DO ENVIRONMENTAL MARKETS CAUSE ENVIRONMENTAL INJUSTICE?
EVIDENCE FROM CALIFORNIA'S CARBON MARKET

Danae Hernandez-Cortes
Kyle C. Meng

Do Environmental Markets Cause Environmental Injustice? Evidence from California's Carbon
Market
Danae Hernandez-Cortes and Kyle C. Meng
NBER Working Paper No. 27205
May 2020, Revised April 2021
JEL No. H4,I14,Q5,Q51,Q52,Q53,Q54

ABSTRACT

Market-based environmental policies are widely adopted for their allocative efficiency. However, there is growing concern that market-induced spatial reallocation of pollution could widen existing pollution concentration gaps between disadvantaged and other communities. We estimate how this "environmental justice" (EJ) gap changed following the 2013 introduction of California's greenhouse gas (GHG) market, the world's second largest and most subjected to EJ critiques. We find the policy has reduced GHG and criteria air pollution emissions. Applying an atmospheric dispersal model to determine resulting pollution concentration changes, we detect the EJ gap, which was widening before 2013, has since fallen across criteria pollutants.

Danae Hernandez-Cortes
Department of Economics
University of California, Santa Barbara
Santa Barbara, CA 93106
hernandezcortes@umail.ucsb.edu

Kyle C. Meng
Bren School of Environmental
Science and Management
Department of Economics
University of California, Santa Barbara
4416 Bren Hall
Santa Barbara, CA 93106
and NBER
kmeng@bren.ucsb.edu

1 Introduction

Over the last three decades, policy makers have increasingly relied on market-based environmental policies - such as pollution trading and taxes - to address environmental problems. Expanded use of market-based policies followed each major amendment to the U.S. Clean Air Act since the 1970s (Schmalensee and Stavins, 2019). Widespread adoption has occurred in other environmental domains: today, market-based policies cover 30% of global fisheries (Costello et al., 2016), account for over $36 billion in global ecosystem service payments (Salzman et al., 2018), and govern 20% of global greenhouse gas (GHG) emissions (World Bank Group, 2019).

The central appeal of market-based environmental policies is allocative efficiency. In theory, such policies reduce the total abatement cost of meeting an environmental objective by inducing less abatement from polluters with higher abatement costs (Crocker, 1966; Dales, 1968; Montgomery, 1972). This contrasts with traditional command-and-control regulations, which typically require heterogeneous polluters to adopt uniform abatement actions.

At the same time, the particular reallocation of emissions induced by market-based policies also spatially alters who is harmed by pollution. This is of particular concern as a growing "environmental justice" (EJ) literature has documented that communities with lower income, higher minority share, and/or otherwise disadvantaged, systematically experience higher pollution concentrations than other communities, a statistic we refer to as the environmental justice gap (or EJ gap).[1] Could the adoption of environmental markets be compounding existing EJ gaps?

Whether a market-based environmental policy widens or narrows the EJ gap depends on the joint spatial distribution of polluting facilities, their abatement costs, and disadvantaged communities. Market-based policies induce relatively less abatement from facilities with steeper marginal abatement cost curves. If these facilities are upwind of disadvantaged communities, such policies will widen an existing EJ gap. Conversely, if these facilities are upwind of non-disadvantaged communities, a market-based policy will narrow the EJ gap (Burtraw et al., 2005).[2] Unfortunately, facility-level marginal abatement cost curves are usually unobserved, making it hard to anticipate the direction of EJ gap effects ex-ante. This difficulty underscores the need for ex-post empirical approaches, for which prior studies have largely found inconclusive EJ gap effects (Fowlie, Holland and Mansur, 2012; Grainger

---

[1]EJ gaps across many settings have been shown through case (Bullard, 2000; Bowen, 2002; Ringquist, 2005; Mohai, Pellow and Roberts, 2009; Banzhaf, Ma and Timmins, 2019) and population-level (Tessum et al., 2019; Colmer et al., 2020; Currie, Voorheis and Walker, 2020) studies.

[2]Additionally, for a policy regulating global pollutants like greenhouse gases, the EJ gap effect depends on the extent in which GHG and local pollutants are co-produced.

and Ruangmas, 2018; Shapiro and Walker, 2021).

This paper estimates the EJ gap consequences of California's greenhouse gas (GHG) cap-and-trade (C&T) program, which since 2013 has created the world's second largest carbon market. This program has also been a focal point of EJ concerns, as local air pollution emissions are typically co-produced with GHG emissions.[3] The possibility that the program could widen California's existing EJ gaps in local air pollution has, among other critiques, led to political opposition that temporary paused the program's initial development in 2011 and nearly halted renewal efforts in 2017. However, to date, there has been limited causal evidence on whether the program has indeed widened EJ gaps.

We make two contributions, one empirical and another methodological, in order to establish the EJ gap consequences of California's C&T program. First, we find that the C&T program has lowered GHG and criteria air pollution (i.e., $PM_{2.5}$, PM_{10}, NO_x, and SO_x) emissions for sample facilities. Specifically, we exploit the program's facility-level eligibility rule based on historical emissions and its timing to estimate a break in differential emission trends between regulated and unregulated facilities after 2013. This research design is possible because we observe facility GHG and criteria air pollution emissions for both regulated and unregulated facilities, and for periods before and after the program's introduction, data availability that is not common across cap-and-trade programs. For example, facility-level pre-program emissions are not directly observed for the European Union Emissions Trading System (EU-ETS), the world largest carbon market (Petrick and Wagner, 2014; Martin, Muûls and Wagner, 2016; Colmer et al., 2020). Even in settings where emissions data is available, emissions-based eligibility thresholds can sometimes be too low for there to be sufficient control units within the same jurisdiction, as in the case of Southern California's RECLAIM NO_x C&T program (Fowlie, Holland and Mansur, 2012). We compare regulated and unregulated units within the same jurisdiction. Our identifying assumption requires that any existing differential emission pre-trends between regulated and unregulated facilities would have continued after 2013 if not for the C&T program.

We estimate that C&T reduced emissions annually at a rate of 3-9% across GHG and criteria air pollutants during 2012-2017. To isolate the C&T effect from that of other concurrent climate programs in California, such as renewable portfolio and low carbon fuel standards, we restrict attention to facilities that were only directly regulated by C&T. Emissions abatement induced by these complementary climate programs have in general made it difficult to discern whether C&T contributed to the recent 27.1 million ton CO_2e decline in

---

[3]Similar EJ concerns have arisen elsewhere. Recent efforts to introduce state-level U.S. climate policies and renew the European Union Emissions Trading System were opposed on EJ grounds (Leber, 2016; Herron, 2019; Transnational Institute, 2013).

total California GHG emissions between 2012-2017 (Burtraw et al., 2018; Borenstein et al., 2019). Our estimates imply that GHG emissions across sample facilities declined by 3.2 million tons of CO_2e during this period. The exclusion of C&T-regulated facilities subject to complementary program from our analysis suggests that this figure is a weak lower bound on the effect of C&T on total California GHG emissions.

We demonstrate that C&T emissions effects are robust to various model specification and sample restriction choices; to concerns about spillover effects between regulated and unregulated facilities; and to heterogeneity in emission effects as a function of a facility's average emissions. In a placebo test that systematically imposes trend breaks across sample years, we detect the largest trend break in 2013, the year when the program was actually introduced.

Our second contribution is to develop an empirical approach for determining how policy-driven changes in pollution *emissions* alter the spatial distribution of pollution *concentrations*. The canonical economics framework for evaluating environmental policies requires knowing the link between pollution "source" and "receptors" (Baumol and Oates, 1988). In practice, however, this mapping is rarely characterized and is instead assumed to follow simple spatial patterns such as assigning pollution concentrations to areas within the same geographic unit of a facility or within a distance circle centered at a facility (Banzhaf, Ma and Timmins, 2019). In reality, the spatial and temporal patterns of pollution dispersal are far more complex and depend on topography and time-varying atmospheric conditions. Failure to accurately account for actual dispersal patterns can lead to bias estimates even in otherwise valid quasi-experimental settings (Deschenes and Meng, 2018).

To address this challenge, our estimation framework explicitly embeds an atmospheric dispersal model, a computationally-intensive procedure that involves running over two million pollution trajectories. Specifically, our approach combines estimates of C&T emissions effects (and its uncertainty) at the facility level with an analysis of resulting EJ gap changes at the location level, as determined by the atmospheric dispersal model. In doing so, we build on prior studies using dispersal models which typically only conduct analysis at the location-level (e.g., Sullivan (2017)) or facility-level (e.g., Grainger and Ruangmas (2018)). Location-only studies insert observed (not policy-driven) emissions into a dispersal model and thus do not consider changes in pollution concentration arising from specific policies. Facility-only studies typically examine whether a policy's effect on emissions varies with demographic characteristics of downwind locations from a facility. We formally demonstrate that estimates from such facility-level regressions do not in general equal the EJ gap effect and need not even be of the same sign. We further show that for the EJ gap effect to be recovered from facility-level estimates, one must assume a highly simplistic spatial pattern

4

of pollution dispersal, an assumption which we can reject for our setting.

Employing a definition of a "disadvantaged" zip code that serves as a basis for California's EJ policies, we detect three EJ gap findings. First, consistent with EJ concerns in the lead up to the C&T program's introduction, we find not only were there baseline EJ gaps across criteria air pollutants in 2008, but that gaps were widening in the 2008-2012 period before the program. Second, the C&T program has narrowed EJ gaps since 2013. Third, while EJ gaps have narrowed, they have not been eliminated: by 2017, the C&T program returned EJ gaps roughly to 2008 levels. These EJ gap effects are robust across a variety of checks. In particular, we find that allowing for heterogeneous emissions effects as a function of a facility's average emissions leads to slightly larger declines in EJ gaps. We further demonstrate similar EJ gap effects when employing an alternative atmospheric dispersal model that generates secondary $PM_{2.5}$ concentrations. An analysis of spatial heterogeneity reveals that EJ gaps narrowed most for disadvantaged zip codes in California's Central Valley, while a few disadvantaged zip codes in Los Angeles County experienced widening gaps.

We demonstrate the importance of modeling pollution dispersal for our results. Our EJ gap effects become unstable if instead of modeling pollution dispersal, we were to employ more conventional approaches for assigning pollution emissions to concentrations. We posit that our empirical approach may have broader applicability. In particular, there is a common need across many environmental policy settings to track how policy-driven changes in pollution emissions alter the spatial distribution of pollution concentration (Greenstone and Gayer, 2009; Graff Zivin and Neidell, 2013; Deschenes and Meng, 2018).

The paper is structured as follows: Section 2 considers a conceptual framework for how a C&T program could widen or narrow an existing EJ gap and offers background on California's GHG C&T program. Section 3 summarizes our data. Section 4 details our empirical approach. Section 5 presents our results. Section 6 provides a discussion.

2 Background

We begin by discussing how the introduction of a cap-and-trade (C&T) program can either widen or narrow existing pollution concentration gaps between disadvantaged and other communities. We then review California's greenhouse gas (GHG) cap-and-trade program.

2.1 Cap-and-trade and the environmental justice gap

In a textbook C&T program, the regulator establishes a limit (or cap) on total emissions within a jurisdiction by issuing a fixed supply of emission permits. Regulated facilities are

then either given, or must purchase, permits to cover their emissions. Permit trading allows the marginal abatement cost (MAC) of emissions to be equalized amongst regulated facilities to the permit price.[4]

Two key consequences of C&T are often emphasized. First, by placing a price on pollution, a C&T program induces polluting facilities to internalize (some of) the social costs of their emissions.[5] Second, by equalizing MACs across facilities, a C&T program allocates emissions by inducing relatively less abatement from facilities with steeper MAC curves and more abatement from facilities with flatter MAC curves. In theory, the resulting allocation of abatement achieves the aggregate emissions cap at the lowest total abatement cost across regulated facilities (Montgomery, 1972).

What is less clear is how the allocative efficiency achieved by C&T alters the spatial distribution of pollution concentration. In particular, there is growing concern that the same market forces resulting in allocative efficiency may also be altering the difference in pollution concentrations experienced between disadvantaged and other communities. This difference, which we call the "environmental justice gap" (or EJ gap) has been shown to be positive in the many settings (Bullard, 2000; Bowen, 2002; Ringquist, 2005; Mohai, Pellow and Roberts, 2009; Banzhaf, Ma and Timmins, 2019; Tessum et al., 2019; Colmer et al., 2020; Currie, Voorheis and Walker, 2020).

The introduction of C&T can either widen or narrow an existing EJ gap. Figure 1 illustrates this ambiguous effect for a stylized two-facility setting with emissions (e) on the horizontal axis and permit prices (τ) on the vertical axis. The first facility is upwind of a disadvantaged community (DAC) with a marginal abatement curve labeled "DAC" (in orange). The second facility is upwind of a non-disadvantaged community and has a marginal abatement curve labeled "non-DAC" (in gray).[6] To establish an existing positive EJ gap prior to the introduction of C&T, we allow the DAC facility to have larger emissions in the absence of C&T, or when $\tau = 0$. When C&T is introduced, each facility's MAC is equated to the equilibrium permit price $\tau = \tau^*$. What happens to the EJ gap?

In the left panel of Figure 1, the DAC facility has a steeper MAC curve than the non-DAC facility, causing the DAC facility to abate less than the non-DAC facility under C&T. In this case, C&T widens the EJ gap. The right panel of Figure 1 shows an alternative case whereby the DAC facility has a flatter MAC curve than the non-DAC facility. Following C&T, the DAC facility abates more than the non-DAC facility, narrowing the EJ gap.

Thus, in settings with an existing positive EJ gap, whether C&T widens or narrows the

---

[4]The modern C&T framework was initially developed by Crocker (1966) and Dale (1968).

[5]Whether social costs are fully internalized depends on if the cap is set at the socially optimal level.

[6]The horizontal axes in Figure 1 indicates emissions rather than abatement in order to illustrate emissions levels prior to C&T when $\tau = 0$.

EJ gap depends on whether facilities upwind of DAC communities have relatively steeper or flatter MAC curves. Furthermore, for a cap-and-trade system regulating a global pollutant such as greenhouse gases, the EJ gap effect depends on the extent in which GHG and local air pollutants are co-produced.

Figure 1: EJ gap under cap-and-trade

DAC-upwind facility with steeper MAC   DAC-upwind facility with flatter MAC
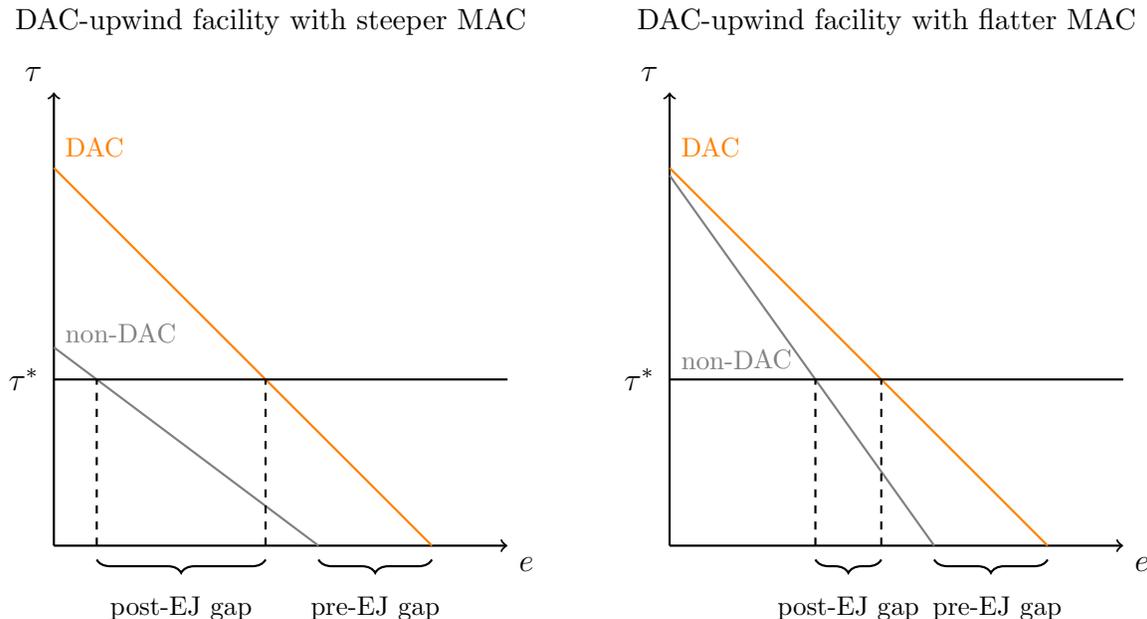
NOTES: Panels illustrate how the introduction of a C&T program can widen or narrow an existing EJ gap in a two facility setting. Horizontal axes indicate emissions. Vertical axes indicate marginal abatement costs, and equivalently the permit price under C&T. The marginal abatement cost curve for facility upwind of a disadvantaged community (labeled DAC) is shown in orange. The marginal abatement cost curve for facility upwind of a non-disadvantaged community (labeled non-DAC) is shown in gray. τ^* indicates the permit price under C&T. In the left panel, the DAC-upwind facility has a relatively steeper MAC curve. In the right panel, the DAC-upwind facility has a relatively flatter MAC curve.

