. This implicitly assumes that the variables in X are exogenous to treatment status. In our case, these variables will include facility level emissions observed prior to the introduction of the RE-

CLAIM program, industry classification, county-level attainment status, and pre-determined zip code-level economic and demographic indicators. 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 are regions of the covariate space where there is only limited overlap in the distributions of X across the treatment and control groups, the imputed missing outcomes will rely heavily on extrapolation. If functional form assumptions are incorrect, estimates may be biased. 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 biases, we turn to semi-parametric matching estimators.

3.1.2 Semi-parametric conditioning strategies

Matching estimators, which are used extensively in non-experimental program evaluation, are an extension of standard regression approaches. Our general approach follows Heckman *et al.* (1997, 1998) who introduce the following generalized DID matching estimator:

$$\widehat{\alpha_{DID}} = \frac{1}{N_1} \sum_{j \epsilon I_1} \{(Y_{j1}(1) - Y_{j0}(0)) - \sum_{k \epsilon I_0} w_{jk}(Y_{k1}(0) - Y_{k0}(0))\}. \tag{3}$$

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.[25] The participants are indexed by j;

---

[25]This assumes a common support which we explore in the next section.

the non-participants are indexed by k. The weight placed on individual k when constructing

the counterfactual estimate for treated facility j is w_{jk}. Different DID matching estimators

adopt different approaches to defining the weights w_{jk} used to scale the contribution of each

non-participant. In general, when the observable characteristics, X_i, of an untreated unit

k are closer to the characteristics of a treated facility j (relative to other facilities in the

control group), the untreated unit k is weighted relatively more heavily in the construction

of a counterfactual estimate for unit j. In what follows, we will be emphasizing a nearest

neighbor estimator and a propensity score-based reweighting estimator.

One clear advantage of matching estimators over standard linear regression techniques

is that they eschew parametric assumptions about the relationship between the outcome

variable and the covariates in X.[26] Additionally, whereas treatment effect heterogeneity may

bias OLS estimates of average treatment effects, non-parametric matching estimators are

unbiased because they impose no restrictions on individual treatment effects.

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 county characteristics (such as historic emissions, industry classification, county

attainment status), is the same among participating and non-participating facilities. If this

---

[26] Smith and Todd (2005) compare results of matching and regression estimates using the same set of covariates and demonstrate that avoiding functional form assumptions can be important in reducing bias. However, if the true outcome equation is linear as in (2), failing to impose linearity will reduce the efficiency of our estimates.

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 order to interpret [3] as an estimate of the effect of RECLAIM vis a vis the CAC regulation that would have prevailed in the SCAQMD absent RECLAIM, it must be the case that the emissions trajectories of RECLAIM facilities and the matched control facilities would have followed parallel trajectories absent RECLAIM. This implies a somewhat stronger version of the unconfoundedness assumption: 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.

Some of these assumptions and identifying conditions 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 pose more of a challenge. Most importantly, unconfoundedness is not directly testable in principle. However, we are able to conduct indirect tests. The weaker unconfoundedness assumption implies that $Y_{it'}(0)$ will be distributed similarly within sub-populations that are

17

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 the two control groups.

The stronger unconfoundedness assumption requires that the control treatment resembles the regulations that RECLAIM facilities would have been subjected to absent RECLAIM. In Appendix A, we assess the plausibility of this assumption by looking at the ozone concentrations reductions mandated in SCAQMD vis a vis other California air basins over the study period.

3.3   Treatment effect heterogeneity

In addition to identifying the average effect of RECLAIM on facility-level emissions, we are interested in investigating whether treatment effects appear to vary systematically across facilities located in neighborhoods with different socio-economic characteristics. Standard DID matching estimators discard potentially useful information about the joint distribution of the outcome variable Y_{it} and the covariates in X that is unrelated to the treatment indicator D_i. We combine matching with parametric regression so as to investigate what this discarded information can tell us about how treatment effects vary with observable neighborhood characteristics.

All of the matching estimators we use can be interpreted as nonparametric weighted regression equations that weight treatment effects according to the proportion of treated facilities associated with each value of X. We specify a weighted regression model that allows the treatment effect to vary systematically with a subset of variables in X_i (denoted

18

Z_i):

$$Y_{i1} - Y_{i0} = \beta' X_i + \alpha D_i + \theta' Z_i D_i + \varepsilon_i \tag{4}$$

To investigate the extent to which emissions trading has exacerbated (or mitigated) environmental injustice vis a vis CAC regulations, we define Z_i as historic emissions, zip code median income, and zip code percent minority. As a robustness check, we extend Z_i to include measures of regulatory variables, macroeconomic shocks, and industry fixed effects (see Appendix C).

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 the local districts. Our primary data comes from this database which also includes information on industry classification beginning in 1990. We use addresses, geocodes, and industry classifications to ensure a consistent coding of facilities across our panel.[27] We also use separate emissions data from RECLAIM to verify the emissions reported to the ARB database.[28]

In the regressions, we use zip code level demographics data from the year 2000 Census

---

[27]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.

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

Summary File 3.[29] The data include median household income in 1999 (variable p53) and population by ethnicity and race (variable p7). We construct a measure of percent minority as the percent of the zip code's total population that is either non-Hispanic black or Hispanic. See Figure 1 for the spatial distribution of this measure.

Figure 3 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 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 for which emissions data are available (2004-2005). The sample includes all facilities reporting positive emissions in each period. Overall, annual facility-level

---

[29]See http://factfinder.census.gov decennial census data sets.

emissions are significantly larger among RECLAIM firms *vis a vis* the comparison group. Average emissions among RECLAIM facilities fell 70 percent between period 1 and period 4. We will subsequently refer to these unadjusted differences as "baseline" measures of the average effect of RECLAIM on facility-level emissions.[30] This table also illustrates that annual emissions are distributed differently across RECLAIM facilities and others in all periods.

Table 2 examines the distributions of treated and control facilities within the eight 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% of emissions are from firms in other four-digit SIC codes. Regarding the distributions of the treatment and control groups, there are several points worth noting. First, the industry-specific means of the RECLAIM facilities are not always larger. In fact, for seven of these eight industries, the RECLAIM facilities are smaller on average than the control facilities in the same industry. Second, the common support assumption is satisfied for most of these eight industries in the sense that the intervals containing the minimums and maximums of the distributions are overlapping.[31] Third, the table suggests a word of caution about our matching strategy. Ideally we would like to match each treated facility with a large number of control facilities to average out idiosyncratic shocks in our estimate of counterfactual emissions. However, the number of control facilities within each industry suggest that some industries have a limited number of

---

[30]When the sample is limited to just those facilities who were in compliance with RECLAIM over the entire study period, average emissions fell from 72.2 to 31.5 (a similar percentage reduction).

[31]Of the fourteen relevant minimum or maximum comparisons, eleven are satisfied. Note that this is rather a weak test, since it may be driven by outliers. Ideally, we would like sufficient density over the common support.

potential controls.[32]

Our panel of facility-level data is unbalanced. Of those facilities reporting emissions after the RECLAIM program was introduced, only 60 (45) percent of RECLAIM firms and 32 (13) percent of control facilities report emissions in the first three (four) periods. Facility-level emissions data in a given period may be missing for a number of reasons, including firm entry and exit, errors in the data, or simply a facility's failure to report emissions in a given period. On average, treated facilities reporting emissions in all periods were larger emitters in period 1, although not significantly so.[33] Section 5 discusses sample selection issues in more detail.

In Table 3, we calculate the change in tons of NOx emissions exposure per capita. We use census tract data from 2000 to determine the number of people within a demographic group who live in a tract whose centroid is within 1/2, 1 and 2 miles of each facility.[34] For each demographic group, we aggregate the change in emissions across people in that group and facilities and then normalize by the group's total population in the LA area. The first three columns report the actual change in emissions from period 2 to period 3. We find that all groups experienced a reduction of emissions.[35] Within a 1/2 mile, the group that saw the largest actual reductions was high income Asians while the group that saw

---

[32]A limited number of controls may lead to poor match quality. We discuss match quality below.