Unfortunately, facility-level MAC curves are rarely observed, which limits the ability to anticipate EJ gap effects of proposed C&T programs. Ex-post studies also face several empirical challenges. First, isolating the effect of reallocation from cap-and-trade requires restricting attention to facilities that are only regulated by cap-and-trade and not additionally by complementary climate programs.[7] Second, to remove the influence of macroeconomic conditions, one needs to estimate the effect of cap-and-trade for regulated facilities relative to unregulated facilities. Third, estimated C&T-driven facility emissions must be mapped onto location-level pollution concentrations in order to examine resulting EJ gap changes. Section 4 details how we overcome these challenges.

---

[7]Returning to Figure 1, suppose a complementary climate program binds for one facility such that its emissions are unchanged following the introduction of cap-and-trade. Any subsequent change in the EJ gap following cap-and-trade now depends on the complementary program, and in particular whether it binds for the DAC or non-DAC facility and at what emissions level.

2.2 California's GHG cap-and-trade program

California's has one of the world's most sophisticated and ambitious climate policies. In 2006, California passed Assembly Bill 32 (AB 32), requiring total GHG emissions across the state to reach 1990 emissions level by 2020. AB 32 remains the only economy-wide climate policy in the U.S.: all other state or national climate policies regulate specific sectors, whereas AB 32 covers all GHG emission sources.

To meet this GHG target, AB 32 established a suite of climate programs. One key program was cap-and-trade, introduced in 2013 and administered by the California Air Resources Board (CARB).[8] The program requires participation by all stationary GHG-emitting facilities producing at least 25,000 metric tons of annual carbon dioxide equivalent emissions, or CO_2e, during any year between 2009-2012.[9] This eligibility criteria covers all sectors that directly emit GHGs from stationary sources and is unique amongst other AB 32 climate programs.[10,11] California's C&T program has since created the world's second largest carbon market by permit value, following the European Union Emissions Trading System (EU-ETS).

In 2016, California met AB 32's 2020 GHG target four years early. That same year, the state extended its GHG target to 40% below 1990 levels by 2030. This was shortly followed by a 2030 extension of the C&T program. However, critical questions remain regarding the performance and consequences of the C&T program.

First, it remains unclear whether C&T has lower GHG emissions. In particular, when C&T coexists with complementary climate programs, an overall GHG emissions cap can be met with little or no abatement induced by C&T if these complementary programs bind. Indeed, an ex-ante analysis of California's GHG C&T program demonstrated a potentially large role played by such complementary programs on overall GHG abatement (Borenstein et al., 2019).[12] Second, even if C&T lowered GHG and local air pollution emissions, it remains unclear whether the resulting change in the spatial distribution of air pollution concentrations widens or narrows California's EJ gap.

---

[8]Prominent complementary programs to C&T under AB 32 include a Renewable Portfolio Standard for electricity generation and a Low Carbon Fuel Standard for refineries.

[9]Greenhouse gases covered by the program were CO_2, CH_4, N_2O, HFCs, PFCs, SF6, NF3 and other fluorinated GHGs. The 25,000 metric ton eligibility criteria is re-evaluated

[10]The 2013 timing of the C&T program is also unique. Most other AB 32 climate programs were introduced earlier.

[11]The GHG C&T program does not directly regulate local criteria air pollution emissions. Any changes in the spatial distribution of local air pollution concentration due to the program is driven by the program's reallocation of local air pollution emissions that is co-produced with GHG emissions.

[12]Furthermore, even a positive GHG permit price does not ensure that the C&T program caused GHG emissions to fall. Suppose, for example that there was some form of restriction on GHG emissions prior to the C&T program leading to a pre-program positive shadow price on GHG abatement. A C&T program with an overall cap set equal to total emissions under the prior restriction would generate a positive permit price despite no change in overall GHG emissions.

3 Data

Our analysis involves two primary datasets: 1) GHG and criteria air pollution emissions at the facility-by-year level and 2) an indicator of whether a zip code is considered to be "disadvantaged" according to California legislation.

Facility emissions   We obtain 2008-2017 facility-level annual emissions of GHG (or CO2e), $PM_{2.5}$, PM_{10}, NO_x, and SO_x, all in metric tons, from CARB's Pollution Mapping Tool.[13] We observe GHG as well as criteria air pollutions emissions for both C&T-regulated and non-regulated stationary facilities, before and after the introduction of the C&T program.[14]

Several additional facility-level variables serve as inputs for the atmospheric dispersal model. CARB provides facility latitude and longitude as well as pollution-specific stack heights for a subset of facilities. For other facilities, we impute missing pollution-specific stack heights using sector averages constructed from non-missing observations.

Definition of a disadvantaged community   There is no established definition of a "disadvantaged" community. Previous papers in other settings use a location's median income or minority share of population as proxy measures (Fowlie, Holland and Mansur, 2012; Grainger and Ruangmas, 2018; Mansur and Sheriff, 2019). For our setting, we select a policy-relevant definition of a "disadvantaged" community. Senate Bill 535 (SB 535), passed in 2012, requires a portion of the revenue from the auction of C&T permits to be directed towards benefiting disadvantaged communities. SB 535 formally defines a "disadvantaged community" using CalEnviroScreen, a scoring system based on multiple indicators developed by the California Environmental Protection Agency. Specifically, a zip code is considered disadvantaged if it contains all or part of a census tract with a CalEnviroScreen score above the top 25th percentile. Zip codes designated as disadvantaged are shaded in dark blue in Figure 2a. Importantly, pre-2013 data was used in constructing CalEnviroScreen, which mitigates the concern that cap-and-trade may have affected zip code designation. We further augment

---

[13]Available here: https://ww3.arb.ca.gov/ei/tools/pollution_map/

[14]Stationary facilities with annual emissions past a certain threshold must report emissions. For GHGs, the data reporting threshold is 10,000 metric tons of CO_2e, set by CARB. For criteria air pollutants, CARB sets a reporting threshold of 10 metric tons per year, but each air district can set lower data reporting thresholds. As a consequence, we observe criteria air pollution emissions below 10 metric tons, with no evidence of bunching at 10 tons (see histograms of sample facility-year emissions in Figure S1). We confirmed that emissions data in CARB's Pollution Mapping Tool matches values found in source datasets: CARB's Mandatory Reporting Regulation (MRR) dataset for GHG emissions and the California Emissions Inventory Development and Reporting System (CEIDARS) for criteria air pollution emissions. Details on California's emissions reporting requirements can be found: https://ww3.arb.ca.gov/ei/tools/pollution_map/doc/caveats%20document12_22_2017.pdf

our zip code level data with average 2008-2012 population obtained from the U.S. Census Bureau.

4 Empirical approach

Our analysis proceeds along three steps. First, we use facility-by-year-level data to estimate how the GHG C&T program altered GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x emissions. Second, we feed C&T-driven $PM_{2.5}$, PM_{10}, NO_x, and SO_x emissions predicted from the first step into an atmospheric dispersal model to generate zip code-by-year-level concentrations of these pollutants due to the program. Finally, we examine whether the C&T program changed the concentration gap for these pollutants between disadvantaged and other communities following its 2013 introduction.

Step 1: Estimating C&T effects on emissions We exploit the facility-level eligibility criteria based on pre-program GHG emissions and the 2013 timing of the C&T program to identify its effects on GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x facility-level emissions during 2008-2017. Because the program's eligibility criteria is based on pre-C&T GHG emissions, we expect regulated and unregulated facilities to differ in pre-program emissions levels and perhaps also in pre-program emission trends. Our empirical test therefore examines whether differential emission trends exhibit a break after 2013. For this test to have a causal interpretation, our identifying assumption requires that any existing differential emission pre-trends to have continued if not for the introduction of the C&T program.[15]

Specifically, let j index facilities. $C_j \in \{0,1\}$ is GHG C&T regulatory status with $C_j = 1$ indicating facility j is regulated.[16] For facility j in year t, Y_{jt}^p is annual emissions of pollutant $p \in \{GHG, PM_{2.5}, PM_{10}, NO_x, SO_x\}$. Because emissions exhibit a skewed distribution and contain zero values, we apply an inverse hyperbolic sine transformation, which like a log

---

[15] Because there is no overlap in pre-program GHG emissions for regulated and unregulated facilities, we are unable to implement a matching estimator that matches on pre-program emissions, as is done in Fowlie, Holland and Mansur (2012) and Martin, Muûls and Wagner (2016). Implementing such a matching approach would require emissions data from facilities outside of California. That comparison, however, may be confounded by systematic unobserved differences between California and non-California facilities.

[16] All but 39 facilities that emit local air pollution found in CARB's Pollution Mapping Tool have time-invariant GHG C&T regulatory status between 2008-2017. These 39 facilities all switched status in 2017. Under the C&T program, a regulated (unregulated) facility can become unregulated (regulated) if annual GHG emissions fall below (above) the 25,000 metric tons threshold in any year during a prior compliance period. Of the 39 facilities that switched status in 2017, 8 switched even though annual GHG emissions during the previous 2015-2016 compliance period should not have permitted a regulatory status change. Because we do not know if these switches are due to actual changes in regulatory status or coding errors, we retain these 39 facilities in our sample and re-assign them their previous (time-invariant) regulatory status for 2017. In a robustness check, we drop observations from these 39 facilities in our estimation.

transformation lends a percentage effect interpretation, but with the added advantage of retaining zero-valued observations (Bellemare and Wichman, 2020). To examine differential emission trends driven by the C&T program, we estimate the following specification:

$$asinh(Y_{jt}^p) = \kappa_1^p[C_j \times t] + \kappa_2^p[C_j \times \mathbf{1}(t \geq 2013) \times t] + \phi_j^p + \gamma_t^p + \nu_{jt}^p \qquad (1)$$

Facility-specific dummy variables ϕ_j^p removes time-invariant determinants of pollution p for facility j. Year-specific dummy variables γ_t^p remove common determinants of emissions affecting all sample facilities in year t, such as California-wide economic conditions.

κ_1^p captures the differential emission pre-trend for pollutant p between facilities that would and would not eventually be regulated by the C&T program during 2008-2012, reported in annual percentage point changes. κ_2^p is the change, or break, in the differential emission trend after the program's introduction during 2013-2017. ν_{jt}^p is clustered at the county-level to allow for arbitrary forms of heteroskedasticity and serial correlation within a county.

We employ two sample restrictions to strengthen identification of trend break effects in equation (1). First, despite the C&T program's unique eligibility criteria and timing, the presence of other major climate programs under AB 32, such as the Renewable Portfolio Standard for electricity generators and the Low Carbon Fuel Standard for refineries, may confound C&T effects for these facilities. We remove electricity generators and refineries from our sample to avoid this possibility.[17] Second, to ensure better comparability between treated and control facilities, we restrict our sample to facilities with sample average annual GHG emissions below the 75th percentile.[18] As a robustness check, we consider smaller and larger cutoff percentiles.

Our benchmark sample contains 106 regulated and 226 unregulated facilities. Each regulated facility is shown as a black dot in Figure 2a. Table S1 shows average 2008-2012 annual GHG and criteria emissions and sectoral distribution for sample regulated and unregulated facilities. Since C&T regulatory status is defined by historical GHG emissions, it is unsurprising that regulated and unregulated facilities exhibit different average pre-program emissions, nor does this invalidate our differential emissions trend break design, per se. Table S1 also shows a slight sectoral imbalance between regulated and unregulated facilities, with more regulated facilities in extraction and more unregulated facilities in services. In a robustness check, we replace year fixed effects in equation (1) with sector-by-year fixed effects to address concerns that this sectoral imbalance may confound our estimates.

To construct facility-by-year emissions driven by the C&T program (relative to California-

---

[17] This restriction also addresses concerns about the the 2013 closure of the San Onofre Nuclear Generating Station, a power plant in southern California (Davis and Hausman, 2016).

[18] The 75th percentile corresponds to average annual emissions of 62,770 metric tons of CO_2e.

wide determinants of pollution), we apply a hyperbolic sine transformation to the first two terms of equation (1) and the estimated facility-level fixed effect.[19] Because facilities differ by average emission levels, the inclusion of facility-level fixed effects allows us to generate heterogeneous C&T-driven pollution abatement across regulated facilities despite estimating a common percentage effect.[20] This implicitly assumes that larger emitting facilities abate more under C&T. To examine this assumption, in a robustness check, we estimate variants of equation (1) that allow the post-C&T trend break to vary as linear and quadratic functions of facility-level average annual emissions.

Step 2: Modeling pollution dispersal Our second step determines how C&T-driven criteria air pollution disperses spatially across California. The standard approach is for the researcher to prescribe the set of locations affected by emissions from a particular source, either by assuming emissions only disperses within areas in the same administrative unit of the source or within a radially uniform distance from the source. For example, one may assume emissions from a facility in Los Angeles County only affect Los Angeles County or areas within a certain radial distance of that facility. Actual affected areas, however, may not conform to these assumptions and instead may vary depending on topography or time-varying meteorological conditions. To fully capture the complexity of pollution dispersal, we turn to an atmospheric dispersal model.

We feed predicted facility-by-year $PM_{2.5}$, PM_{10}, NO_x, and SO_x emissions from step 1, together with the location and stack height of each facility, into the Hybrid Single Particle Lagrangian Integrated Trajectory Model (HYSPLIT), an atmospheric dispersal model developed by the U.S. National Oceanographic and Atmospheric Administration (NOAA) with meteorological conditions from NOAA's 40-km resolution North American Model Data Assimilation System (NAMDAS) (Draxler and Hess, 1998). An emerging literature uses HYSPLIT to convert pollution emissions to concentrations (Grainger and Ruangmas, 2018; Henneman, Mickley and Zigler, 2019; Casey et al., 2020).

We choose HYSPLIT because it provides a middle-of-the-road approach for our application, balancing atmospheric realism with computational tractability. HYSPLIT is less

---

[19] Specifically, C&T-driven emissions is:

$$\widehat{Y}_{jt}^p = sinh\left(\widehat{\kappa}_1^p[C_j \times t] + \widehat{\kappa}_2^p[C_j \times \mathbf{1}(t \geq 2013) \times t] + \widehat{\phi}_j^p\right) * e^{(RMSE)^2/2}$$

where hat notation indicates estimated parameters and RMSE is the root mean squared error from equation (1). In theory, the hyperbolic sine transformation can generate negative emission values. In practice, our benchmark model predicts negative emissions for 1%, 1%, 0.2%, and 0.3% of sample observations for $PM_{2.5}$, PM_{10}, NO_x, and SO_x, respectively. We replace these negative values with zeros.

[20]For example, a 10% abatement effect implies 10 tons of abatement for a facility with 100 tons of average annual emissions and 5 tons of abatement for a facility with 50 tons of average annual emissions.

computationally intensive than chemical dispersal models such as WRF-Chem, but at the cost of not incorporating atmospheric chemistry which is important for modeling secondary pollutant formation. At the same time, HYSPLIT is more reliable for modeling pollution dispersal beyond distances of 50 kilometers, which less computationally-intensive Gaussian-plume models like AERMOD or APEEP do poorly (EPA, 2015).

We note several features of our HYSPLIT implementation. First, to account for high-frequency variation in meteorological conditions, we run forward particle trajectories at four hour intervals, implicitly assuming that annual emissions are distributed uniformly within the year. Each trajectory runs for 24 hours, a duration long enough to ensure most emitted particles leave California.[21] Second, because HYSPLIT does not explicitly account for particle decay, we apply half-life parameters from the atmospheric chemistry literature set at 24 hours for $PM_{2.5}$ and PM_{10}(U.S. EPA, 2018), 3.8 hours for NO_x (Liu et al., 2016), and 13 hours for SO_x (Lee et al., 2011). Third, we assume that a particle no longer contributes to surface pollution concentrations once it exits the planetary boundary layer, beyond which there is far less turbulent mixing. We conservatively set the boundary layer height at 1 km above the surface, which is about double the typical height for California (Rahn and Mitchell, 2016). As a robustness check, we also consider boundary layer heights of 0.5 and 2 km. As an illustration of pollution dispersal modeled by HYSPLIT, Figure 2b shows the trajectories of pollution emitted by a regulated facility in Los Angeles during 2016. In total, we compute over 2 million particle trajectories from the roughly one hundred regulated facilities in our sample during the 2008-2017 period. This procedure takes about 24 hours to complete using over one thousand facility-by-year parallelized nodes on a high-performance computing cluster.