[33]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. Among the control group, these averages are 87.8 tons and 52.9 tons, respectively.

[34]This technique is similar in spirit to Banzhaf and Walsh (2008) but is more closely related to earlier techniques used in the environmental justice literature.

[35]Standard errors are computed by assuming the facility-level changes in emissions are iid. Simple T-tests find that most groups within 0.5 miles had statistically significant reductions in emissions from period 2 to period 3.

the smallest reductions was middle income Asians. These results change depending on the distance from facilities. In section 5, we look at the relative changes in emissions, namely the actual changes less the changes from our counterfactual control group, that we attribute to RECLAIM trading.

5 Results

In this section, we present estimates of treatment effects for the long-term view of the program as well as for the window surrounding the price spike when we expect to see emissions reductions from RECLAIM. The estimates are broadly similar across our three estimators and are subjected to several falsification tests. We then present estimates of heterogeneous treatment effects and environmental justice and discuss selection issues.

5.1 Differences-in-differences estimates

Table 4 summarizes our main results. The control group is restricted to facilities located in counties that, during 1990 and 1993, were not in attainment with the 1-hour ozone NAAQS standards. The columns on the right pertain to differences constructed using facility-level annual emissions data. The columns on the left construct differences using log transformed values. The top panel (Panel A) presents a more long term view of program impacts; facility-level emissions in period 4 are subtracted from period 1 emissions. Panel B compares facility-level emissions immediately before and after the emissions cap began to bind in aggregate. Recall that some electricity generators did not comply with RECLAIM when permit prices spiked in period 3. We thus exclude these electricity producers from this period

2-3 comparison so as to focus on those facilities that were participating in (and complying with) the cap-and-trade program.

We use a simple linear regression framework to generate DID estimates that condition on observable facility characteristics that are that are likely to be correlated with unobservable determinants of emissions. Covariates include industry fixed effects and NOx emissions in period 1. Conditional on the identifying assumptions of this linear regression model, we find that RECLAIM program effects on emissions changes over period 1 to 4 are highly significant. Emissions changes immediately following the cross over period (from period 2 to 3) are also statistically significant in logs. The period 1 NOx 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

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 (Morgan and Harding, 2006). In this section, we present results using two semi-parametric matching estimators: a bias adjusted "nearest neighbor" estimator and the so-called "double robust" estimator.

Nearest neighbor 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. 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.[36] 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. Following Abadie and Imbens (2006), we augment our covariate matching estimation with a regression-based bias adjustment so as to eliminate bias introduced by poor match quality. After matching the treated facilities with m nearest neighbors, within pair differences are bias-adjusted using a parametric regression of the control outcome on X.[37]

Table 4 presents these nearest-neighbor estimates. Matching covariates include pre-treatment (*i.e.*, period 1) NO_x emissions, four-digit SIC code, and county attainment status. These variables are likely to be correlated with unobservable determinants of facility-level emissions, including production technology characteristics, firm size, demand for the products produced by the facility. Matching is exact on the categorical variables. Standard error estimates are constructed using the variance formula of Abadie and Imbens (2006). Leave-one-out cross-validation methods are used to select the neighbor parameter m. Appendix tables A1 and A3 demonstrate that our results are not overly sensitive to the choice of m or the bias adjustment.

---

[36]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. The bias of this estimator is of order $N^{1/z}$, where z is the number of covariates in X.

[37]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 C.

When the overall change in emissions between the pre-treatment period (period 1) and period 4 is used as the dependent variable, the NN estimate, -17.79 tons per year, is statistically significant at the five percent level. This is approximately 17 percent of the average annual emissions at RECLAIM firms 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 firms versus matched control facilities. Because the estimated average annual percentage reduction is somewhat larger than the estimated average reduction (in levels) expressed as a percentage of period 1 emissions, this suggests that percentage reductions are somewhat larger at smaller firms.

Making the period 3 period 2 comparison, the NN estimate is -9.42 and statistically significant. This represents 13 percent of the average annual emissions at RECLAIM firms in period 2. The SATT estimate estimated using log transformed data is 0.26.

Propensity score matching

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 units, reweighting estimators outperform nearest neighbor or kernel matching in

finite samples (Busso *et al.*, 2009). In our case, 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. The process that determined whether a facility located inside or outside of SCAQMD is not directly observed. Consequently, the propensity score is an unobserved statistical object that needs to be estimated. 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 *et al.,*1995; Robins and Ritov, 1997). By combining propensity score matching with regression, we can reduce bias introduced by poor match quality.[38] Table 4 summarizes the main results. These SATT estimates are larger in absolute value as compared to the NN

---

[38]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).

estimates and somewhat noisier.

Robustness of semi-parametric matching estimates

Our empirical 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-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 firms are significantly greater on average as compared to the matched controls. In order to interpret these estimates as an unbiased measure of RECLAIM program impacts, some important assumptions must hold. We investigate the plausibility of these assumptions here.

First, our analysis is predicated on the assumption that the emissions trajectories of firms in the control group are representative of the emissions trajectories that would have been observed at similar RECLAIM facilities had RECLAIM not been implemented. Put differently, we assume that facilities with the same observed covariates should follow parallel emissions trajectories if they are subject to CAC regulation in a non-attainment county. Discontinuities in RECLAIM participation requirements allow us to indirectly test this unconfoundedness assumption. We do this by means of a false experiment. 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, the estimated "treatment effects" in this false experiment should not be statistically significantly different from zero.

Table 5 summarizes the results from this experiment. We find that the change in the average emissions (in levels) 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 similar facilities located in other California air basins follow parallel paths. This is consistent with our unconfoundedness assumption.

Differences in percentage changes in emissions (*i.e.*, the estimates obtained using log-transformed emissions data) tell a somewhat different story. In the more long run comparison, estimated effects of being located in SCAQMD (yet exempt from RECLAIM) versus located in another California non-attainment area are negative and substantially smaller in absolute value than the estimates reported in panel A. The double robust estimate is statistically significant, suggesting that the percentage reduction in emissions observed over this ten year period was larger among low-emitting firms subject to SCAQMD CAC regulation as compared to small emitters located in other non-attainment areas of California. We find the reverse is true in the period 2-period 3 comparison. In sum, although results using untransformed data generally support the unconfoundedness assumption, results generated using log-transformed data are less clear

One might also be concerned that, in response to the introduction of the RECLAIM program, production and associated emissions moved from RECLAIM facilities to those

exempt from the program. This would bias upwards our counterfactual emissions estimates, and exaggerate our estimates of program impacts. To investigate this issue, we experiment with using different subsets of the control group to identify the sample average treatment effect. In one exercise, we exempt Los Angeles facilities from the control group. In another exercise we include only Northern California facilities. Table 6 summarizes these results. Our point estimates are not substantially impacted. Our nearest neighbor estimates remain highly statistically significant. Our double robust estimates are quite noisy, in part due to the smaller sample size, and are harder to statistically distinguish from zero.

The final row of table 6 report SATT estimates obtained using only data from facilities in severe (versus moderate) non-attainment areas as controls. Estimated program effects are substantially larger in absolute value. This result is consistent with the idea that the political palatability of emissions trading programs *vis a vis* CAC regulations makes it feasible to introduce more aggressive environmental standards. In the early 1990s, California counties that had historically suffered from more severe air quality problems had more strict regulations in place (Figure A1). In these counties, industrial stakeholders would presumably have be more resistant to the introduction of more stringent emissions regulations, as compared to facilities in regions where the marginal costs of air quality improvements was relatively low. If this were the case, we would expect that the effects of market-based regulations (in terms of emissions reductions achieved) would be greater when comparisons are made against counties facing similar (versus less) political opposition to increased regulatory stringency.

5.3 Environmental justice

The literature on environmental justice is concerned that polluters in poor or minority neighborhoods may have been able to avoid reducing emissions by purchasing permits (Bansal, 1998). In this section, we ask whether trading resulted in changes in emissions that are correlated with the demographics of the polluters location. From Tables 5 and 6, we conclude that emissions reductions were significantly greater among facilities in the RECLAIM program relative to the CAC counterfactual we construct. We now investigate whether these RECLAIM program impacts were more or less felt by traditionally disadvantaged neighborhoods in the SCAQMD.