HYSPLIT generates particle-level trajectories. To convert this into concentration units, we sum HYSPLIT trajectories for each zip code and year and divide by the volume of the atmosphere between a zip code's surface and the boundary layer. We further divide by 365 days. This gives us a zip code-by-year measure of average daily C&T-driven pollution concentration for the 1 km-high air column above each zip code in units of $\mu g/m^3/day$.[22] Figure 2c shows our benchmark HYSPLIT-generated daily concentration (in $\mu g/m^3/day$) for each zip code, averaged across 2008-2017 for $PM_{2.5}$. Figure S4 similarly shows average 2008-

---

[21]Unlike Henneman et al. (2019), we do not discard the first hour of each particle trajectory because doing so may omit highly localized pollution concentrations that may be important for our distributional analysis.

[22]Other HYSPLIT applications convert HYSPLIT particles into concentration units by regressing HYSPLIT output onto concentration output from a different atmospheric dispersal model using the same emissions sources (see for example: Henneman, Choirat and Zigler (2019)) to obtain predicted concentrations using that fitted relationship. We are unable to perform that adjustment as there are no alternative measures of C&T-driven pollution concentrations in the literature.

Figure 2: Modeling air pollution concentrations driven by the cap-and-trade program
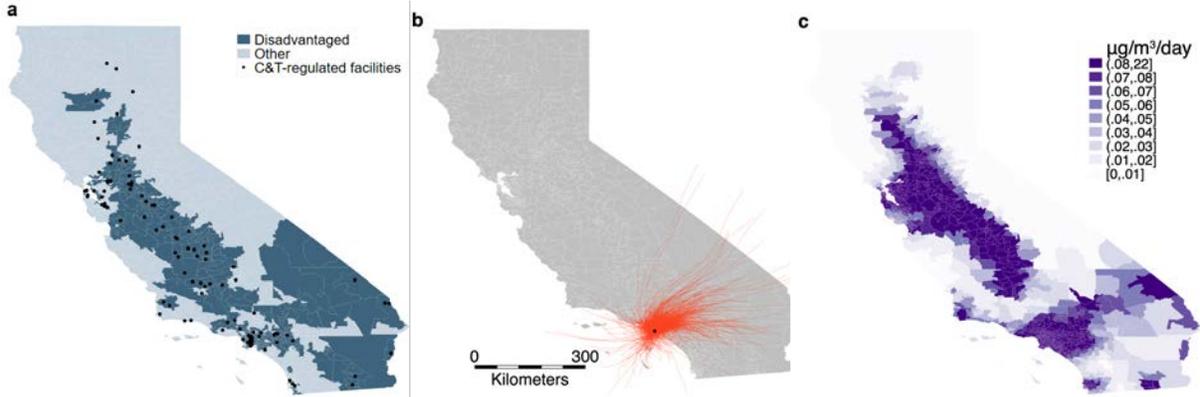
NOTES: Panels illustrates how facility-level emissions is converted to zip code-level pollution concentrations using an atmospheric dispersal model. Shading in panel (a) shows California zip codes that are designated as disadvantaged (dark blue) and zip codes that are not (light blue) according to California policy. Black dots show sample facilities regulated by California's GHG C&T program. Panel (b) shows HYSPLIT-generated particle trajectories every 4-hours from a regulated facility during 2016. Panel (c) shows zip code-level average daily $PM_{2.5}$ concentrations (in $\mu g/m^3$/day) during 2008-2017 driven by facilities regulated by the C&T program as modeled by HYSPLIT.

2017 zip-code concentrations for PM_{10}, NO_x, and SO_x.[23] Note that pollution concentration levels in Figure S4 are generally below those recorded in ambient monitors because we are only considering pollution concentrations driven by C&T-driven emissions from sample regulated facilities.

Lastly, as noted, a major limitation with HYSPLIT is that it does not model secondary pollution formation. To see if secondary $PM_{2.5}$ concentrations exhibits a different spatial pattern than primary $PM_{2.5}$ concentrations, in a robustness check, we replace HYSPLIT with InMAP, a reduced-complexity dispersal model based on the WRF-Chem model which generates secondary pollutants (Tessum, Hill and Marshall, 2017).

Step 3: Estimating C&T-driven change in EJ gap trends In our third step, we examine whether the C&T program altered the difference in pollution concentrations between disadvantaged and other communities, or the EJ gap. Let $D_i \in \{0,1\}$ denote disadvantaged status, with $D_i = 1$ indicating that zip code i contains all or part of a "Disadvantaged Community Census Tract," as defined by Senate Bill 535. For zip code i in year t, we take C&T-driven pollution concentration from HYSPLIT, E_{it}^p, for criteria air pollutant $p \in \{PM_{2.5}, PM_{10}, NO_x, SO_x\}$, and estimate the following specification:

$$E_{it}^p = \beta_1^p[D_i \times t] + \beta_2^p[D_i \times \mathbf{1}(t \geq 2013) \times t] + \psi_i^p + \delta_t^p + \epsilon_{it}^p \tag{2}$$

---

[23]Figure 2, Figure S4, and Table S6 show that criteria air pollution from GHG C&T-regulated facilities disperses across all of California and not just zip codes designated as disadvantaged.

where ψ_i^p are zip code-specific dummies and δ_t^p are year-specific dummies. β_1^p, or the pre-C&T EJ gap trend, captures the linear trend in the EJ gap (from facilities that would eventually be regulated by the C&T program) during 2008-2012, before the program was introduced. A positive trend (i.e., $\beta_1^p > 0$) would indicate that the EJ gap was widening prior to the C&T program. Our main parameter of interest is β_2^p, which captures the change in the EJ gap trend after the program's introduction, or the post-C&T EJ gap trend break. Conditional on $\beta_1^p > 0$, $\beta_2^p < 0$ implies that the introduction of the C&T program slowed the previous positive EJ gap trend. We consider two additional trend break statistics. The first statistic asks whether the post C&T EJ gap trend break is sufficiently large such that the EJ gap has actually narrowed in level terms after the C&T program. This would be captured by $\beta_1^p + \beta_2^p$, or the post-C&T EJ gap trend, with $\beta_1^p + \beta_2^p < 0$ indicating that the EJ gap is narrowing.[24] A second statistic examines the relative degree in which C&T program has slowed the prior EJ gap trend. Specifically, $\frac{\beta_2^p}{\beta_1^p} * 100 = (\frac{(\beta_1^p + \beta_2^p) - \beta_1^p}{\beta_1^p}) * 100$ captures the percentage change in the EJ gap trend following the introduction of the C&T program.

C&T-driven pollution concentration, E_{it}^p, the outcome variable in equation (2), is predicted C&T-driven emissions from equation (1) via HYSPLIT. As a consequence, ϵ_{it}^p, the error term in equation (2), does not account for statistical uncertainty in C&T emission effects from equation (1). Instead, ϵ_{it}^p may capture residuals that arise when estimating an average EJ effect in the presence of heterogeneous EJ effects. To address inference concerns, we conduct two standard error adjustments. First, we cluster ϵ_{it} at the county level to allow for arbitrary forms of heteroskedasticity and serial correlation when heterogeneous treatment effects are not independent and identically distributed. Second, to incorporate statistical uncertainty in predicted C&T-driven emissions from equation (1), we conduct a bootstrap procedure drawing multiple vectors of C&T-driven emissions from the estimated empirical distributions of κ_1^p and κ_2^p, which are then fed into steps 2 and 3. In practice, we implement 250 bootstrap draws to generate a component of the standard error for β_1^p and β_2^p that accounts for statistical uncertainty in equation (1). We add this component to the standard error from directly estimating equation (2) when reporting uncertainty for β_1^p and β_2^p. Figure S5 plots the empirical distribution of β_1^p and β_2^p across bootstrapped draws.[25]

---

[24]Observe that while $\beta_2^p < 0$ alone implies that the C&T program resulted in EJ gap benefits by slowing the growth in the EJ gap, it does not necessarily imply that this post-trend break effect is strong enough to offset the magnitude of the pre-trend such that EJ gap is narrowing in absolute terms following the program. For that to occur, one needs $\beta_2^p < -\beta_1^p$, or $\beta_1^p + \beta_2^p < 0$.

[25]As with prior literature, we omit uncertainty associated with atmospheric dispersal, or the mapping between facility-level emissions and zip code-level concentration. One possibility involves resampling meteorological conditions in HYSPLIT via a bootstrapping algorithm. Given that our use of HYSPLIT takes 24 hours, overlaying such an approach to the existing 3-step procedure is currently unrealistic under available computational resources.

Appendix A.1 provides more details on this bootstrap procedure.

Finally, to estimate an average EJ gap effect across individuals in California, we weight each zip code-by-year observation in equation (2) by average zip code population during 2008-2012, the period prior to the program.

Comparison with prior uses of pollution dispersal models   Our empirical approach is part of a broader effort across natural and social sciences to use atmospheric dispersal models to map pollution emissions to concentrations. Prior studies can be broadly classified into two groups: whether the analysis is done at the location-level or at the facility-level.

Location-level analyses typically feed observed emissions into a dispersal model, but without first estimating the emissions effects of environmental policies (Ash and Fetter, 2004; Morello-Frosch and Jesdale, 2006; Sullivan, 2017; Cummiskey et al., 2019; Henneman et al., 2019; Henneman, Choirat and Zigler, 2019; Kim et al., 2020). Because these studies omit estimation of policy-driven emissions (i.e., our Step 1), they cannot attribute changes in pollution concentrations to specific policies.[26]

Facility-level studies examine whether a policy's effect on emissions varies with the demographic characteristics of households downwind of facilities, as determined by the atmospheric dispersal model (Grainger and Ruangmas, 2018; Mansur and Sheriff, 2019). This approach augments the facility-level in equation (1) by adding a term that interacts the policy treatment with demographic characteristics of downwind locations. However, given the complex spatial nature of pollution disperal whereby concentrations in multiple locations may be affected by emissions from multiple facilities, it is not obvious whether one can recover EJ gap changes, the estimand of interest, from such an approach.

In Appendix A.2, we formally demonstrate that the coefficient on the interaction term from such dispersal-augmented facility-level regressions does not in general equal the EJ gap effect, nor does it necessarily have the same sign, making it hard to draw EJ gap conclusions from such regressions. We then show one special case where equality does hold but which requires - among other assumptions - a particularly strong assumption on the spatial pattern of pollution dispersal: emissions from each regulated facility must affect either *only* disadvantaged communities or *only* non-disadvantaged communities. That is, facilities cannot alter pollution concentrations in both types of locations. This assumption can be rejected in our setting: panels a and b of Figure 2 show how emissions from one regulated sample facility alters pollution concentrations in both disadvantaged and other

---

[26]For example, Henneman et al. (2019) and Henneman, Choirat and Zigler (2019) insert observed air pollution emissions from coal-fired power plants into a version of HYSPLIT to examine how much U.S. $PM_{2.5}$ concentrations are due emissions from these plants, but cannot speak to the policies that are affecting coal-fired power plant emissions.

communities.

Our approach combines both facility- and location-level analyses. As such, we are able to attribute changes in emissions due to the C&T program and quantify the resulting change in the EJ gap as a consequence of these emissions.

5   Results

This section presents our results. Section 5.1 shows the effect of C&T on differential emission trends between regulated and unregulated facilities. Section 5.2 examines how these C&T-driven emissions altered trends in the pollution concentration gap between disadvantaged and other communities across California.

5.1   Cap-and-trade effects on emissions

Main results   Table 1 reports the pre-C&T differential emissions trend (i.e., κ_1^p from equation (1)) and the post-C&T differential emissions trend break (i.e., κ_2^p from equation (1)) for GHG and criteria air pollutants. Column 1 shows a statistically significant trend break in GHG emissions, indicating that the C&T program led to a reduction in GHG emissions. Prior to the program, the gap in GHG emissions between regulated and unregulated facilities increased at an annual rate of 19 percentage points. Following the introduction of the program, this trend slowed by 30 percentage points leading the gap in GHG emissions to fall at an annual rate of 11 percentage points between 2012-2017. For criteria air pollutants, columns 2-4 show a statistically significant, negative emissions trend break following the program's introduction for $PM_{2.5}$, PM_{10}, NO_x. For SO_x, the trend break is negative but not statistically significant, suggesting that all subsequent SO_x results should be interpreted with caution.

We predict C&T-driven emissions using estimates in Table 1 together with facility-level fixed effects. This generates heterogeneous facility-level C&T-driven abatement between 2012-2017, or $\widehat{Y}_{j,2017}^p - \widehat{Y}_{j,2012}^p$ as defined in footnote 19, and shown in Figure S3 for GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x. Averaged across sample regulated facilities, between 2012 and 2017, the C&T program reduced emissions annually at a rate of 9%, 5%, 4%, 3%, and 9% for GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x, respectively.[27] GHG permit prices for the California C&T program have largely hovered above the program's price floor since 2013. Detecting emissions abatement from sectors directly regulated by only C&T, however, is consistent

---

[27]This is calculated by averaging $(\frac{\widehat{Y}_{j,2017}^p - \widehat{Y}_{j,2012}^p}{\widehat{Y}_{j,2012}^p})/5$, as defined in footnote 19, across regulated sample facilities for each pollutant p.

17

Table 1: Trend break in emissions

	(1) GHG	(2) $PM_{2.5}$	(3) PM_{10}	(4) NO_x	(5) SO_x
		Outcome is (asinh) emissions			
κ_1^p	0.187	0.058	0.083	0.075	0.006
	(0.052)	(0.043)	(0.033)	(0.039)	(0.035)
	[0.001]	[0.183]	[0.016]	[0.061]	[0.875]
κ_2^p	-0.297	-0.097	-0.117	-0.104	-0.037
	(0.077)	(0.048)	(0.039)	(0.050)	(0.043)
	[0.000]	[0.053]	[0.005]	[0.042]	[0.393]
$\kappa_1^p + \kappa_2^p$	-0.111	-0.039	-0.034	-0.029	-0.031
	(0.036)	(0.018)	(0.018)	(0.019)	(0.019)
	[0.004]	[0.039]	[0.068]	[0.138]	[0.108]
Facilities	316	302	302	303	303
Counties	41	40	40	40	40
Observations	2,054	1,968	1,968	1,970	1,965

NOTES: Estimates of pre-C&T differential emissions trend (i.e., κ_1^p from equation (1)) and post-C&T differential emissions trend break (i.e., κ_2^p from equation (1)) for GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x across columns. All models include facility-specific and year-specific dummy variables. Standard errors clustered at the county-level in parentheses, p-value in brackets.

with permit prices at the price floor when complementary climate programs bind for other C&T-covered sectors. When such programs bind, aggregate demand for GHG permits fall, causing permit prices to hit the price floor. However, provided that total abatement driven by complementary programs is insufficient for meeting the total GHG cap, C&T will still induce abatement from sectors that are only regulated by C&T.

In total, sample regulated facilities reduced 3.2 million tons of CO_2e between 2012-2017. This figure is likely a weak lower bound on the effect of C&T on total California GHG emissions as C&T is likely to have weakly negative emissions effects on regulated facilities subject to complementary climate programs but are excluded from our analysis, As a point of reference, California's statewide GHG emissions fell by 27.1 million tons of CO_2e (Board, 2020) between 2012-2017.

Robustness checks   We subject these emission effects to several robustness checks. First, Figure S2 considers placebo program start years, plotting κ_2^p for GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x emissions from variants of equation (1) that impose alternative C&T start years

18

across 2009-2016. With the exception of the SO_x results, we detect the strongest trend break coefficient when we assign the treatment year to its actual occurrence in 2013.

Table S2 considers several alternative specification and sample restriction choices. Table S1 shows that regulated and unregulated facilities are not perfectly balanced across sectors. To address concerns that differential trends across sectors may confound our estimates, column 1 of Table S2 replaces year fixed effects with sector-by-year fixed effects. Column 2 drops the handful of facilities whose treatment status switched only in 2017. Columns 3 and 4 change the 75th percentile average GHG emissions cutoff to the 70th and 80th percentiles.[28] None of these robustness checks produces estimates that differ meaningfully from our benchmark estimates in Table 1.