In Tables 7 and 8, we use non-parametric matching methods to construct a counterfactual change in emissions for each regulated (and compliant) facility and then regress the difference between the actual change and the predicted change on a number of covariates measuring environmental justice. We present results from several alternative specifications for both the overall effects of the program. Table 7 presents results from the period 2 and 3 comparison; Table 8 investigates the period 1 to 4 comparison. Results are generated using both nearest neighbor and propensity score matching. The results are qualitatively similar.

Specification 1 includes only a treatment dummy, period 1 NOx emissions, and an interaction between these two covariates as explanatory variables. This specification is similar to the linear regression equation introduced in section 5.1, although less restrictive because it includes group (versus industry) fixed effects and allows the *Period 1 NOx* coefficient to vary across the treatment and control group. The *Treat * Period 1 NOx* coefficient is statistically

significant. Conditional on group fixed effects and the *Period 1 NOx* variable, this is the effect of historic emissions on the change in emissions at RECLAIM facilities relative to that effect for all firms. For example, in Panel A of Table 7, the coefficient of -0.06 implies that for each additional ton that a facility emitted in period 1, it reduced emissions between periods 2 and 3 by an additional 0.06 tons relative to the control group.

Specifications (2) through (7) include income and race covariates. Specification (7) includes historic NOx emissions and all demographic variables. In this specification, neither median income nor percent minority are statistically significant predictors of the changes in emissions in RECLAIM. The same is true within the control group; neither income nor race are significant determinants of changes in emissions. The change in emissions per capita reported in Table 3 show similar findings. For both the actual change in emissions as well as the relative change (the difference between the actual change and the change for the control group), emissions fall for almost all demographic groups. Furthermore, the changes do not have a clear pattern based on income or race.

In all specifications containing *Treat * Period 1 NOx*, the variable 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 RE-CLAIM and other facilities in more detail. As with Table 7, we use the $\widetilde{N}$ sample resulting from Table 5.[39] In Figure A3, we smooth the observations, separately for RECLAIM and for

---

[39]In particular, we use the results from the top row, period 1 to 4, $m = 3$. For each treated observation, we construct a measure of what the change in emissions would have been 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 C).

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.

In all specifications, the *Period 1 NOx* coefficient is statistically significant. Ideally, our within group matching on historic NOx emissions would be perfect and the *Period 1 NOx* 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. In light of these results, we are concerned about bias introduced by poor match quality. All of our matching estimation incorporates a parametric adjustment to mitigate this bias (Abadie and Imbens, 2006).

Appendix C explores the robustness of the results presented in Table 7. The results are robust to including noncompliant facilities and to the inclusion of other variables like regulatory variables, macroeconomic shocks, and industry fixed effects (Table A4). The demographic variables are not significant if subsets of the demographic variables are excluded (Table A3).

5.4   Selection issues

We are also concerned about the unbalanced nature of our panel. In order to interpret our results as a measure of the average effect of the RECLAIM program on facility-level emissions, the subset of facilities reporting emissions should be representative of the populations of

RECLAIM and control facilities, respectively. The group of facilities reporting emissions changes over our sample period. Non-random selection into the data could imply that our estimates of the average program impacts among the population of RECLAIM facilities may be biased. Intuitively the direction of selection 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 does make more stringent emissions reductions more politically feasible, we might expect to see facilities with relatively high abatement costs exiting the SCAQMD with higher frequency, biasing our results in the opposite direction.

One common approach to addressing sample selection issues involves augmenting the outcome equation with an inverse Mills ratio estimated using a parametric selection model (Heckman, 1979). We use a probit model to represent the selection process. When we include facility-specific inverse Mills ratios as an additional explanatory variable in our parametric regressions, the estimated coefficient on this ratio is not statistically significant. Although this suggests that selection bias does not contaminate our estimates, we report these results with some important caveats. First, there are no variables included in the selection equation that can be credibly excluded from the outcome equation. Moreover, we offer no intuitive motivation for the parametric assumptions underlying our model of the selection process.

Another approach to assessing whether the unobserved factors that drive the selection process are biasing our results is to investigate whether our results change significantly if we

generate our results using only those units that are observed over the entire sample. Our estimates of program impacts over the full time period, and in the late (post cross-over) period are not significantly different.

6   Conclusions

In this paper, we exploit some unique design features of the RECLAIM program in order to bring new and illuminating evidence to bear on two important questions. First, did emissions reductions at facilities subject to Southern California's RECLAIM 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. However, during the period of great permit price volatility, several facilities did not comply with the regulation. When these facilities are included in our analysis, we find that RECLAIM did *not* reduce emissions relative to command and control during this volatile time. For those complying, emissions fell significantly more at large facilities, as is consistent with scale economies in abatement technology. We find no evidence that neighborhood demographic characteristics were insignificant determinants of their changes in emissions.

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] Bansal, Shipra and Sam Davis. 1998. "Holding our Breath: Environmental Injustice Exposed in Southeast Los Angeles," pp. 17-28.

[3] 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), pp. 843-63.

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

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

[6] 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.*

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

[8] 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.

[9] 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.

[10] 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.* Freeman, Jody and Charles Kolstad Eds. Oxford University Press.

[11] 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.

[12] 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.

[13] Executive Order 12898. 1994. "Federal Actions To Address Environmental Justice in Minority Populations and Low-Income Populations."

[14] Freedman, David A. and R.A. Berk. 2008. "On weighting regressions by propensity scores." *Evaluation Review* 32: 392-409.

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

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

[17] 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 2007.

[18] 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.

[19] 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. Morgenster, and Thomas Sterner, eds., RFF Press, Washington, DC.

[20] 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.

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

[22] 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.

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

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

[25] 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.

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

[27] 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.

[28] Keohane, Nathaniel O., Erin T. Mansur, and Andrey Voynov. Forthcoming. "Averting Regulatory Enforcement: Evidence from New Source Review," *Journal of Economics and Management Strategy.*

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

[30] 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.

[31] Moore, Curtis A. 2004. "RECLAIM: Southern California's Failed Experiment With Air Pollution Trading." *Clean Air Trust* http://www.cleanairtrust.org/release.121902.html, http://www.cleanairtrust.org/pdf/reclaim.pdf.

[32] 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.

[33] Morgan, Stephen L. and David J. Harding. 2006. "Matching Estimators of Causal Effects: Prospects and Pitfalls in Theory and Practice." *Sociological Methods & Research* 35(1): 3-60.

[34] 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.

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

[36] 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.

[37] 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," in *Environment and Health.*

[38] South Coast Air Quality Management District. 1998. *Three Year Audit and Progress Report*, Diamond Bar, California.

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

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

[41] 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.

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

[43] 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.

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

[45] Shapiro, M. (2005). Equity and Information: Information Regulation, Environmental Justice, and Risks from Toxic Chemicals. *Journal of Policy Analysis and Management*, 24(2), 373–398.

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

[47] 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.

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

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

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

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

[52] 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.

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

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

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

Figures and Tables
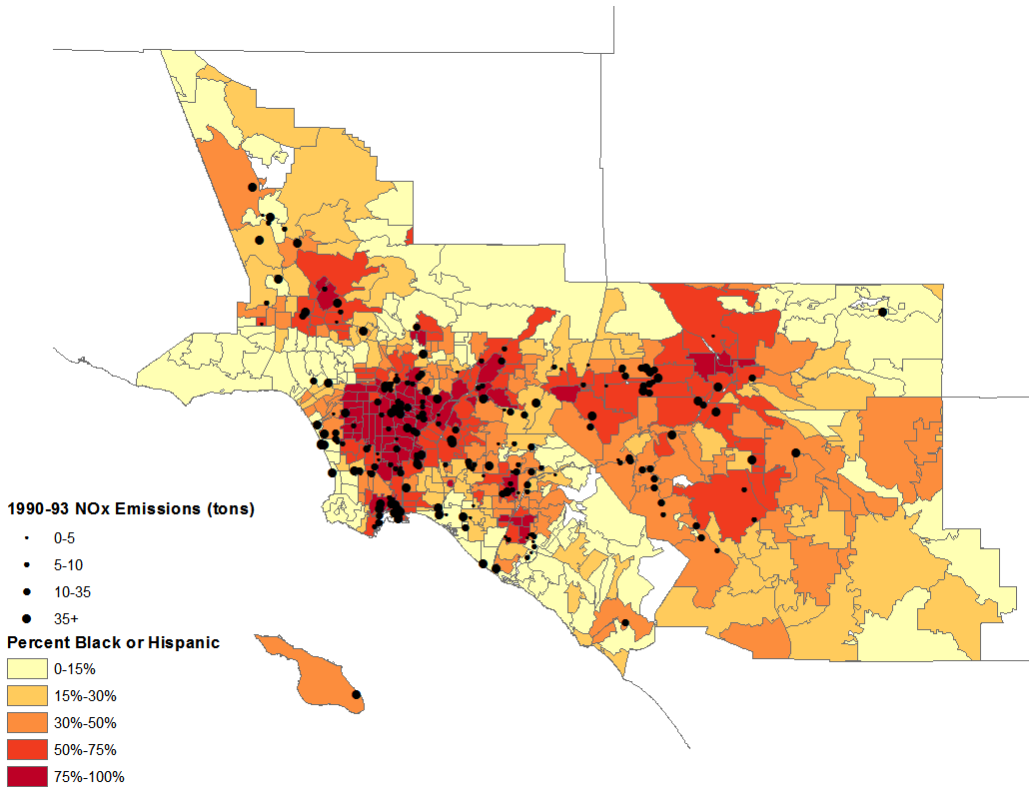
Figure 1: The South Coast Air Quality Management District.
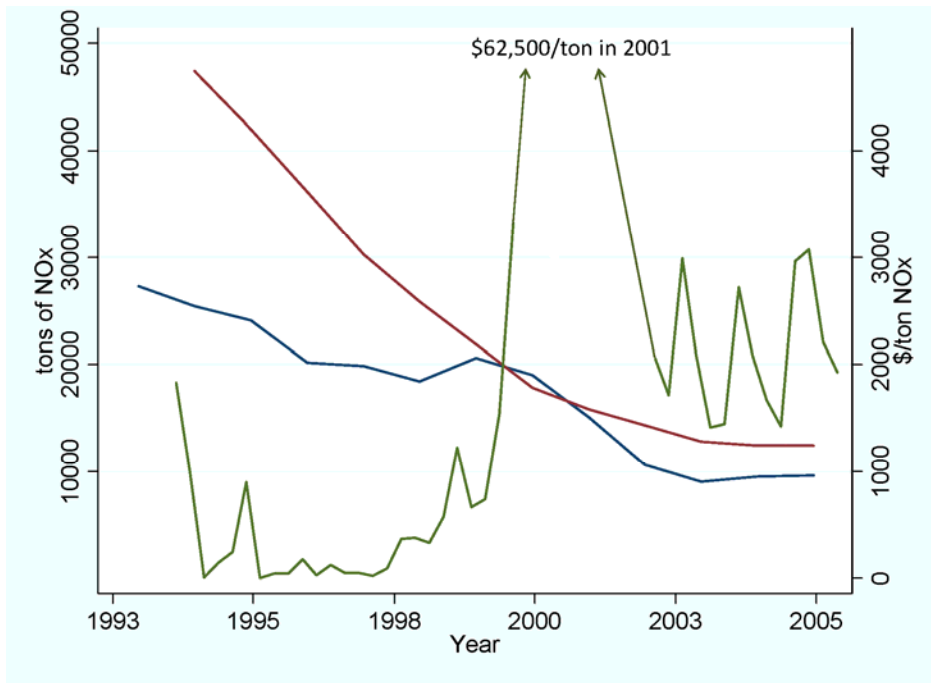
Figure 2: Trends in Nitrogen Oxides Emissions (blue), Allocations (red), and Permit Price (green).
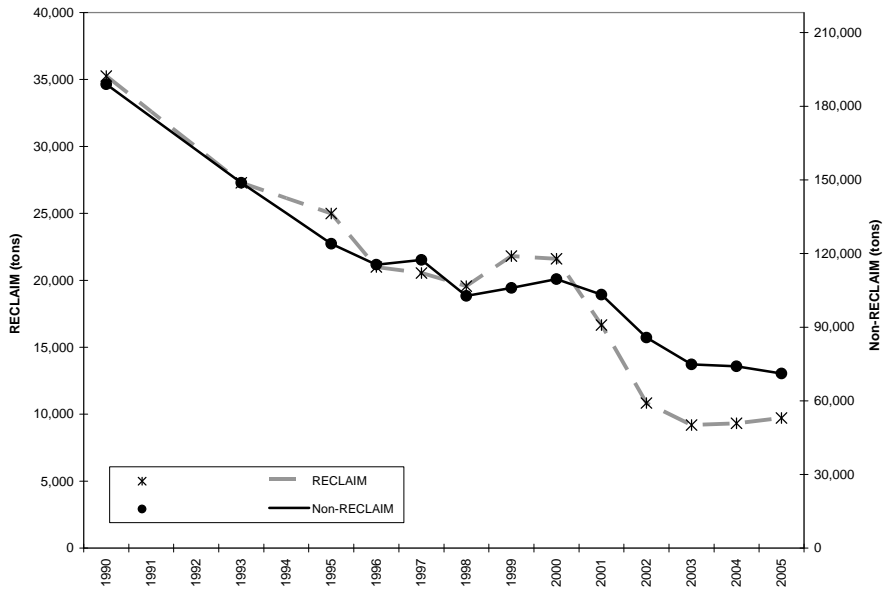
Figure 3: Total NOx Emissions in RECLAIM and in the rest of California.
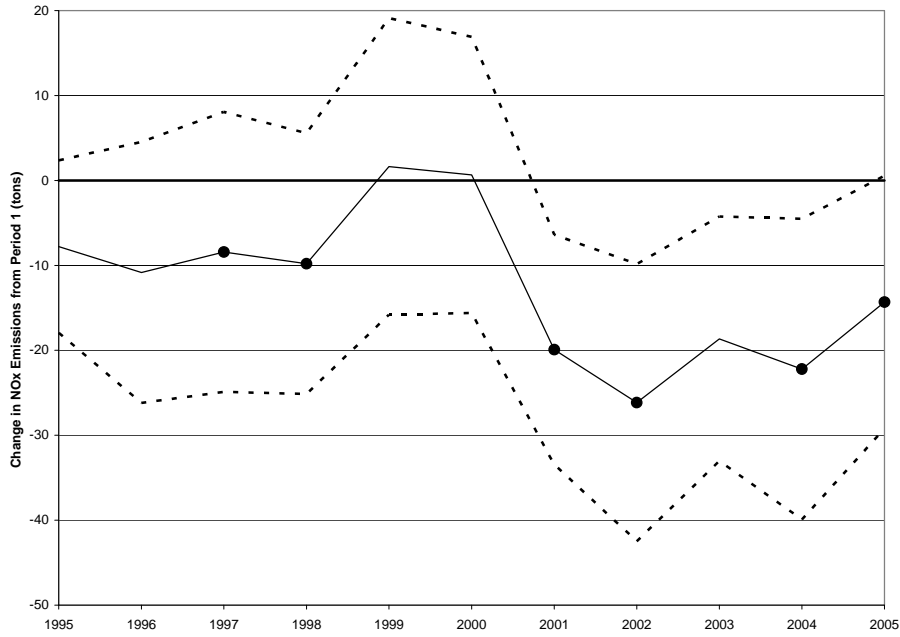
Figure 4: Average Cumulative Treatment Effect by Year (relative to Period 1 emissions), matching m=3.

Table 1: Summary Statistics of NOx Emissions.