Our C&T-driven emissions which includes facility fixed effects, implicitly assumes more pollution abatement from facilities that emit more on average. To examine whether this assumption is reasonable, column 2 of Table S3 reports a variant of equation (1) that further includes an interaction between the trend break term and a linear function of facility-level average emissions. A positive interaction coefficient would imply that larger emitting facilities are abating less, contradicting our assumption. With the exception of GHG emissions for which the linear interaction term is positive but of very small magnitude, the coefficient on this interaction term for every criteria air pollution is negative. This suggests that our benchmark model, which estimates an average trend break coefficient across facilities (regardless of size) is understating the degree in which large-emitting facilities are also abating more under C&T. Column 3 of Table S3 shows that heterogeneity by average emissions does not exhibit nonlinearity, as indicated by statistically imprecise quadratic interaction terms.

Finally, there may be a Stable Unit Treatment Value Assumption (SUTVA) as pollution may shift from a regulated to unregulated facilities following the introduction of C&T. If so, the resulting increase in unregulated facility emissions may lead to more negative estimates of the trend break parameter κ_2^p. Following Fowlie, Holland and Mansur (2012), we consider two robustness checks in Table S4 to examine this possibility. In the first test, we observe that a facility located in a county under U.S. Clean Air Act nonattainment for a particular pollutant is largely unable to increase pollution levels. This idea is implemented in column 2, which restricts the sample of unregulated facilities to those located in nonattainment counties for that pollutant under the Clean Air Act.[29] Our second test notes that firms with

---

[28]The 70th and 80th percentiles for sample average annual GHG emissions corresponds to 48,834 and 82,173 tons of CO_2e, respectively.

[29]In Table S4, column 2 does not apply to GHG emissions because it is not a criteria pollutant regulated under the Clean Air Act. For SO_x, there are no counties in nonattainment during our sample period. For NO_x, because there were not enough counties under NO_2 nonattainment to construct a control group, we follow Fowlie, Holland and Mansur (2012) by looking at nonattainment under Clean Air Act's one-hour ozone standard as NO_x is a precursor pollutant to ozone.

19

multiple facilities can more readily reallocate pollution across their facilities. In column 3, we restrict the control group of unregulated facilities to those whose parent company only operates a single plant.[30] If treatment spillovers were present, the trend break coefficient κ_2^p should be of smaller magnitude in columns 2 and 3 than in our benchmark estimate, shown in column 1. This is not the case.

5.2 Cap-and-trade effects on EJ gaps

Validating pollution dispersal modeling We consider two sensibility checks for our measure of C&T-driven pollution concentrations via HYSPLIT before turning to our main EJ gap results. First, we examine whether HYSPLIT-generated criteria air pollution concentrations correlate with monitored ambient air pollution concentrations. Specifically, we match zip code-level HYSPLIT-generated pollution concentration averaged over 2008-2017 to the average ambient pollution concentration of that zip code as recorded by pollution monitors averaged over the same period, obtained from the U.S. Environmental Protection Agency.[31] we do not expect a perfect fit between these two variables as ambient pollution at any location is composed of emissions originating from many more sources (i.e., stationary and non-stationary, within and beyond California) than our subset of stationary sources regulated by California's GHG C&T program. However, a positive correlation between the two pollution concentration measures would provide reassurance that HYSPLIT-generated pollution concentration from C&T regulated facilities is detected by ambient pollution monitors. The positive correlations shown in Table S5 indicate that is indeed the case.[32]

Next, we examine the EJ gap in 2008 driven by facilities that would eventually be regulated by the C&T program. Prior work documented strong baseline EJ gaps in California (Cushing et al., 2018). Indeed, this baseline EJ gap informed initial EJ concerns regarding California's C&T program. Table S6 shows that steps 1 and 2 of our approach reproduces EJ gaps in 2008. Disadvantaged communities experienced higher levels of $PM_{2.5}$, PM_{10}, NO_x, and SO_x concentrations in 2008 than other communities on average due to emissions from facilities that would eventually be regulated by the C&T program.

---

[30]We link each facility from CARB with its parent company as indicated by the EPA. We employ a fuzzy string matching algorithm as facility names are not standardized across the two datasets.

[31]Available here: https://aqs.epa.gov/aqsweb/airdata/download_files.html

[32]We are interested in modeling where C&T-driven pollution is dispersed. As such, we do not directly use ambient pollution data (either from ground-based monitoring stations or remotely-sensed satellites) in our analysis as it is often difficult to determine which component of any location's ambient pollution originates from C&T-regulated facilities. Such "backwards" atmospheric modeling often yield indeterminate results.

Table 2: Trend break in the environmental justice gap

	(1) PM$_{2.5}$	(2) PM$_{10}$	(3) NO$_x$	(4) SO$_x$
β_1^p	0.042	0.065	0.085	0.037
	(0.015)	(0.017)	(0.037)	(0.025)
	[0.006]	[0.000]	[0.026]	[0.151]
β_2^p	-0.063	-0.090	-0.143	-0.101
	(0.022)	(0.029)	(0.074)	(0.051)
	[0.006]	[0.003]	[0.060]	[0.053]
$\beta_1^p + \beta_2^p$	-0.021	-0.026	-0.058	-0.064
	(0.015)	(0.020)	(0.050)	(0.027)
	[0.159]	[0.203]	[0.252]	[0.024]
$(\beta_2^p/\beta_1^p) * 100$	-149.699	-139.739	-168.282	-272.291
	(36.368)	(29.971)	(53.375)	(66.043)
	[0.000]	[0.000]	[0.002]	[0.000]
Zip codes	1649	1649	1649	1649
Counties	58	58	58	58
Observations	16,416	16,416	16,416	16,416

NOTES: Estimates of the pre-C&T EJ gap trend (i.e., β_1^p from equation (2)), the post-C&T EJ gap trend break (i.e., β_2^p from equation (2)), the post-C&T EJ gap trend (i.e, $\beta_1^p + \beta_2^p$), and the percentage change in the EJ gap trend following the introduction of the C&T program (i.e., $\frac{\beta_2^p}{\beta_1^p} * 100$) for PM$_{2.5}$, PM$_{10}$, NO$_x$, and SO$_x$, across columns. All models include zip code-specific and year-specific dummy variables. Observations weighted by zip code-level average population during 2008-2012. Parentheses indicate standard errors that account for statistical uncertainty in C&T predicted emissions (ν_{it}^p from equation (1) via the bootstrap procedure in Appendix A.1) and county-level heterogeneity in EJ gap effects of arbitrary form (ϵ_{it}^p from equation (2)). P-value in brackets.

Main results   We now turn to our main results examining the time evolution of EJ gaps between 2008-2017, shown in Table 2 and Figure 3. Across PM$_{2.5}$, PM$_{10}$, NO$_x$, and SO$_x$, the EJ gap widens during 2008-2012, the period prior to the C&T program, as indicated by the positive pre-C&T EJ gap trend (i.e., β_1^p from equation (2)). Following 2013, the EJ gap trend falls: the post-C&T EJ gap trend break (i.e., β_2^p from equation (2)) is negative and statistically significant. This drop in the EJ gap trend is sufficiently large such that the EJ gap is actually narrowing following C&T, as indicated by the negative post-C&T EJ gap trend across pollutants (i.e., $\beta_1^p + \beta_2^p$). In percentage terms (i.e., $\frac{\beta_2^p}{\beta_1^p} * 100$), the EJ gap

trend fell between 140-270% across pollutants after the program's introduction. Figure 3 plots this trend break as well as annual EJ gap coefficients from a more flexible version of equation (1) using year-specific EJ gap coefficients.[33] Figure 3 also highlights that while the C&T program has led EJ gaps to narrow since 2012, it has not eliminated them. By 2017, EJ gaps are roughly at 2008 levels across pollutants.

Figure 3: Environmental justice gap before and after the cap-and-trade program
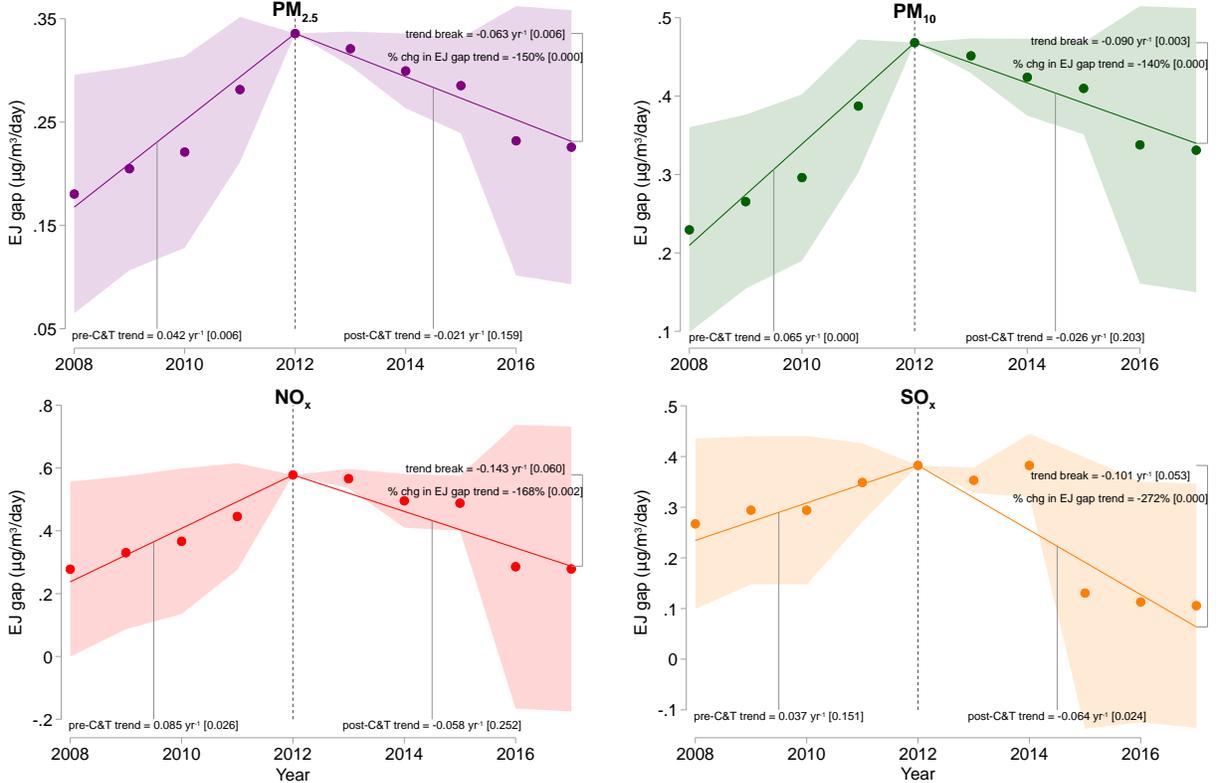
NOTES: Panels show the estimated average daily pollution concentration gap (in $\mu g/m^3$/day) between disadvantaged and other zip codes (i.e., "EJ gap") during 2008-2017 for $PM_{2.5}$, PM_{10}, NO_x, and SO_x, respectively. Dots show year-specific EJ gap. Solid lines show linear fit for EJ gap trend before (2008-2012) and after (2013-2017) the C&T program. Associated text indicates point estimates and standard errors for the pre-C&T linear trend, post-C&T trend break, post-C&T linear trend, and the percentage change in the EJ gap trend (i.e., β_1^p, β_2^p, $\beta_1^p + \beta_2^p$, $\frac{\beta_2^p}{\beta_1^p} * 100$). 95% confidence interval and p-values (in brackets) account for statistical uncertainty in C&T predicted emissions (ν_{it}^p from equation (1) via the bootstrap procedure in Appendix A.1) and county-level heterogeneity in EJ gap effects of arbitrary form (ϵ_{it}^p from equation (2)). Trend break estimates also reported in Table 2.

---

[33]Specifically, the annual coefficients in Figure 3 are β_τ^p from

$$E_{it}^p = \sum_{\substack{2008 \leq \tau \leq 2017 \\ \tau \neq 2012}} \beta_\tau^p [D_i \times \mathbf{1}(t = \tau)] + \psi_i^p + \delta_t^p + \epsilon_{it}^p$$

Figure 4: Spatial heterogeneity in EJ gap effects
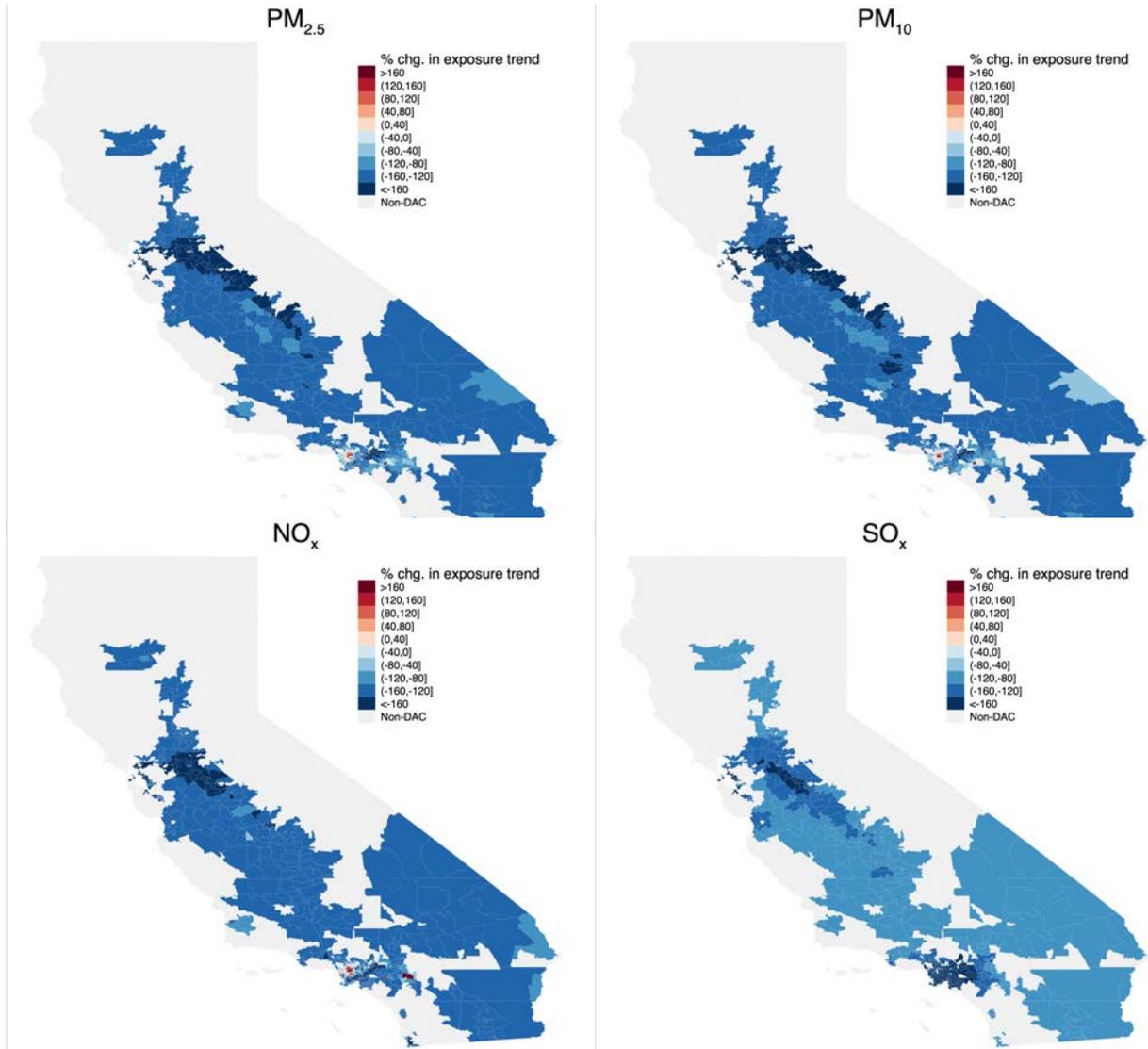
NOTES: Panels maps the zip code-specific percentage change in the EJ gap trend following the introduction of the C&T program for disadvantaged zip codes across $PM_{2.5}$, PM_{10}, NO_x, and SO_x. Blue (red) shading indicates reduced (increased) EJ gap trends following C&T for disadvantaged zip codes. Gray shading shows non-disadvantaged zip codes.

Spatial heterogeneity Estimates from equation (2) shown in Table 2 and Figure 3 examine the time evolution of EJ gaps averaged across disadvantaged and other zip codes. Additionally, one may be interested in how EJ gap effects vary spatially, particularly given the localized nature of EJ concerns. To examine spatial heterogeneity in trend break effects across disadvantaged zip codes, we estimate a variant of equation (2) allowing zip code-

specific post-C&T EJ gap trend break coefficients.[34] Figure 4 shows the percentage change in the EJ gap trend following the introduction of C&T for each disadvantaged zip code for $PM_{2.5}$, PM_{10}, NO_x, and SO_x. Across pollutants, post-C&T EJ gaps narrowed the most for disadvantaged zip codes in California's Central Valley. For $PM_{2.5}$, PM_{10}, and NO_x, Figure 4 also shows a cluster of zip codes in Los Angeles County that experienced widening post-C&T EJ gaps. Figure S6 shows histograms for the distribution of percentage changes in EJ gap trends across disadvantaged zip codes.