Period	RECLAIM	Control	Total
Period 1	101.8	87.8	89.8
(1990-1993)	(304.4)	(394.0)	(382.4)
Period 2	62.7	67.7	67.0
(1997-1998)	(179.8)	(338.4)	(320.5)
Period 3	43.8	57.4	55.5
(2001-2002)	(125.4)	(309.6)	(290.5)
Period 4	30.8	44.9	42.9
(2004-2005)	(117.1)	(265.0)	(249.3)

Notes: We report the summary statistics on the balanced sample of facilities with positive emissions in all four periods. For RECLAIM facilities, this includes both those that complied with the regulation as well as those that did not. We report the mean tons of NOx emissions per facility (e.g., 101.8) as well as the standard deviation 304.4). There are 213 facilities in RECLAIM and 1273 in the control group. The control group is restricted to facilities 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.

Industry	SIC	RECLAIM Share	Treatment obs	min	mean	max	Control obs	min	mean	max
Petroleum Refining	2911	37.5%	10	9.8	880	2492	17	1.1	1046	4685
Electric Services	4911	23.9%	21	5.3	378	1370	82	0.1	407	5545
Crude Petroleum and Natural Gas	1311	7.1%	10	8.4	116	374	154	0.1	83	1945
Cement, Hydraulic	3241	4.1%	2	55.4	699	1342	9	501.4	1885	2978
Glass Containers	3221	3.8%	1	611.1	611	611	5	287.9	856	1169
Natural Gas Trans. and Distribution	4923	2.3%	8	7.0	85	214	4	8.9	474	1361
Paper Mills	2621	1.8%	6	4.5	83	422	5	0.3	121	392
Electric and Other Services Combined	4931	1.6%	4	20.9	107	188	64	0.1	335	5545
National Security	9711	0.9%	3	48.1	81	129	29	0.5	68	453
Industrial Inorganic Chemicals, NEC	2819	0.9%	5	8.3	31	80	10	0.3	223	2166
Steel Works, Blast Furnaces	3312	0.9%	3	5.7	103	237	4	2.3	20	74
Steam and Air-Conditioning Supply	4961	0.9%	7	6.6	39	108	2	16.0	55	94
Products of Petroleum and Coal, NEC	2999	0.8%	1	259.1	259	259	1	579.6	580	580
Total for Major Industries		87%	81	4.5	281	2492	386	0.1	296	5545

Notes: "RECLAIM Share" is the 4-digit SIC industry share of initial, period 1 NOx emissions. Most of the electric services did not comply (86% of the initial emissions). For facilities with positive emissions, we also report the summary statistics of tons of NOx emissions during period 1 for both treated and the control facilities.

Table 3: Change in Emissions (tons) per Capita by Demographic Group

Group	Actual change			Relative change		
	0.5 miles	1 mile	2 miles	0.5 miles	1 mile	2 miles
White, Low Income	-0.61 (0.30)**	-1.62 (0.64)**	-6.99 (2.22)***	-0.68 (0.39)*	-0.96 (0.66)	-4.27 (2.59)
White, Middle Income	-0.21 (0.11)*	-1.45 (0.72)**	-8.48 (4.41)*	-0.02 (0.14)	-0.56 (0.38)	-3.37 (1.33)**
White, High Income	-0.88 (0.58)	-3.12 (1.98)	-10.72 (5.34)**	-0.12 (0.19)	-0.97 (0.70)	-4.52 (2.12)**
Black, Low Income	-0.38 (0.17)**	-2.08 (0.93)**	-15.26 (6.44)**	-0.38 (0.20)*	-1.61 (0.94)*	-12.30 (5.83)**
Black, Middle Income	-0.21 (0.10)**	-1.39 (0.62)**	-14.34 (5.92)**	-0.07 (0.16)	-0.75 (0.53)	-9.08 (4.16)**
Black, High Income	-0.33 (0.20)	-8.80 (7.57)	-27.13 (18.69)	-0.07 (0.09)	-6.43 (6.08)	-21.53 (15.03)
Asian, Low Income	-0.90 (0.67)	-4.65 (3.34)	-17.20 (10.07)*	-0.59 (0.63)	-3.27 (2.88)	-12.11 (8.71)
Asian, Middle Income	-0.20 (0.10)**	-1.99 (1.23)	-13.99 (8.58)	-0.06 (0.16)	-0.20 (0.44)	-3.06 (1.65)*
Asian, High Income	-1.14 (0.91)	-3.57 (2.13)*	-14.13 (7.87)*	0.14 (0.22)	-0.76 (0.55)	-4.20 (1.82)**
Hispanic, Low Income	-0.65 (0.27)**	-5.26 (1.89)***	-24.19 (8.05)***	-0.71 (0.37)*	-1.99 (2.27)	-10.57 (9.24)
Hispanic, Middle Income	-0.33 (0.12)**	-1.63 (0.49)***	-12.56 (3.71)***	-0.02 (0.24)	-0.68 (0.57)	-5.77 (3.57)
Hispanic, High Income	-0.77 (0.49)	-2.04 (1.11)*	-10.35 (4.07)**	-0.04 (0.13)	-0.48 (0.35)	-5.52 (2.40)**
All Whites	-0.63 (0.32)*	-2.36 (1.16)**	-9.42 (3.81)**	-0.18 (0.13)	-0.84 (0.42)**	-4.12 (1.39)***
All Blacks	-0.32 (0.10)***	-3.05 (1.56)*	-17.03 (7.98)**	-0.23 (0.12)**	-2.19 (1.29)*	-12.93 (6.54)**
All Asians	-0.77 (0.41)*	-3.29 (1.52)**	-14.82 (6.63)**	-0.10 (0.20)	-1.17 (0.78)	-5.71 (2.45)**
All Hispanics	-0.55 (0.16)***	-3.46 (0.99)***	-17.88 (4.69)***	-0.36 (0.20)*	-1.29 (1.14)	-8.07 (5.19)
All Low Income	-0.63 (0.26)**	-4.09 (1.49)***	-19.09 (6.09)***	-0.65 (0.33)*	-1.82 (1.66)	-9.63 (6.69)
All Middle Income	-0.26 (0.10)**	-1.61 (0.61)**	-11.35 (4.37)**	-0.02 (0.18)	-0.58 (0.44)	-4.81 (2.28)**
All High Income	-0.88 (0.58)	-3.22 (1.74)*	-11.85 (5.19)**	-0.06 (0.16)	-1.07 (0.64)	-5.38 (2.20)**
Total	-0.59 (0.21)***	-2.97 (0.87)***	-14.09 (3.74)***	-0.25 (0.14)*	-1.16 (0.63)*	-6.61 (2.82)**

Notes: Change in emissions from Period 2 to Period 3. Electric facilities are included. We denote significance with *** at the 1-percent level, ** at the 5-percent level, and * at the 10-percent level.

Table 4: Average Treatment Effect using Nearest Neighbors Matching.

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

	Levels		Logs	
Method	Coef.	Std.Err.	Coef.	Std.Err.
OLS	-38.14	(15.06)**	-0.37	(0.11)***
Nearest Neighbor Matching	-17.79	(7.63)**	-0.25	(0.09)***
Propensity Score Matching	-24.81	(13.86)*	-0.27	(0.12)**

Panel B: Change in NOx Emissions between Periods 2 and 3 for Complying Facilities.

	Levels		Logs	
Method	Coef.	Std.Err.	Coef.	Std.Err.
OLS	-6.82	(6.66)	-0.23	(0.05)***
Nearest Neighbor Matching	-9.42	(3.85)**	-0.26	(0.06)***
Propensity Score Matching	-14.78	(2.22)***	-0.26	(0.03)***

Period 1: average of positive emissions in 1990 and 1993
Period 2: average of positive emissions in 1997 and 1998
Period 3: average of positive emissions in 2001 and 2002
Period 4: average of positive emissions in 2004 and 2005

Notes: The OLS estimates control for average NOx emissions during Period 1 (1990-1993) and four-digit SIC code indicator variables, with standard errors clustered by air basin. The nearest neighbor matching models match on historic emissions and four-digit SIC codes and use the preferred number of matches (see Table A3). The propensity score matching models match on historic emissions and two-digit SIC codes. For all models, the control group is restricted to facilities located in counties that, during 1990 and 1993, were not in attainment with the 1-hour ozone NAAQS standards. In Panel A, there are 212 treatment observations of the sample of 1523. For Panel B, there are 255 treated, which excludes the 13 facilities not complying with RECLAIM, of 1932. For the log specifications, emissions differences are defined as ln(EmitX+1)-ln(EmitY+1) and all matching is on ln(Emit1+1). 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.

Description	Levels		Logs	
	Coef.	Std.Err.	Coef.	Std.Err.
Change in NOx Emissions between Periods 1 and 4				
Nearest Neighbor Matching	-0.91	(2.13)	-0.07	(0.06)
Propensity Score Matching	-3.03	(1.68)	-0.11	(0.04) **
Change in NOx Emissions between Periods 2 and 3 for Complying Facilities				
Nearest Neighbor Matching	-0.76	(1.98)	0.08	(0.06)
Propensity Score Matching	-0.31	(1.51)	0.08	(0.03) **

Notes: Treatment is L.A. facilities not in RECLAIM, but in the same four-digit SIC industries as in RECLAIM.

Table 6: Robustness to Control Group.

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

	Levels		Logs	
Description	Coef.	Std.Err.	Coef.	Std.Err.
Nearest Neighbor Matching				
Exclude L.A. facilities	-20.86	(7.96) ***	-0.34	(0.09) ***
Northern CA only	-15.62	(6.04) **	-0.29	(0.11) **
Southern CA only	-23.67	(7.58) ***	-0.23	(0.11) **
Severe Non-Attainment only	-43.48	(9.74) ***	-0.39	(0.13) ***
Propensity Score Matching				
Exclude L.A. facilities	-29.63	(17.59)	-0.28	(0.12) **
Northern CA only	-23.79	(9.12) **	-0.17	(0.08) *
Southern CA only	-23.86	(27.68)	-0.30	(0.22)
Severe Non-Attainment only	-44.91	(25.67)	-0.33	(0.17) *

Panel B: Change in NOx Emissions between Periods 2 and 3 for Complying Facilities.

	Levels		Logs	
Description	Coef.	Std.Err.	Coef.	Std.Err.
Nearest Neighbor Matching				
Exclude L.A. facilities	-9.04	(4.40) **	-0.21	(0.07) ***
Northern CA only	-4.89	(7.51)	-0.19	(0.07) ***
Southern CA only	-8.71	(3.90) **	-0.28	(0.07) ***
Severe Non-Attainment only	-10.49	(5.31) **	-0.23	(0.09) ***
Propensity Score Matching				
Exclude L.A. facilities	-17.25	(4.95) ***	-0.24	(0.09) **
Northern CA only	-22.88	(6.78) **	-0.14	(0.11)
Southern CA only	-13.43	(1.71) ***	-0.20	(0.06)
Severe Non-Attainment only	-19.77	(7.77) **	-0.35	(0.04) ***

Table 7: Environmental Justice Results from Emissions Trading

LHS is Change in NOx Emissions from Period 2 to Period 3 for Compliant Firms

Panel A: Nearest Neighbor Matching

	1	2	3	4	5	6	7
Treatment	-6.70***	-7.51***	-7.22**	-7.66***	-7.81***	-7.82**	-8.46***
	(1.43)	(2.21)	(2.45)	(1.90)	(2.00)	(2.91)	(2.26)
Treat * Period 1 NOx	-0.06***			-0.07***	-0.07***		-0.07***
	(0.02)			(0.01)	(0.01)		(0.01)
Treat * Income		-0.23		-0.15		-0.37	-0.47
		(0.20)		(0.13)		(0.31)	(0.32)
Treat * %Minority			0.03		-0.06	-0.11	-0.25
			(0.09)		(0.07)	(0.15)	(0.19)
Period 1 NOx	-0.35***	-0.34***	-0.34***	-0.37***	-0.37***	-0.34***	-0.37***
	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)	(0.09)
Income		0.22		0.16		0.29	0.31
		(0.24)		(0.19)		(0.37)	(0.33)
%Minority			-0.05		0.02	0.05	0.14
			(0.10)		(0.09)	(0.17)	(0.17)
R^2	0.64	0.63	0.62	0.64	0.64	0.63	0.64

Panel B: Propensity Score Matching

	1	2	3	4	5	6	7
Treatment	-12.02***	-19.74***	-18.67***	-16.52***	-16.64***	-18.55***	-16.91***
	(2.84)	(3.06)	(3.20)	(3.42)	(3.35)	(2.69)	(2.93)
Treat * Period 1 NOx	-0.15***			-0.19***	-0.19***		-0.19***
	(0.04)			(0.05)	(0.05)		(0.05)
Treat * Income		-0.18		-0.14		-0.02	-0.37
		(0.31)		(0.23)		(0.50)	(0.35)
Treat * %Minority			0.17		-0.05	0.16	-0.20
			(0.10)		(0.10)	(0.22)	(0.16)
Period 1 NOx	-0.08*	-0.15***	-0.15***	-0.03	-0.03	-0.15***	-0.03
	(0.04)	(0.04)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)
Income		0.09		0.07		-0.04	0.13
		(0.30)		(0.21)		(0.44)	(0.31)
%Minority			-0.12		0.02	-0.13	0.06
			(0.11)		(0.11)	(0.20)	(0.17)
R^2	0.35	0.35	0.35	0.46	0.46	0.35	0.46

Notes: For the nearest neighbor matching, there are 974 observations with demographic data. Group fixed effects are not shown. Weight control observations by control group size. For the propensity score matching, there are 1697 observations with demographic data and we do not show the two-digit SIC fixed effects. Significance is noted at the 10% (*), 5% (**) and 1% (***) levels. %Minority is percent black or Hispanic.

Table 8: Environmental Justice Results from Overall RECLAIM Policy

LHS is Change in NOx Emissions from Period 1 to Period 4 for All Firms

Panel A: Nearest Neighbor Matching

	1	2	3	4	5	6	7
Treatment	-20.64 **	-15.35 **	-12.04 **	-15.75 **	-13.02 **	-9.68 ***	-10.84 **
	(7.81)	(6.14)	(4.00)	(5.03)	(5.25)	(2.56)	(4.37)
Treat * Period 1 NOx	-0.19			-0.11 **	-0.10 **		-0.10 **
	(0.11)			(0.04)	(0.04)		(0.04)
Treat * Income		-0.45		-0.17		1.21	0.98
		(0.57)		(0.59)		(1.01)	(0.91)
Treat * %Minority			0.91 *		0.61	1.35 *	0.96
			(0.46)		(0.37)	(0.69)	(0.54)
Period 1 NOx	-0.48 ***	-0.34 ***	-0.34 ***	-0.36 ***	-0.36 ***	-0.34 ***	-0.36 ***
	(0.11)	(0.06)	(0.06)	(0.06)	(0.05)	(0.06)	(0.06)
Income		-0.21		-0.23		-0.97	-0.88
		(0.54)		(0.50)		(0.93)	(0.84)
%Minority			-0.39		-0.31	-0.73	-0.62
			(0.29)		(0.29)	(0.52)	(0.48)
R^2	0.90	0.93	0.93	0.93	0.93	0.93	0.93

Panel B: Propensity Score Matching

	1	2	3	4	5	6	7
Treatment	-20.11 **	-22.16 **	-20.93 **	-18.04 **	-18.01 **	-18.59 ***	-16.24 ***
	(8.23)	(9.55)	(8.91)	(6.25)	(6.51)	(6.08)	(4.99)
Treat * Period 1 NOx	-0.12			-0.12	-0.11		-0.11
	(0.10)			(0.09)	(0.09)		(0.09)
Treat * Income		-0.28		-0.17		0.65	0.37
		(0.33)		(0.37)		(0.80)	(0.63)
Treat * %Minority			0.59 **		0.36 **	0.83	0.49
			(0.25)		(0.13)	(0.51)	(0.33)
Period 1 NOx	-0.61 ***	-0.69 ***	-0.68 ***	-0.62 ***	-0.62 ***	-0.68 ***	-0.62 ***
	(0.11)	(0.05)	(0.05)	(0.09)	(0.09)	(0.05)	(0.09)
Income		-0.34		-0.32		-0.61	-0.47
		(0.50)		(0.44)		(0.82)	(0.70)
%Minority			-0.08		0.00	-0.30	-0.17
			(0.14)		(0.11)	(0.41)	(0.33)
R^2	0.87	0.89	0.89	0.89	0.89	0.89	0.89

Notes: For the nearest neighbor matching, there are 829 observations with demographic data. Group fixed effects are not shown. Weight control observations by control group size. For the propensity score matching, there are 1396 observations with demographic data and we do not show the two-digit SIC fixed effects. Significance is noted at the 10% (*), 5% (**) and 1% (***) levels. %Minority is percent black or Hispanic.