Robustness checks We subject our EJ gap trend effects to several robustness checks. Most robustness checks forgo the bootstrap procedure across steps 1-3 (detailed in Appendix A.1) given the computational demands of that procedure. Instead, Figure 5 presents only point estimates of the percentage change in the EJ gap trend following C&T (i.e., $\frac{\beta_2^p}{\beta_1^p} * 100$) for each robustness check and compares that with the point estimate and 95% confidence interval of our benchmark result for which inference does account for statistical uncertainty from equation (1) via our bootstrap procedure.[35]

Within step 1, we conduct eight EJ gap robustness checks, drawing on C&T emissions effects shown in Tables S2-S4. Equation (1) models changes in the emissions difference between C&T regulated and non-regulated facility as linear trends. We find a similar result when we estimate a more flexible version of equation (1) with year-specific emission differences (M2 of Figure 5 and column 1 of Table S7); when we replace year fixed effects with sector-by-year fixed effects in equation (1) (M3 of Figure 5 and column 2 of Table S7); and when we drop facilities that switched regulatory status in 2017 (M4 of Figure 5 and column 3 of Table S7). Next, we consider restricting facilities to those with sample average annual GHG emissions below the 70th and 80th percentiles, respectively (M5-6 of Figure 5 and columns 4-5 of Table S7). These alternative facility sample restrictions do not alter EJ gap trend effects.

We further allow the post C&T emissions trend break to vary as a linear function of sample average emissions. Recall from the heterogeneous emissions effects shown in column 2 of Table S3 that large-emitting facilities are abating more than is assumed in our baseline

---

[34]Specifically, we estimate the following variant of equation (2)

$$E_{it}^p = \beta_1^p [D_i \times t] + \sum_i \beta_{2i}^p [D_i \times \mathbf{1}(t \geq 2013) \times t] + \psi_i^p + \delta_t^p + \epsilon_{it}^p$$

where β_{2i}^p is the post-C&T trend break for zip code i. Figures 4 and S6 plot $\frac{\beta_{2i}^p}{\beta_1^p} * 100$, the percentage change in the EJ gap trend following the introduction of the C&T program for zip code i relative to the average pre-C&T EJ gap trend across disadvantaged zip codes.

[35]Coefficients β_1^p and β_2^p in accompanying Tables S7 and S8 cluster standard errors ϵ_{it}^p from equation (2) at the county-level but are not adjusted for statistical uncertainty in equation (1).

model which assumes a common percentage emissions change. For $PM_{2.5}$, PM_{10}, and NO_x, allowing for heterogeneity in emissions effects results in a slightly larger, though not statistically different, percentage change in the EJ gap trend (M7 of Figure 5 and column 6 of Table S7). That is, our baseline model without heterogeneous emissions effects is slightly understating the EJ gap fall as a consequence of C&T. For SO_x, this dimension of heterogeneity implies much larger drops in the post-C&T EJ gap trend. Lastly, we examine EJ gap effects after restricting the set of unregulated C&T facilities to those in counties under Clear Air Act nonattainment and those whose parent company only operates a single facility (M8-9 of Figure 5 and columns 7 and 8 of Table S7). SUTVA concerns do not alter EJ gap trend effects.

We conduct four robustness checks within step 2. We use pollution half-life parameters taken from the atmospheric chemistry literature because HYSPLIT does not model pollution decay over time. Our results are relatively stable to whether we allow for a 10% larger half-life parameter which implies a slower decay rate (M10 of Figure 5 and column 1 Table S8) or a 10% smaller half-life parameter which implies a faster decay rate (M11 of Figure 5 and column 2 of Table S8). Likewise our results are little affected when we lower the height of the planetary boundary layer to 0.5 km (M12 of Figure 5 and column 3 Table S8) or raise it to 2 km (M13 of Figure 5 and column 4 Table S8).

We conduct three robustness checks within step 3. The first set of checks consider alternative error structures for ϵ_{it}^p. We find that precision increases when we allow ϵ_{it}^p to be spatially correlated within a uniform kernel across a distance of 500 km distance (Conley, 1999), roughly the longitudinal width of California, and serially correlated across 5 years (Newey and West, 1987) (column 5 of Table S8). Likewise, precision increases when we allow for error terms to be correlated across the four local pollutants using a Seemingly Unrelated Regression (SUR) procedure (column 6 of Table S8). Equation (2) examines the EJ gap in daily pollution levels of $\mu g/m^3/day$, the concentration unit typically used for air pollution policy and by the public health literature. In Table S9, we detect a post-C&T EJ gap trend break after applying an inverse hyperbolic sine transformation to our outcome variable, showing C&T-driven concentrations in disadvantaged communities decreased as a percentage of concentration in other communities after 2013. Standard errors reported in Table S9 are adjusted for statistical uncertainty from equation (1) using our bootstrap procedure.

Finally, to examine the potential role of secondary $PM_{2.5}$, we replace HYSPLIT in step 2 of our procedure with InMAP, a reduced-complexity dispersal model based on output from WRF-Chem, which incorporates atmospheric chemistry in order to model total (i.e., primary and secondary) $PM_{2.5}$ concentrations from C&T-driven facility-level $PM_{2.5}$, NO_x,

25

Figure 5: Robustness checks for EJ gap effects
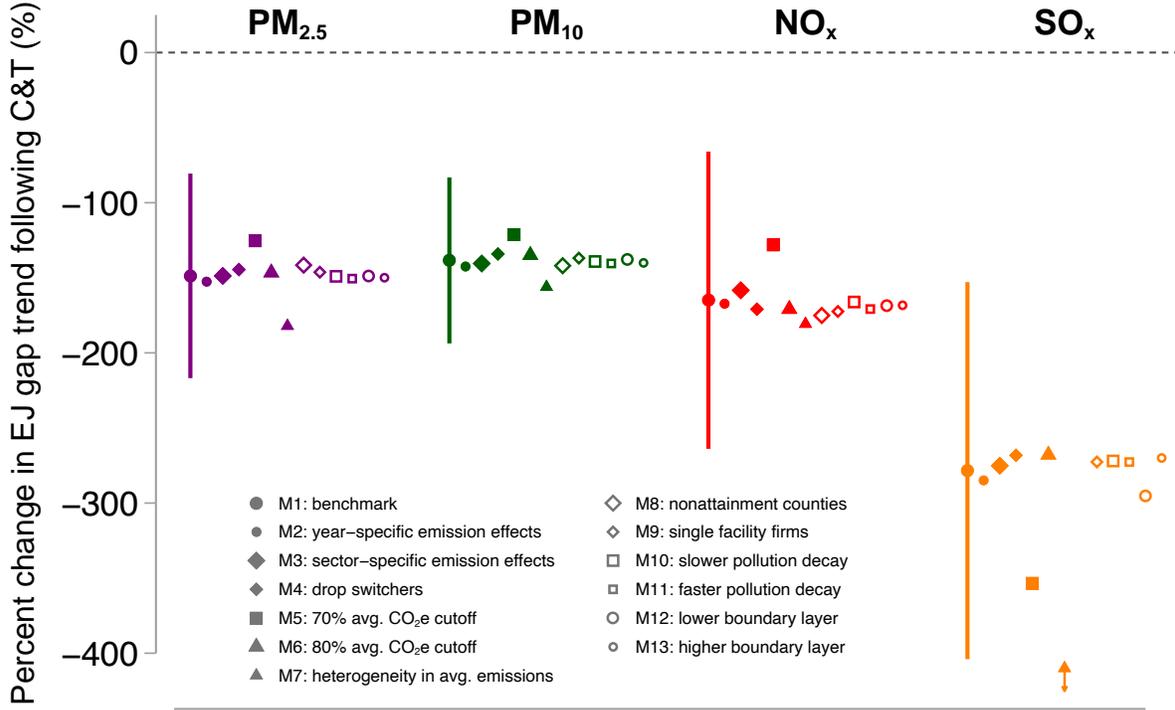
NOTES: Percentage change in the EJ gap trend following the introduction of the C&T program (i.e., $\frac{\beta_2^p}{\beta_1^p} * 100$) for $PM_{2.5}$, PM_{10}, NO_x, and SO_x across robustness checks. M1: benchmark model point estimate and 95% confidence interval accounting for uncertainty in equations (1) and (2). Point estimate shown for all other models. M2: using year-specific effects to estimate C&T-driven emissions. M3: C&T-driven emissions effects estimated using sector-by-year fixed effects. M4: C&T-driven emissions effects estimated without facilities that switched status in 2017. M5: restricting sample to facilities with average annual GHG emissions below the 70th percentile. M6: restricting sample to facilities with average annual GHG emissions below the 80th percentile. M7: allowing heterogeneous emissions effects by average annual emissions. M8: restricting unregulated facilities to those in counties under Clear Air Act nonattainment. M9: restricting unregulated facilities to those whose parent company only operates a single plant. M10: applying a slower pollution decay (i.e., 10% larger half-life parameter). M11: applying a faster pollution decay (i.e., 10% smaller half-life parameter). M12: applying a planetary boundary layer set at 0.5 km. M13: applying planetary boundary layer set at 2 km. Point estimates also reported in Tables S7-S8.

and SO_x, emissions (Tessum, Hill and Marshall, 2017).[36] InMAP, however, has one major limitation: it uses dispersal patterns in 2005, whereas our sample period is 2008-2017. Because InMAP does not model dispersal patterns during our sample period, we are unable to directly compare estimates using InMAP-generated concentrations with that using

---

[36]In addition to the inputs used in HYSPLIT, InMap requires the diameter, temperature, and emissions velocity for each smokestack. We obtain these inputs from CARB. In the case of facilities with more than one stack, we use the mean value across stacks. In the case of facilities with missing observations, we use the industry-level average.

HYSPLIT-generated concentrations.[37] Instead, we examine the role of secondary $PM_{2.5}$ by comparing how EJ gap trend estimates differ between InMAP-generated primary $PM_{2.5}$ concentration and InMAP-generated total $PM_{2.5}$ concentrations. If these two InMap estimates are similar, it is plausible that the true EJ gap effect for secondary $PM_{2.5}$ is similar to effects using HYSPLIT-generated primary $PM_{2.5}$ concentrations. Table S10 replicates the structure of Table 2. Column 1 examines InMAP-generated primary $PM_{2.5}$ concentrations while column 2 examines InMAP-generated total $PM_{2.5}$ concentrations. Both show similar EJ gap effects.

The importance of modeling pollution dispersal Our empirical approach explicitly embeds an atmospheric dispersal model within a causal inference framework. Compared with conventional methods for assigning pollution concentration from emission sources, this approach lends two benefits. It accounts for actual pollution dispersal patterns as dictated by topography and time-varying meteorological conditions. It also determines resulting pollution concentrations across all locations in California, rather than a subset of locations assumed to be exposed to policy-driven emissions. To demonstrate the importance of accounting for pollution dispersal for our results, we compare estimates from using our approach with that of more conventional methods of assigning pollution concentrations from emission sources.

Figure 6 plots estimates of the pre-C&T trend, or β_p^1 (left panel), and the post-C&T trend break, or β_p^2 (right panel), across criteria pollutants under different assumptions about how facility-level emissions alter location-specific concentrations.[38] In M1, we show our benchmark estimate where pollution dispersal is modeled by HYSPLIT every 4 hours throughout the 2008-2017 period. In M2 (and column 1 of Table S11), we assume that the area affected by a facility's emissions is limited to the zip code of that facility, referred to in the literature as "unit-hazard coincidence" (Banzhaf, Ma and Timmins, 2019). In M3-5 (and columns 2-4 of Table S11), we employ a distance-based measure by assuming that the area affected by a facility's emissions is limited to zip codes with centroids that are within 1.6, 4, and 10 km circles around the facility. These radial distances appear in the literature but nonetheless are chosen largely arbitrarily. Point estimates of β_p^1 and β_p^2 vary greatly across these alternative methods for assigning pollution concentrations. Not only do some estimates fall well outside the 95% confidence intervals of our benchmark results, but they also have different signs.

---

[37]Furthermore, there is a difference in units between HYSPLIT and InMap. For any given location, HYSPLIT produces the stock of pollution concentration during a given period, whereas InMAP produces that period's average flow of pollution concentration.

[38]Unlike Figure 5, Figure 6 does not plot $\frac{\beta_2^p}{\beta_1^p} * 100$ because β_1^p and β_2^p do not have consistent sign across the different methods for assigning emissions to concentrations.

Figure 6: Importance of modeling pollution dispersal
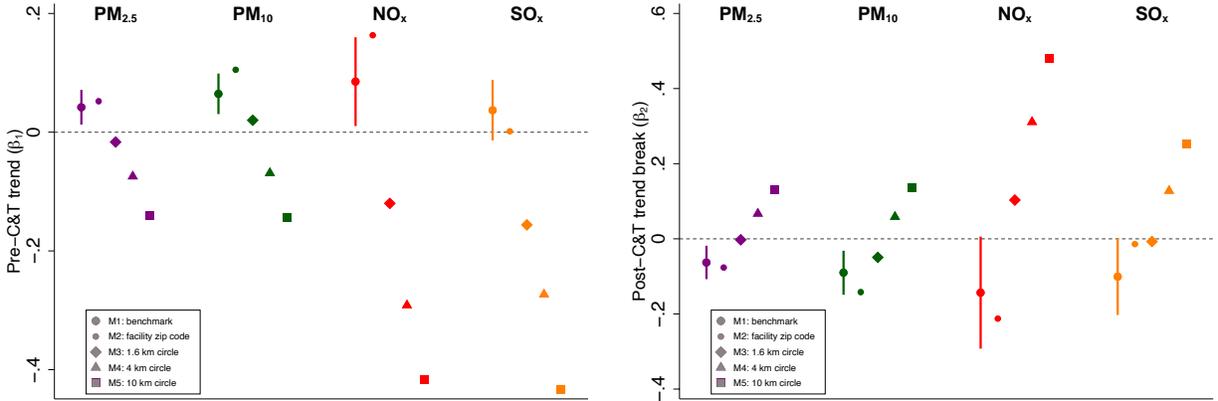
NOTES: Left panel shows estimates of pre-C&T trend (i.e., β_1^p) and right panel shows estimates of post-C&T trend break (i.e., β_2^p) for $PM_{2.5}$, PM_{10}, NO_x, and SO_x across different methods for assigning pollution concentrations from emissions. M1: benchmark model with point estimate and 95% confidence interval accounting for uncertainty in equations (1) and (2). Point estimate shown for all other models. M2: pollution concentration assigned only to zip code of emitting facility. M3-5: pollution concentration assigned to zip codes with centroid within 1.6 km, 4 km and 10 km circle of emitting facility, respectively. Point estimates also reported in Table S11.

6   Discussion

Many market settings are characterized by efficiency-equity tradeoffs. We find that California's carbon market led to an equity co-benefit by narrowing the criteria air pollution concentration gap between disadvantaged and other communities. This result brings causal evidence to a debate that continues to shape one of the world's most ambitious climate policies and climate policies elsewhere. Moreover, the integration of pollution dispersal modeling and causal inference employed in this paper may have broader applications across a variety of environmental valuation settings.

Equity concerns regarding California's cap-and-trade program remain. First, while we show that the program has led the pollution concentration gap between disadvantaged and other communities to fall, this gap has not been eliminated five years into the program. Second, pollution concentration constitutes only one component of the many distributional concerns regarding the program. Questions remain over how the program may have altered the distribution of health outcomes as well as the distribution of the program's cost burden. A comprehensive understanding of welfare inequality must also account for sorting as households move in response to changes in pollution concentrations (Depro, Timmins and O'Neil, 2015; Banzhaf, Ma and Timmins, 2019) and for entry decisions by polluters (Weber,

2020). Third, a broader notion of equity must also consider the ability of disadvantaged communities to partake in decision-making around environmental policies. Such procedural justice issues remain in California though recent policies such as AB 617 are beginning to engage disadvantaged communities directly in the design of local pollution regulations (Fowlie, Walker and Wooley, 2020).

More generally, despite these findings for California, market-based environmental policies should not be used explicitly to address environmental justice concerns. Market-based policies are intended for allocative efficiency and not distributional objectives, per se. The EJ gap consequences detected in California emerges from the state's spatial distribution of polluting facilities and demographic characteristics. In other settings where facilities with steeper marginal abatement cost curves are upwind of disadvantaged communities, an environmental market could widen the environmental justice gap. Difficulties with observing facility-level marginal abatement cost curves make it hard to anticipate ex-ante how proposed market-based policies will alter existing EJ gaps. As a safeguard against potential widening EJ gaps, policies that specifically address environmental justice concerns should be considered in tandem with market-based policies. In short, environmental justice problems need environmental justice policies.