Appendices: Not for Publication

Appendix A: Emissions trends under the CAAAs

In 1990, Congress turned its attention to the widespread failure of US cities to attain health-based national ambient air quality standards (NAAQS). Title I of the 1990 CAAAs sought to address urban ozone problems in particular.[40] The 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. Figure A1 graphically 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. Because SCAQMD was much further from attainment as compared to other air basins, the district was given more time to comply.

The stronger unconfoundedness assumption described in Section 3.2 requires that the control treatment resembles the regulations that RECLAIM facilities would have been subjected to absent RECLAIM. 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 A1 illustrates the compliance requirements

---

[40] Of all the criteria air pollutants regulated under the auspices of the Federal Clean Air Act, ground level ozone has proven to be the most recalcitrant. Ozone is the primary constituent of smog. It is a highly reactive compound that can damage the linings of the respiratory tract, reduce lung function, and aggravate pre-existing respiratory conditions. In 1990, an estimated sixty percent of Americans lived in metropolitan areas that were failing to meet Federal health-based ozone standards.

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. The broken lines represent the more recently required ozone concentration reduction trajectories that pertain to the federal 8-hour ozone standard.[41] 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. Figure A1 illustrates that mandated ozone concentration reduction trajectories were similar across California's non-attainment counties. This is consistent with, but certainly not proof of, the hypothesis that regulations of NOx emissions from industrial sources in SCAQMD and other non-attainment areas would have followed parallel paths over the time period we study absent RECLAIM.

Appendix B: 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

---

[41]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.

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, 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).[42]

---

[42]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

Unresolved disagreements about what constitutes an appropriate measure of counterfactual emissions has 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 C: Robustness to Estimates

This appendix tests the robustness of the main results. First we examine the overall match quality (see Figure A2). In Table A1, we explore the robustness of our nearest neighbor results in Table 4 to the bias adjustments we make using no correction, a linear correction, and a quadratic correction. Patterns of coefficient significance are unaffected by the bias adjustments.

For the observations used in Panel A of Table 8, Table A2 summarizes the demographic data and other covariates we use to test the robustness of the environmental justice results. These include: a variable measuring whether toxics were measured on site from the California ARB's Air Toxics "Hot Spots" Program (http://www.arb.ca.gov/AB2588/overview.htm); annual county employment, payroll and establishments variables; an indicator of whether the RECLAIM facility was located in the coastal zone with restricted permit access; and

---

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).

indicators of two-digit SIC codes for major industries. When including these in the environmental justice analysis, the only coefficients for the treatment effects that are significant are the average effect, historic emissions, and petroleum refining (which is positive).

Table A3 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.

Appendix Figures and Tables
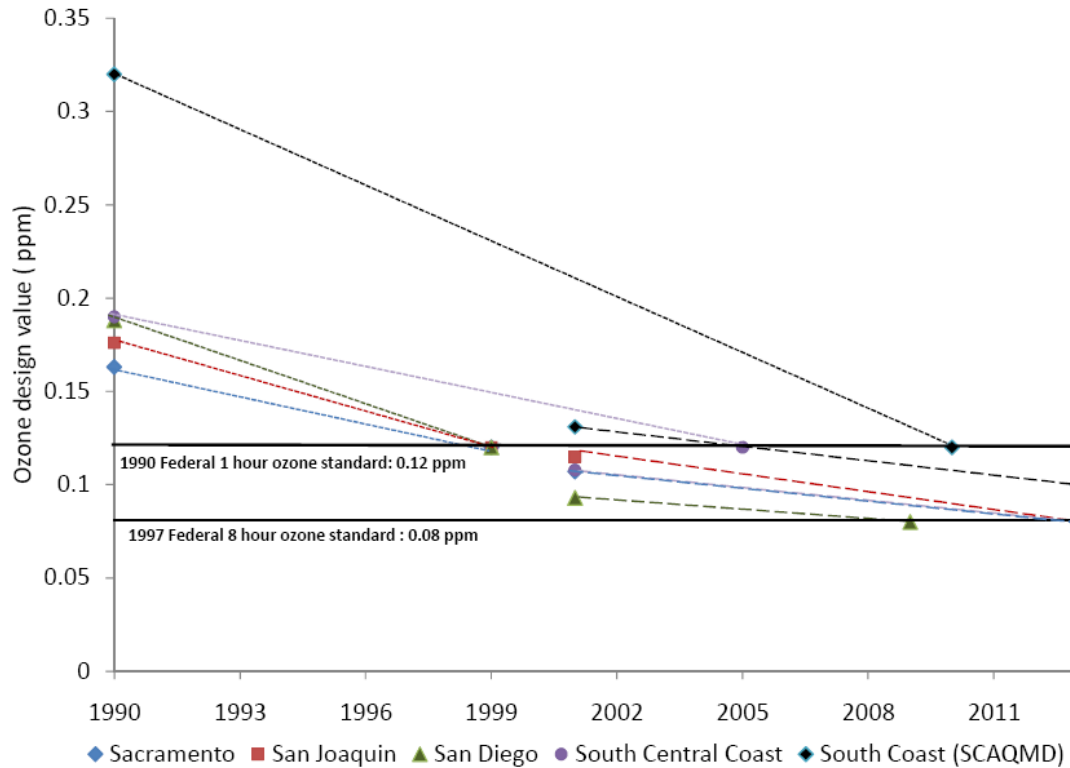
Figure A1: Required Ozone Concentration Reductions for Five Californian Air Basins.