References

Ash, Michael, and T. Robert Fetter. 2004. "Who Lives on the Wrong Side of the Environmental Tracks? Evidence from the EPA's Risk-Screening Environmental Indicators Model." *Social Science Quarterly*, 85(2): 441–462.

Banzhaf, Spencer, Lala Ma, and Christopher Timmins. 2019. "Environmental Justice: The Economics of Race, Place, and Pollution." *Journal of Economic Perspectives*, 33(1): 185–208.

Baumol, William J., and Wallace E. Oates. 1988. *The Theory of Environmental Policy, Second Edition.* Cambridge University Press.

Bellemare, Marc F, and Casey J Wichman. 2020. "Elasticities and the inverse hyperbolic sine transformation." *Oxford Bulletin of Economics and Statistics*, 82(1): 50–61.

Board, California Air Resources. 2020. "California Greenhouse Gas Emissions for 2000 to 2018: Trends of Emissions and Other Indicators ."

Borenstein, Severin, James Bushnell, Frank A. Wolak, and Matthew Zaragoza-Watkins. 2019.

"Expecting the Unexpected: Emissions Uncertainty and Environmental Market Design." *American Economic Review*, 109(11): 3953–77.

Bowen, William. 2002. "An Analytical Review of Environmental Justice Research: What do we Really Know?" *Environmental Management*, 29(1): 3–15.

Bullard, Robert. 2000. *Dumping in Dixie: Race, Class, and Environmental Quality.* Westview Press.

Burtraw, Dallas, A Carlson, D Cullenward, Q Foster, and M Fowlie. 2018. "2018 Annual Report of the Independent Emissions Market Advisory Committee."

Burtraw, Dallas, David A Evans, Alan Krupnick, Karen Palmer, and Russell Toth. 2005. "Economics of Pollution Trading for SO2 and NO x." *Annu. Rev. Environ. Resour.*, 30: 253–289.

Casey, Joan A, Jason G Su, Lucas RF Henneman, Corwin Zigler, Andreas M Neophytou, Ralph Catalano, Rahul Gondalia, Yu-Ting Chen, Leanne Kaye, Sarah S Moyer, et al. 2020. "Improved asthma outcomes observed in the vicinity of coal power plant retirement, retrofit and conversion to natural gas." *Nature Energy*, 5(5): 398–408.

Colmer, Jonathan, Ralf Martin, Mirabelle Muûls, Ulrich J Wagner, et al. 2020. "Does pricing carbon mitigate climate change? Firm-level evidence from the European Union emissions trading scheme." *Center for Economic Performance Discussion Paper*, , (1728).

Conley, Timothy G. 1999. "GMM Estimation with Cross Sectional Dependence." *Journal of Econometrics*, 92(1): 1–45.

Costello, Christopher, Daniel Ovando, Tyler Clavelle, C. Kent Strauss, Ray Hilborn, Michael C. Melnychuk, Trevor A. Branch, Steven D. Gaines, Cody S. Szuwalski, Reniel B. Cabral, Douglas N. Rader, and Amanda Leland. 2016. "Global Fishery Prospects under Contrasting Management Regimes." *Proceedings of the National Academy of Sciences*, 113(18): 5125–5129.

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

Cummiskey, Kevin, Chanmin Kim, Christine Choirat, Lucas R. F. Henneman, Joel Schwartz, and Corwin Zigler. 2019. "A Source-Oriented Approach to Coal Power Plant Emissions Health Effects."

Currie, Janet, John Voorheis, and Reed Walker. 2020. "What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence from the Clean Air Act and Satellite-Based Measures of Air Quality." National Bureau of Economic Research Working Paper 26659.

Cushing, Lara, Dan Blaustein-Rejto, Madeline Wander, Manuel Pastor, James Sadd, Allen Zhu, and Rachel Morello-Frosch. 2018. "Carbon Trading, Co-pollutants, and Environmental Equity: Evidence from California's cap-and-trade program (2011–2015)." *PLoS medicine*, 15(7): e1002604.

Dale, John H. 1968. "Pollution, Property, and Prices: An Essay in Policy-Making."

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

Davis, Lucas, and Catherine Hausman. 2016. "Market Impacts of a Nuclear Power Plant Closure." *American Economic Journal: Applied Economics*, 8(2): 92–122.

Depro, Brooks, Christopher Timmins, and Maggie O'Neil. 2015. "White Flight and Coming to the Nuisance: Can Residential Mobility Explain Environmental Injustice?" *Journal of the Association of Environmental and Resource Economists*, 2(3): 439–468.

Deschenes, Olivier, and Kyle C Meng. 2018. "Quasi-experimental Methods in Environmental Economics: Challenges and Opportunities." *Handbook of Environmental Economics*, 4: 285.

Draxler, Roland R, and GD Hess. 1998. "An Overview of the $HYSPLIT_4$ Modelling System for Trajectories." *Australian Meteorological Magazine*, 47(4): 295–308.

EPA, US. 2015. "Revision to the Guideline on Air Quality Models: Enhancements to the AERMOD Dispersion Modeling System and Incorporation of Approaches to Address Ozone and Fine Particulate Matter." US Environmental Protection Agency Working Paper 2060-AS54.

Fowlie, Meredith, Reed Walker, and David Wooley. 2020. "Climate policy, environmental justice, and local air pollution." Brookings Economic Studies.

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

Graff Zivin, Joshua, and Matthew Neidell. 2013. "Environment, health, and human capital." *Journal of Economic Literature*, 51(3): 689–730.

Grainger, Corbett, and Thanicha Ruangmas. 2018. "Who Wins from Emissions Trading? Evidence from California." *Environmental and Resource Economics*, 71(3): 703–727.

Greenstone, Michael, and Ted Gayer. 2009. "Quasi-experimental and Experimental Approaches to Environmental Economics." *Journal of Environmental Economics and Management*, 57(1): 21 – 44. Frontiers of Environmental and Resource Economics.

Henneman, Lucas RF, Christine Choirat, and Corwin M Zigler. 2019. "Accountability assessment of health improvements in the United States associated with reduced coal emissions between 2005 and 2012." *Epidemiology*, 30(4): 477–485.

Henneman, Lucas R.F., Christine Choirat, Cesunica E. Ivey, Kevin Cummiskey, and Corwin M. Zigler. 2019. "Characterizing population exposure to coal emissions sources in the United States using the HyADS model." *Atmospheric Environment*, 203(203): 271–280.

Henneman, Lucas RF, Loretta J Mickley, and Corwin M Zigler. 2019. "Air pollution accountability of energy transitions: the relative importance of point source emissions and wind fields in exposure changes." *Environmental Research Letters*, 14(11): 115003.

Herron, Elise. 2019. "Oregon Clean Energy Jobs Bill Has Mixed Support From Environmental Justice Groups." *Willamette Week*.

Kim, Chanmin, Lucas RF Henneman, Christine Choirat, and Corwin M Zigler. 2020. "Health effects of power plant emissions through ambient air quality." *Journal of the Royal Statistical Society: Series A (Statistics in Society)*.

Leber, Rebecca. 2016. "The most Dramatic Climate Fight of the Election is in Washington State." *Grist*.

Lee, Chulkyu, Randall V Martin, Aaron van Donkelaar, Hanlim Lee, Russell R Dickerson, Jennifer C Hains, Nickolay Krotkov, Andreas Richter, Konstantine Vinnikov, and James J Schwab. 2011. "SO_2 Emissions and Lifetimes: Estimates from Inverse Modeling using in situ and global, space-based (SCIAMACHY and OMI) Observations." *Journal of Geophysical Research: Atmospheres*, 116(D6).

Liu, Fei, Steffen Beirle, Qiang Zhang, Steffen Dörner, Kebin He, and Thomas Wagner. 2016. "NO_x Lifetimes and Emissions of Cities and Power Plants in Polluted Background Estimated by Satellite Observations." *Atmospheric Chemistry and Physics*, 16(8): 5283–5298.

Mansur, Erin T., and Glenn Sheriff. 2019. "Do Pollution Markets Harm Low Income and Minority Communities? Ranking Emissions Distributions Generated by California's RE-CLAIM Program." National Bureau of Economic Research Working Paper 25666.

Martin, Ralf, Mirabelle Muûls, and Ulrich J Wagner. 2016. "The impact of the European Union Emissions Trading Scheme on regulated firms: what is the evidence after ten years?" *Review of environmental economics and policy*, 10(1): 129–148.

Mohai, Paul, David Pellow, and J. Timmons Roberts. 2009. "Environmental Justice." *Annual Review of Environment and Resources*, 34(1): 405–430.

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

Morello-Frosch, Rachel, and Bill M Jesdale. 2006. "Separate and Unequal: Residential Segregation and Estimated Cancer Risks Associated with Ambient Air Toxics in US Metropolitan Areas." *Environmental Health Perspectives*, 114(3): 386.

Newey, Whitney K., and Kenneth D. West. 1987. "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55(3): 703–708.

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

Rahn, David A., and Christopher J. Mitchell. 2016. "Diurnal Climatology of the Boundary Layer in Southern California Using AMDAR Temperature and Wind Profiles." *Journal of Applied Meteorology and Climatology*, 55(5): 1123–1137.

Ringquist, Evan J. 2005. "Assessing Evidence of Environmental Inequities: A Meta-analysis." *Journal of Policy Analysis and Management*, 24(2): 223–247.

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

Schmalensee, Richard, and Robert N. Stavins. 2019. "Policy Evolution under the Clean Air Act." *Journal of Economic Perspectives*, 33(4): 27–50.

Shapiro, Joseph S, and Reed Walker. 2021. "Where is Pollution Moving? Environmental Markets and Environmental Justice." National Bureau of Economic Research Working Paper 28389.

Sullivan, Daniel M. 2017. "The True Cost of Air Pollution: Evidence from the Housing Market." *mimeo.*

Tessum, Christopher W., Jason D. Hill, and Julian D. Marshall. 2017. "InMAP: A model for air pollution interventions." *PLOS ONE*, 12(4): 1–26.

Tessum, Christopher W., Joshua S. Apte, Andrew L. Goodkind, Nicholas Z. Muller, Kimberley A. Mullins, David A. Paolella, Stephen Polasky, Nathaniel P. Springer, Sumil K. Thakrar, Julian D. Marshall, and Jason D. Hill. 2019. "Inequity in consumption of goods and services adds to racial–ethnic disparities in air pollution exposure." *Proceedings of the National Academy of Sciences*, 116(13): 6001–6006.

Transnational Institute. 2013. "It is Time to Scrap the ETS! ."

U.S. EPA. 2018. "User's Guide for the AMS/EPA Regulatory Model AERMOD."

Weber, Paige. 2020. "Dynamic Response to Carbon pricing in the Electricity Sector." mimeo.

World Bank Group. 2019. "State and Trends of Carbon Pricing."

A Appendix

A.1 Bootstrap procedure for incorporating uncertainty in C&T emission effects

This section details our bootstrap procedure over steps 1-3 to account for statistical uncertainty in C&T-driven emission effects from equation (1), reproduced here:

$$asinh(Y_{jt}^p) = \kappa_1^p[C_j \times t] + \kappa_2^p[C_j \times \mathbf{1}(t \geq 2013) \times t] + \phi_j^p + \gamma_t^p + \nu_{jt}^p$$

We obtain point estimates $\widehat{\kappa}_1^p$, $\widehat{\kappa}_2^p$ and standard errors $\widehat{\sigma}_{\kappa_1^p}$ and $\widehat{\sigma}_{\kappa_2^p}$ from equation (1). We then iterate the following procedure for draws $b = 1...250$:

1. Draw $\widehat{\kappa}_1^p(b) \sim N(\widehat{\kappa}_1^p, \widehat{\sigma}_{\kappa_1^p})$ and $\widehat{\kappa}_2^p(b) \sim N(\widehat{\kappa}_2^p, \widehat{\sigma}_{\kappa_2^p})$

2. Construct $\widehat{Y}_{jt}^p(b) = sinh\left(\widehat{\kappa}_1^p(b)[C_j \times t] + \widehat{\kappa}_2^p(b)[C_j \times \mathbf{1}(t \geq 2013) \times t] + \widehat{\phi}_j^p\right) * e^{(RMSE)^2/2}$, where RMSE is the root mean squared error from equation (1)

3. Feed $\widehat{Y}_{jt}^p(b)$ into HYSPLIT to generate zip code-by-year pollution concentration, $E_{it}^p(b)$

4. Estimate equation (2) using $E_{it}^p(b)$ as the outcome variable to obtain $\widehat{\beta}_1^p(b)$ and $\widehat{\beta}_2^p(b)$

Figure S5 plots the empirical distributions for $\widehat{\beta}_1^p(b)$ and $\widehat{\beta}_2^p(b)$ for $p \in \{PM_{2.5}, PM_{10}, NO_x, SO_x\}$. Denote standard errors across 250 bootstrap runs as $\widehat{\sigma}_{\beta_1^p}(\nu_{jt}^p)$ and $\widehat{\sigma}_{\beta_2^p}(\nu_{jt}^p)$ where the ν_{jt}^p argument indicates the dependence on statistical uncertainty from equation (1). Denote $\widehat{\sigma}_{\beta_1^p}(\epsilon_{jt}^p)$ as the estimated standard error arising from heterogeneity in β_1^p obtained by directly estimating equation (2) with county-level clustered errors. Our reported standard error for β_1^p is $\widehat{\sigma}_{\beta_1^p} = \widehat{\sigma}_{\beta_1^p}(\epsilon_{jt}^p) + \widehat{\sigma}_{\beta_1^p}(\nu_{jt}^p)$. Likewise, for β_2^p. $\widehat{\sigma}_{\beta_1^p}$ and $\widehat{\sigma}_{\beta_2^p}$ are reported in Table 2 and used to construct the confidence intervals displayed in Figure 3.

A.2 Can the EJ gap effect be recovered using dispersal-augmented facility-level regressions?

Facility-level analyses examine how a policy's emissions effects vary with characteristics of locations that are downwind of facilities, as determined by an atmospheric dispersal model (Grainger and Ruangmas, 2018; Mansur and Sheriff, 2019). For example, one may estimate the following dispersal-augmented facility-level regression for the change in facility j emissions before and after the introduction of a policy:

$$\Delta Y_j = \phi_0 C_j + \phi_1 C_j s_j + \phi_2 s_j + \varepsilon_j \tag{A.1}$$

where $C_j \in \{0,1\}$ is regulatory status and $s_j \in [0,1]$ is the share of affected downwind locations that is disadvantaged, as determined by the dispersal model. In these models, the coefficient of interest is $\widehat{\phi}_1$, the added emissions effect for facilities that disproportionately affect disadvantaged communities.

How does $\widehat{\phi}_1$ relate to the EJ gap effect, the estimand of interest? As in Section 4, let i index locations (e.g., California zip codes in our setting) and $D_i \in \{0,1\}$ denote disadvantaged status. For simplicity, assume there are the same number of disadvantaged and non-disadvantaged locations, $N = \sum_{i:D_i=0}(1 - D_i) = \sum_{i:D_i=1} D_i$. The EJ gap effect is the difference between the change in pollution concentration for disadvantaged communities and that of non-disadvantaged communities, due to C&T-driven emission changes from regulated facilities (relative to unregulated facilities), or $\widehat{\Delta Y}_j$. Formally, this estimand is

$$
\begin{aligned}
\theta = &\underbrace{\frac{1}{N}\sum_j \widehat{\Delta Y}_j s_j}_{\substack{\text{avg. DAC concentration} \\ \text{from C\&T-driven} \\ \text{emissions}}} - \underbrace{\frac{1}{N}\sum_j \widehat{\Delta Y}_j (1 - s_j)}_{\substack{\text{avg. non-DAC concentration} \\ \text{from C\&T-driven} \\ \text{emissions}}} \\
= &\frac{1}{N}\sum_j \widehat{\Delta Y}_j (2s_j - 1) \\
= &\frac{1}{N}\sum_j (\widehat{\phi}_0 + \widehat{\phi}_1 s_j)(2s_j - 1)
\end{aligned}
\tag{A.2}
$$

The last line applies C&T-driven relative emissions change for regulated facilities as $\widehat{\Delta Y}_j = \widehat{\phi}_0 + \widehat{\phi}_1 s_j$, where $\widehat{\phi}_0$ and $\widehat{\phi}_1$ are estimated coefficients from equation (A.1).