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.
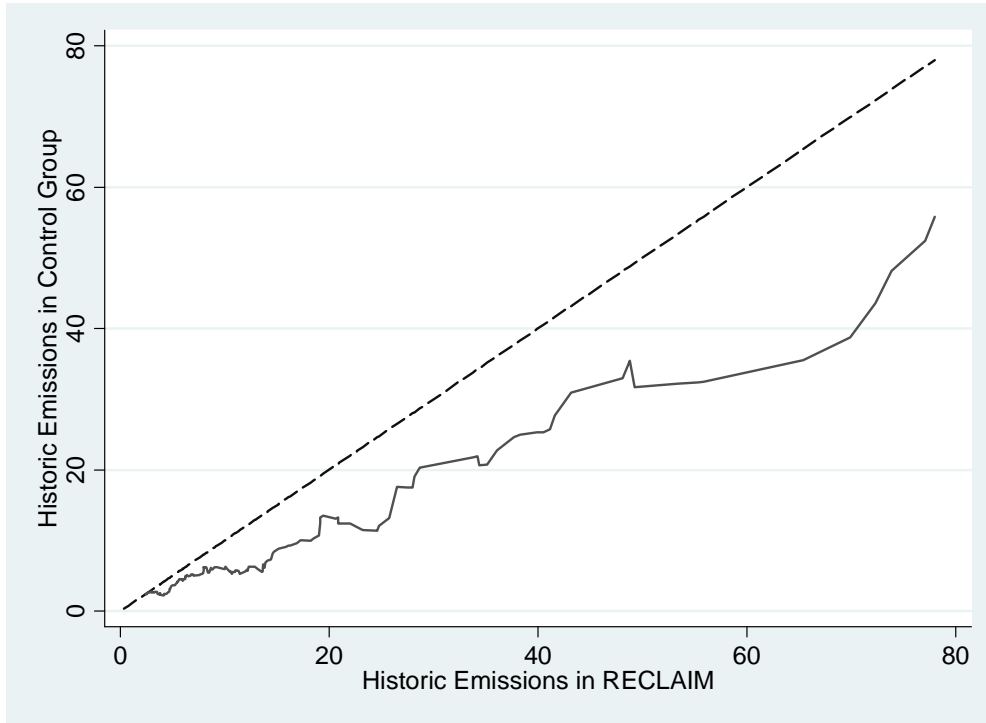
Figure A2: Match Quality for Overall Effect on Historic Emissions.
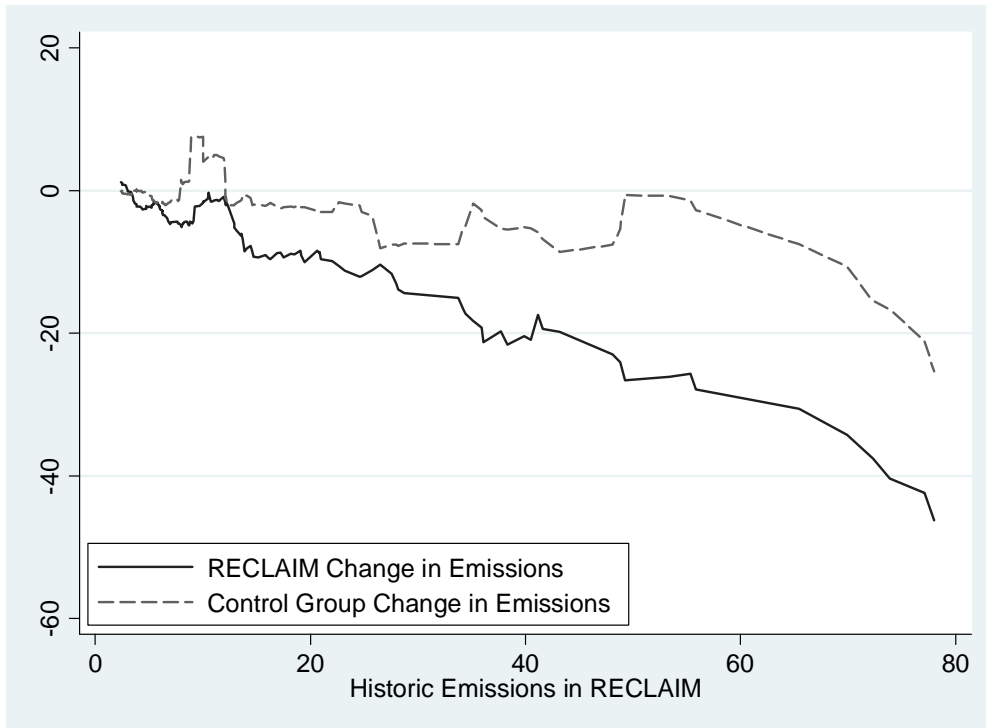
Figure A3: k-Nearest Neighbor Regression of Changes in Emissions from Period 1 to Period 4 in the RECLAIM and Control Groups on Period 1 Emissions. The sample is from the main results shown in Table 4.

Table A1: Robustness to Bias Adjustment and Matching Method.

Panel A: No Bias Adjustment

Description of dependent variable	Levels		Logs			
	Coef.	Std.Err.	Coef.	Std.Err.	n	treat
Overall (Emit4-Emit1)	-25.02	(7.63) ***	-0.32	(0.09) ***	1,745	215
Trading (Emit3-Emit2)	-9.60	(3.85) **	-0.37	(0.08) ***	2,217	255

Panel B: Linear Bias Adjustment

Description of dependent variable	Levels		Logs			
	Coef.	Std.Err.	Coef.	Std.Err.	n	treat
Overall (Emit4-Emit1)	-20.59	(7.63) ***	-0.27	(0.09) ***	1,745	215
Trading (Emit3-Emit2)	-8.29	(3.85) **	-0.34	(0.08) ***	2,217	255

Panel C: Quadratic Bias Adjustment (main specification)

Description of dependent variable	Levels		Logs			
	Coef.	Std.Err.	Coef.	Std.Err.	n	treat
Overall (Emit4-Emit1)	-17.79	(7.63) **	-0.25	(0.09) ***	1,745	215
Trading (Emit3-Emit2)	-9.42	(3.85) **	-0.31	(0.08) ***	2,217	255

Notes: EmitX is emissions in period X. All models match on the same variables as in Table 4. For the log specification, emissions differences are defined as $\ln(EmitX+1)-\ln(EmitY+1)$ and all matching is on $\ln(Emit1+1)$. For each LHS variable, the preferred *m* from Table 4 is used. The kernel matching panel uses a bandwidth of 0.05. We denote significance with *** at the 1-percent level, ** at the 5-percent level, and * at the 10-percent level.

Table A2: Demographic Summary Statistics

Variables for RECLAIM facilities	n	mean	std dev	min	max
median zip code income (in $1000s) in 1999	208	45.2	16.3	8.6	98.1
percent of zip code that is black or Hispanic (in 2000)	208	55.2	27.8	4.0	98.8
indicator of whether toxics were measured on site	208	29%	45%	0%	100%
% change in total county employment from period 1 to 4	208	23%	18%	11%	69%
% change in total county payroll from period 1 to 4	208	76%	30%	56%	153%
% change in total county establishments from period 1 to 4	208	17%	7%	13%	35%
Indicator of coastal permits	208	70%			
Petroleum Refining	208	7%			
Stone, Clay, and Glass Products	208	9%			
Primary Metal Industries	208	9%			
Electric and Gas Services	208	16%			

Variables for control facilities	n	mean	std dev	min	max
median zip code income (in $1000s) in 1999	621	45.9	15.4	17.9	120.1
percent of zip code that is black or Hispanic (in 2000)	621	46.7	26.5	3.1	100.0
indicator of whether toxics were measured on site	621	19%	39%	0%	100%
% change in total county employment from period 1 to 4	621	27%	17%	6%	123%
% change in total county payroll from period 1 to 4	621	94%	33%	56%	277%
% change in total county establishments from period 1 to 4	621	14%	8%	-2%	60%
Indicator of coastal permits	621	0%			
Petroleum Refining	621	9%			
Stone, Clay, and Glass Products	621	9%			
Primary Metal Industries	621	9%			
Electric and Gas Services	621	16%			

Table A3: Robustness of Average Treatment Effect to Number of Nearest Neighbors.

Dependent Variable: Change in NOx Emissions from Period X to Period Y

Dependent variable	m(1)	m(2)	m(3)	m(4)	m(5)
Overall (Emit4-Emit1)	-27.06***	-21.57***	-17.79**	-15.17*	-15.09*
	(8.74)	(8.07)	(7.63)	(7.81)	(8.10)
Sum MSE	44.30	48.10	**42.30**	47.90	46.00
Average percent change	-26%	-21%	-17%	-15%	-15%
Trading (Emit3-Emit2)	-7.83	-9.64**	-9.42**	-9.38**	-9.89**
	(7.08)	(4.13)	(3.85)	(4.64)	(4.06)
Sum MSE	17.44	15.13	**14.48**	15.31	16.02
Average percent change	-16%	-20%	-20%	-19%	-20%

Notes: All matching models match on average NOx emissions during Period 1 (1990-1993) and four-digit SIC codes. In addition, the control group is restricted to facilities located in counties that, during 1990 and 1993, were not in attainment with the 1-hour ozone NAAQS standards. In the first row, the dependent variable is the change in NOx emissions from period 1 to period 4 (215 treatment observations of the sample of 2421). The second row is the change from period 2 to 3 (255 treated, excluding the 13 facilities not complying with RECLAIM, of 2217). m(#) specifies the number of matches to be made per observation. The sum of mean squared errors (in millions) is shown for each regression (see text for discussion of calculation). The preferred model is in bold. We denote significance with *** at the 1-percent level, ** at the 5-percent level, and * at the 10-percent level.