It is clear from (A.2) that $\widehat{\phi}_1$ does not generally equal θ. But is the sign of $\widehat{\phi}_1$ at least consistently the same as the sign of θ? The following example rejects this claim. For simplicity, let $\widehat{\phi}_0 = 0$. Next, suppose $s_1 = 1$ for the first facility and $s_{j>1} < 0.5$ for all other

facilities. The estimand is then:

$$\theta = \frac{1}{N}\left(\widehat{\phi}_1 + \sum_{j>1}\widehat{\phi}_1 s_j(2s_j - 1)\right)$$

$$= \frac{\widehat{\phi}_1}{N}\left(\underbrace{1 + \sum_{j>1}\underbrace{s_j(2s_j - 1)}_{<0}}_{\gtreqless 0}\right)$$

indicating that θ and $\widehat{\phi}_1$ can be of different signs. Thus, simply showing, for example, that emissions are relatively higher for facilities that disproportionately affect disadvantaged communities (i.e., $\widehat{\phi}_1 > 0$) does not imply that the EJ gap has widened (i.e., $\theta > 0$).

Lastly, can $\widehat{\phi}_1$ ever equal θ? Returning to equation (A.1), we explore one such special case. Assume that exactly N facilities only affect disadvantaged communities (i.e., $s_j = 0$), and that another N facilities only affect non-disadvantaged communities (i.e., $s_j = 1$). Then the estimand becomes:

$$\theta = \frac{1}{N}\left(\sum_{j:s_j=0}-\widehat{\phi}_0 + \sum_{j:s_j=1}(\widehat{\phi}_0 + \widehat{\phi}_1)\right)$$

$$= \frac{1}{N}(-N\widehat{\phi}_0 + N\widehat{\phi}_0 + N\widehat{\phi}_1)$$

$$= \widehat{\phi}_1$$

Observe how restrictive the assumptions are for this special case. It requires that facilities *only* affect disadvantaged communities or *only* affect non-disadvantaged communities, that is $s_j \subset \{0, 1\} \ \forall j$. Facilities cannot alter pollution concentrations in both types of locations. Furthermore, this case requires that the number of facilities in both groups equal the number of disadvantaged communities, which must also equal the number of non-disadvantaged communities.

Appendix Figures

Figure S1: Distribution of sample facility-year emissions
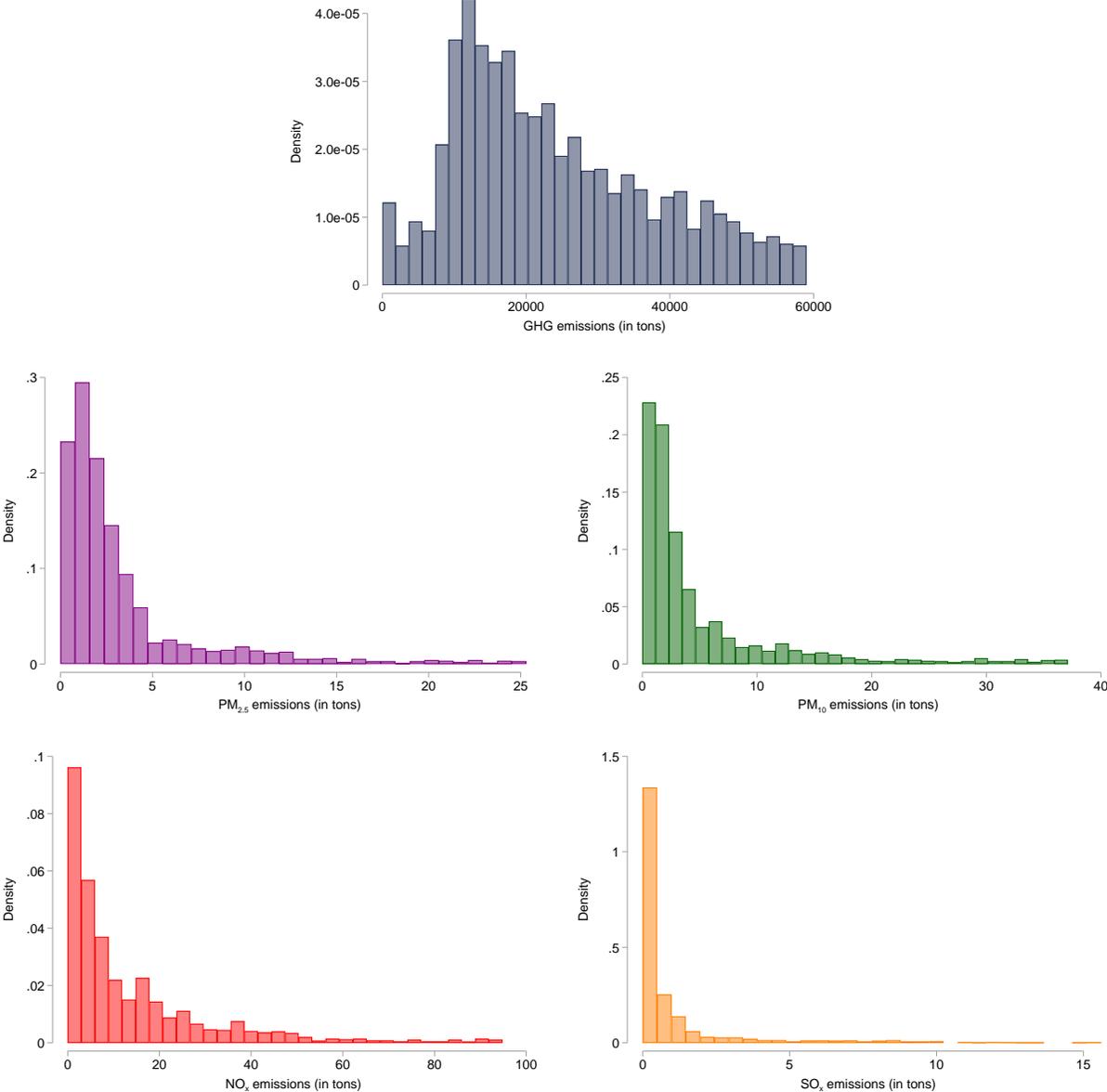
NOTES: Panels show the distribution of facility-year GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x emissions for sample observations. Observations above the 95th percentile are truncated.

Figure S2: Emissions robustness: placebo C&T program timing
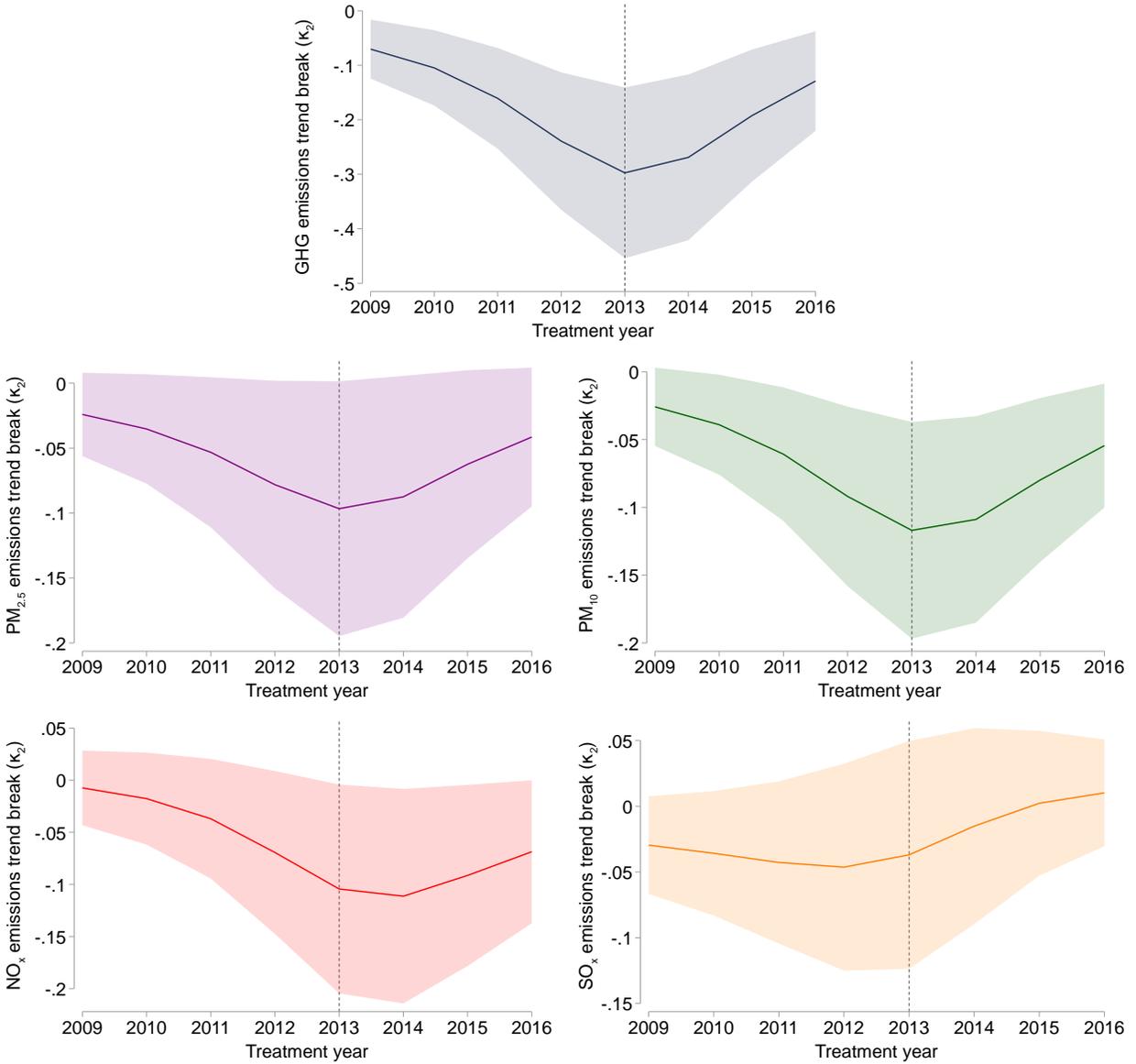
NOTES: Panels show estimated (true and placebo) emissions trend break coefficients (i.e., κ_2 from eq. (1)) for GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x emissions from varying the start year of the C&T program. Vertical line at 2013 indicates actual introduction of the program. Shaded areas indicate 95% confidence intervals.

Figure S3: Facility-level C&T-driven abatement between 2012-2017
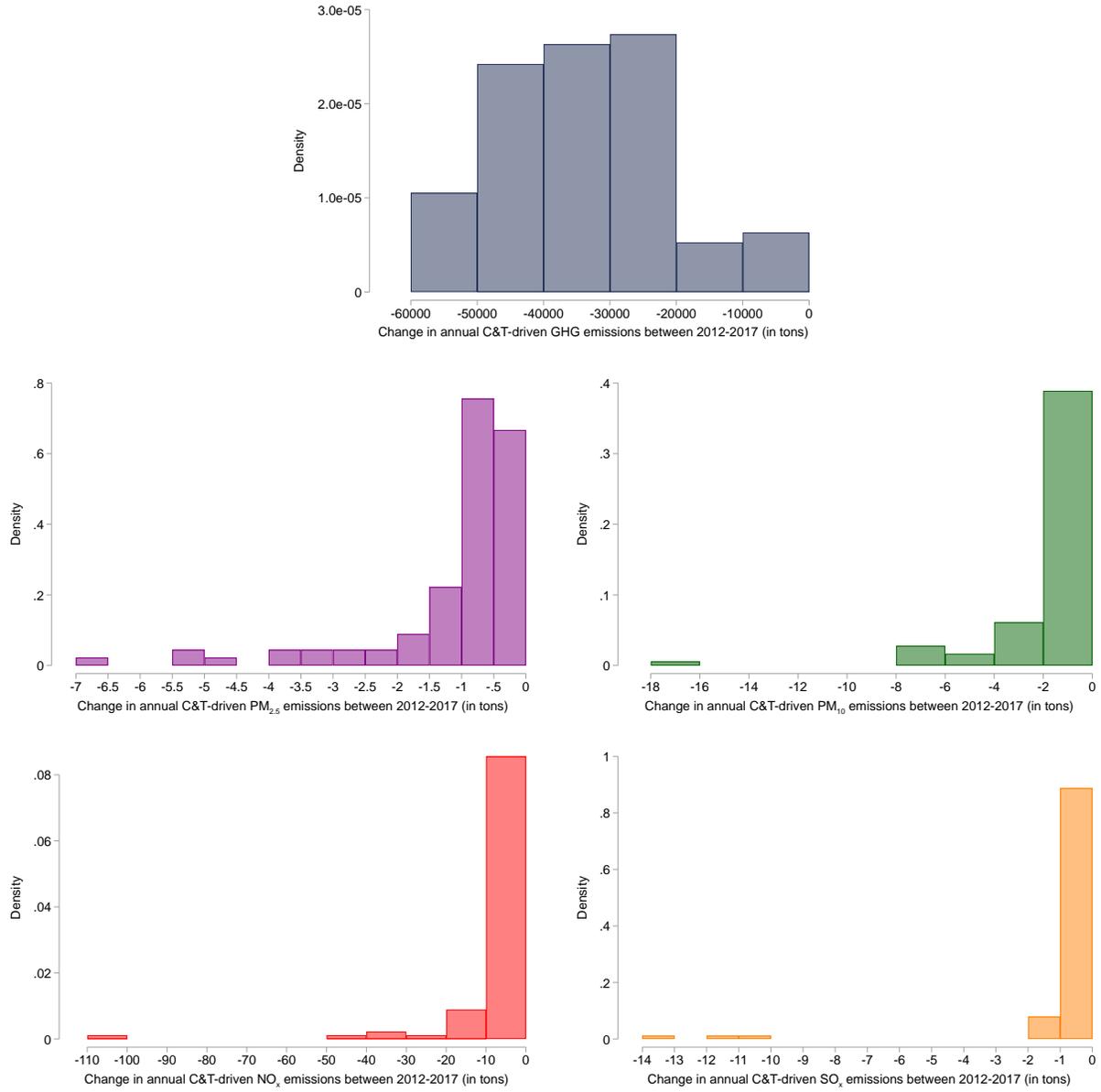
NOTES: Panels show the distribution of facility-level change in C&T-driven pollution between 2012-2017 (or abatement) predicted from step 1 for GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x emissions, respectively.

Figure S4: Average pollution concentrations driven by C&T regulated facilities
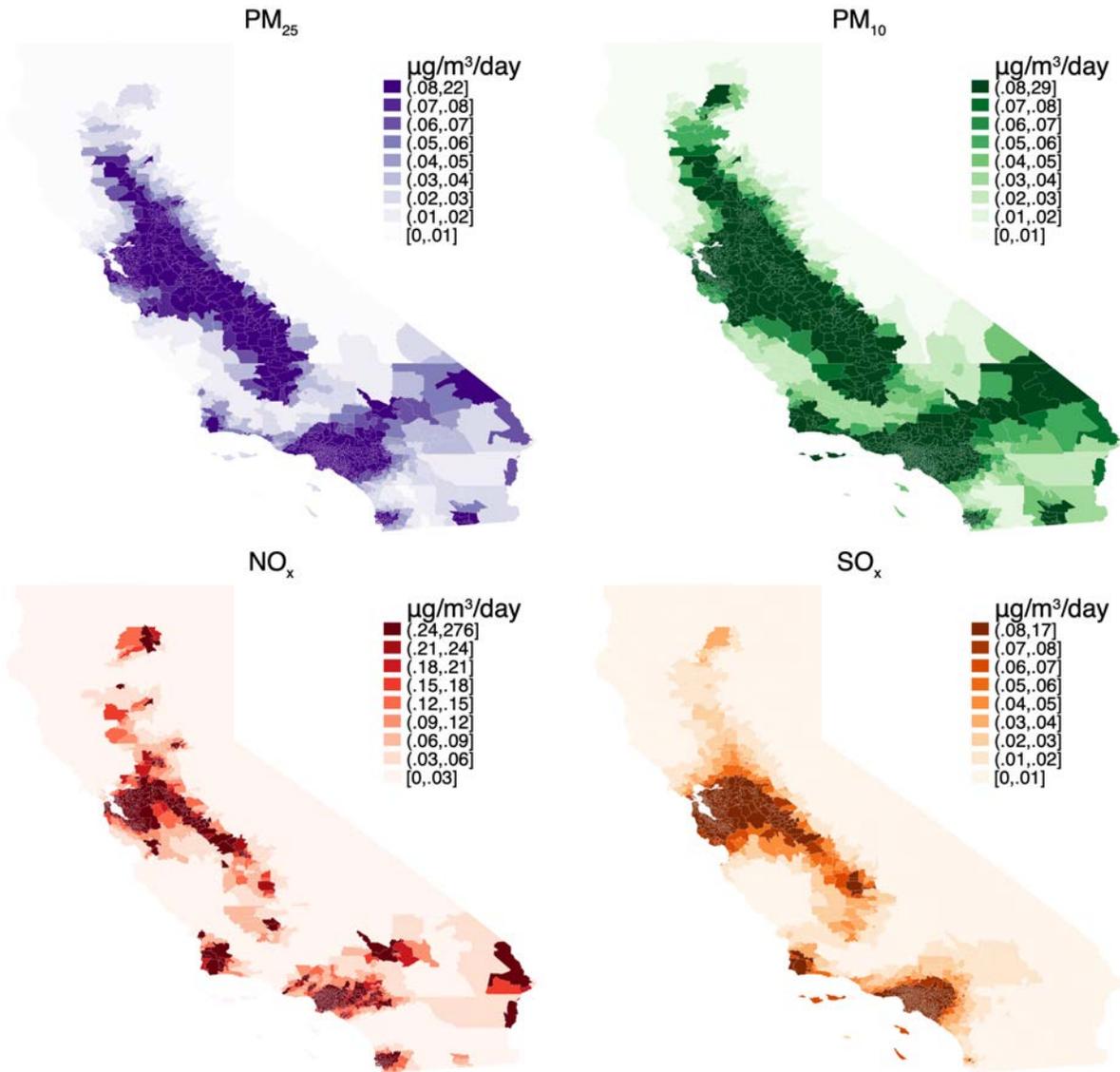
NOTES: Panels show daily concentrations (in $\mu g/m^3/day$) for each zip code averaged across 2008-2017 from GHG C&T-regulated facilities as modeled in step 2 by HYSPLIT for $PM_{2.5}$, PM_{10}, NO_x, and SO_x, respectively.

Figure S5: Empirical distribution of β_1^p and β_2^p from bootstrapping step 1
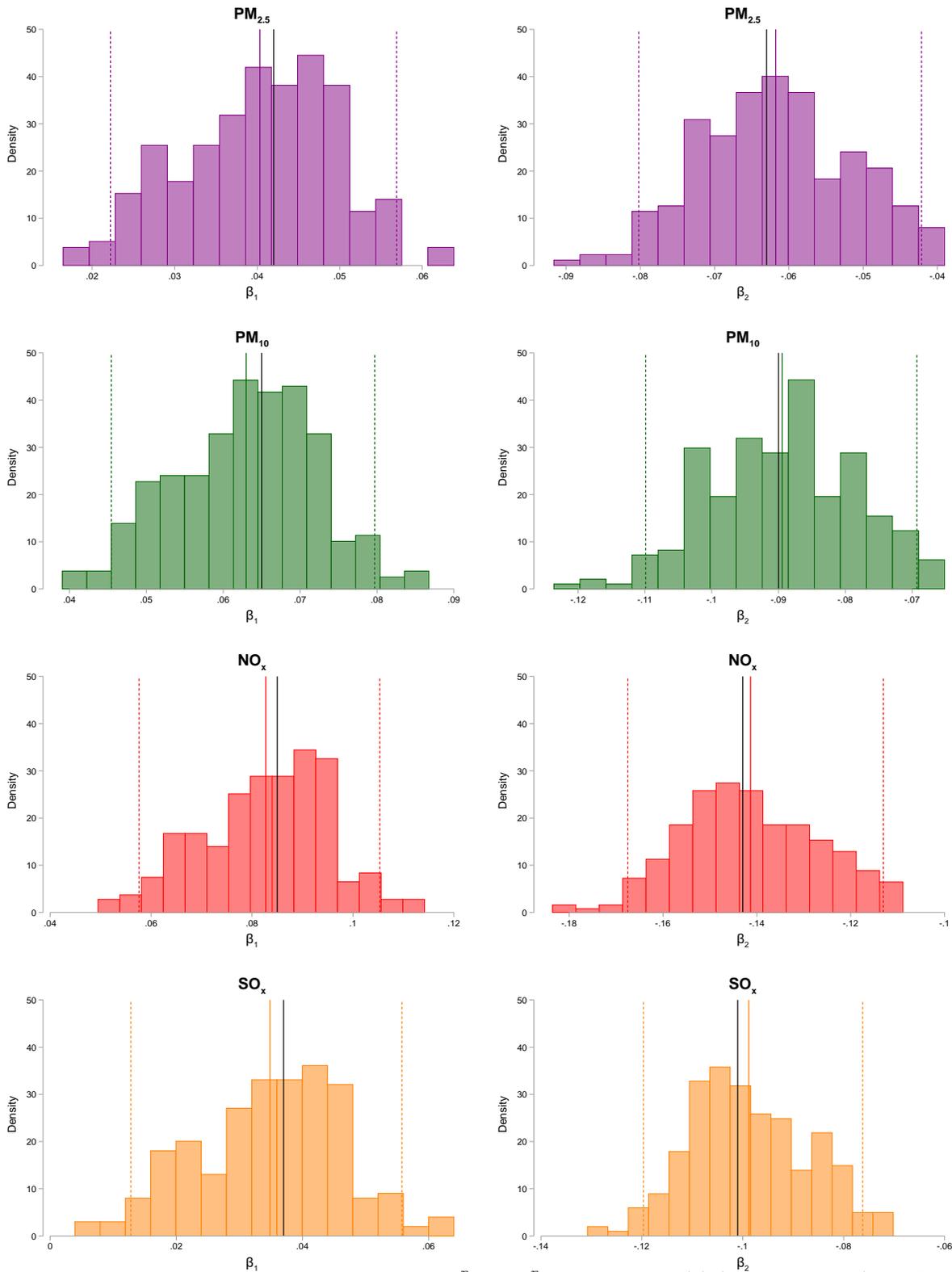
NOTES: Panels show the empirical distribution of β_1^p and β_2^p from equation (1) (across columns) for $PM_{2.5}$, PM_{10}, NO_x, and SO_x (across rows) using the bootstrap procedure detailed in Section A.1 with 250 draws. Solid black line shows parameter from directly estimating equation (1). Solid colored line shows the mean parameter value from the empirical bootstrapped distribution. Dotted colored lines show the 2.5% and 97.5% percentiles of the empirical bootstrap distributions.

Figure S6: Zip code-level percent change in EJ gap trend following C&T
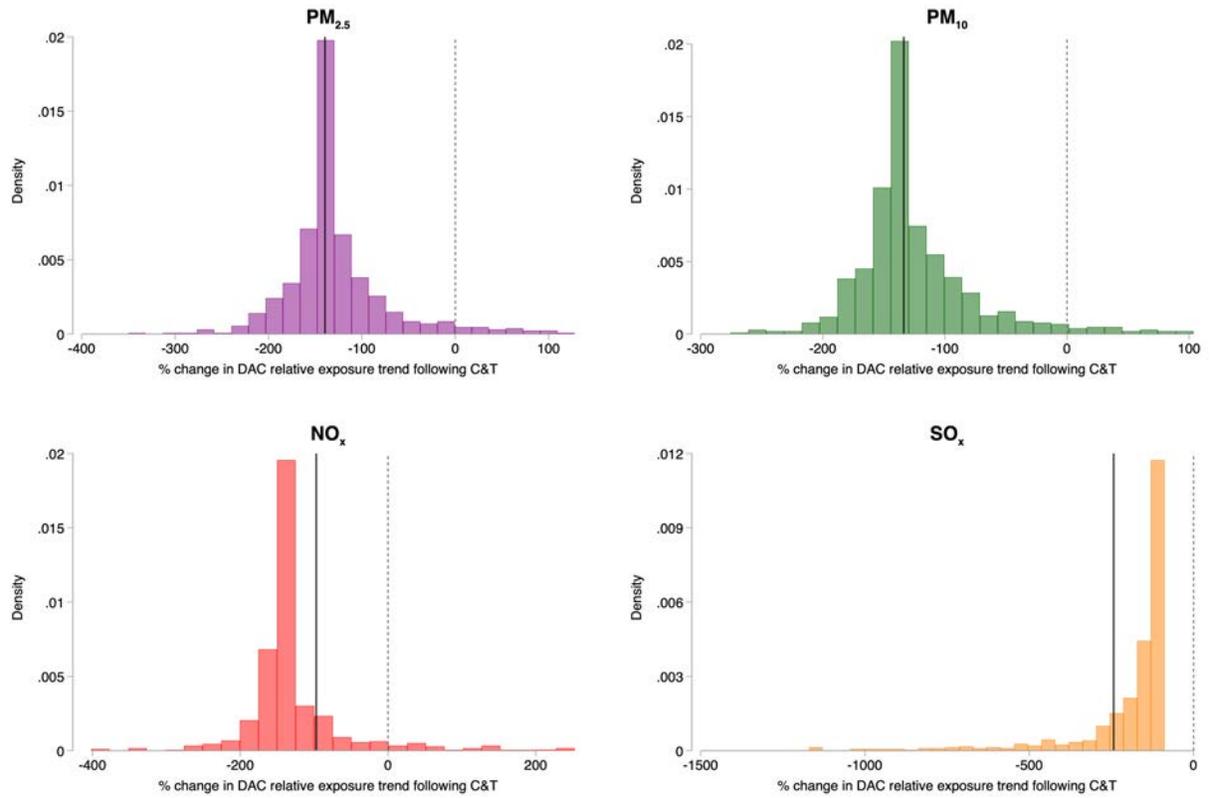
NOTES: Panels show the distribution of zip code-level percentage change in the EJ gap trend following the introduction of the C&T program, for each disadvantaged zip code across $PM_{2.5}$, PM_{10}, NO_x, and SO_x. Solid line shows the average percentage change across disadvantaged zip codes, or $\frac{\beta_2^p}{\beta_1^p} * 100$ from equation (2). Dashed line marks zero.

Appendix Tables

Table S1: GHG cap-and-trade regulated and non-regulated facilities

	C&T regulated facilities	non-C&T regulated facilities
Number of facilities	106	226
Avg. 2008-2012 emissions (in metric tons):		
CO_2	38192.62	17566.48
$PM_{2.5}$	8.08	3.74
PM_{10}	14.47	6.25
NO_x	53.42	16.03
SO_x	10.86	2.8
Shares by sector:		
Agriculture	0	.018
Manufacturing	.623	.5
Mining, oil and gas extraction	.151	.097
Services	.066	.23
Transportation	.075	.053
Utilities	.075	.093
Wholesale	.009	.009

NOTES: Sample C&T regulated and non-regulated facilities by count, average 2008-2012 GHG and criteria air pollution emissions, and by sector shares. Sectors shown adhere to the following definitions: Agriculture: NAICS 11; Manufacturing: NAICS 31-33; Mining, oil, and gas extraction: NAICS 21; Services: NAICS 51, 54, 56, 61, 62, 71, 81, 92; Transportation: NAICS 48, 49; Utilities: NAICS 22; Wholesale: NAICS 42.

Table S2: Emissions robustness: specification and sample restrictions

	(1) sector-year FEs	(2) Drop switchers	(3) Baseline CO_2e cutoff (%) 70	(4) Baseline CO_2e cutoff (%) 80
	Outcome is $asinh(CO_2e)$ emissions			
κ_1^p	0.172	0.220	0.194	0.174
	(0.058)	(0.054)	(0.055)	(0.050)
	[0.005]	[0.000]	[0.001]	[0.001]
κ_2^p	-0.273	-0.317	-0.307	-0.260
	(0.095)	(0.079)	(0.085)	(0.072)
	[0.006]	[0.000]	[0.001]	[0.001]
Facilities	315	298	294	337
Observations	2,052	1,924	1,863	2,234
	Outcome is $asinh(PM_{2.5})$ emissions			
κ_1^p	0.059	0.065	0.071	0.046
	(0.043)	(0.045)	(0.043)	(0.043)
	[0.176]	[0.156]	[0.111]	[0.298]
κ_2^p	-0.099	-0.092	-0.105	-0.079
	(0.048)	(0.051)	(0.050)	(0.050)
	[0.046]	[0.081]	[0.044]	[0.121]
Facilities	301	285	281	323
Observations	1,966	1,847	1,780	2,147
	Outcome is $asinh(PM_{10})$ emissions			
κ_1^p	0.086	0.088	0.097	0.075
	(0.034)	(0.035)	(0.034)	(0.035)
	[0.014]	[0.016]	[0.008]	[0.039]
κ_2^p	-0.124	-0.108	-0.129	-0.103
	(0.040)	(0.043)	(0.043)	(0.042)
	[0.003]	[0.017]	[0.005]	[0.018]
Facilities	301	285	281	323
Observations	1,966	1,847	1,780	2,147
	Outcome is $asinh(NO_x)$ emissions			
κ_1^p	0.085	0.057	0.085	0.058
	(0.042)	(0.039)	(0.033)	(0.037)
	[0.048]	[0.158]	[0.015]	[0.128]
κ_2^p	-0.117	-0.079	-0.126	-0.091
	(0.053)	(0.050)	(0.047)	(0.048)
	[0.035]	[0.123]	[0.010]	[0.066]
Facilities	302	286	282	324
Observations	1,968	1,849	1,782	2,149
	Outcome is $asinh(SO_x)$ emissions			
κ_1^p	0.005	0.008	-0.005	-0.004
	(0.038)	(0.036)	(0.038)	(0.035)
	[0.902]	[0.817]	[0.890]	[0.912]
κ_2^p	-0.040	-0.040	-0.025	-0.020
	(0.048)	(0.046)	(0.048)	(0.045)
	[0.407]	[0.386]	[0.600]	[0.657]
Facilities	302	286	282	324
Observations	1,961	1,847	1,777	2,142

NOTES: Estimates of pre-C&T differential emissions trend (i.e., κ_1^p from equation (1)) and and post-C&T differential emissions trend break (i.e., κ_2^p from equation (1)) for GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x across panels. All models include facility-specific and year-specific dummy variables (except for column 1). Column 1 replaces year fixed effects with sector-by-year fixed effects with sectors defined as shown in Table S1. Column 2 drops facilities that switched C&T regulatory status in 2017. Columns 3 and 4 restrict facilities to those with sample average annual GHG emissions below the 70th and 80th percentile, respectively. Standard errors clustered at the county-level in parentheses, p-value in brackets.

Table S3: Emissions effect robustness: heterogeneity by average emissions

	(1)	(2)	(3)
	Outcome is asinh(GHG) emissions		
κ_1^p	0.187	0.176	0.172
	(0.052)	(0.052)	(0.052)
	[0.001]	[0.002]	[0.002]
κ_2^p	-0.297	-0.361	-0.354
	(0.077)	(0.092)	(0.097)
	[0.000]	[0.000]	[0.001]
trend break $\times$ avg. emissions		0.000	0.000
		(0.000)	(0.000)
		[0.053]	[0.090]
trend break $\times$ avg. emissions^2			-0.000
			(0.000)
			[0.158]
	Outcome is asinh(PM$_{2.5}$) emissions		
κ_1^p	0.058	0.060	0.060
	(0.043)	(0.042)	(0.043)
	[0.183]	[0.167]	[0.165]
κ_2^p	-0.097	-0.133	-0.146
	(0.048)	(0.051)	(0.068)
	[0.053]	[0.012]	[0.040]
trend break $\times$ avg. emissions		-0.004	-0.005
		(0.003)	(0.004)
		[0.197]	[0.249]
trend break $\times$ avg. emissions^2			0.000
			(0.000)
			[0.661]
	Outcome is asinh(PM$_{10}$) emissions		
κ_1^p	0.083	0.084	0.086
	(0.033)	(0.033)	(0.033)
	[0.016]	[0.015]	[0.012]
κ_2^p	-0.117	-0.143	-0.172
	(0.039)	(0.042)	(0.048)
	[0.005]	[0.002]	[0.001]
trend break $\times$ avg. emissions		-0.002	-0.003
		(0.001)	(0.002)
		[0.080]	[0.073]
trend break $\times$ avg. emissions^2			0.000
			(0.000)
			[0.197]
	Outcome is asinh(NO$_x$) emissions		
κ_1^p	0.075	0.079	0.080
	(0.039)	(0.038)	(0.039)
	[0.061]	[0.046]	[0.046]
κ_2^p	-0.104	-0.143	-0.157
	(0.050)	(0.045)	(0.079)
	[0.042]	[0.003]	[0.054]
trend break $\times$ avg. emissions		-0.001	-0.001
		(0.000)	(0.001)
		[0.002]	[0.294]
trend break $\times$ avg. emissions^2			0.000
			(0.000)
			[0.793]
	Outcome is asinh(SO$_x$) emissions		
κ_1^p	0.006	0.013	0.013
	(0.035)	(0.035)	(0.035)
	[0.875]	[0.715]	[0.705]
κ_2^p	-0.037	-0.110	-0.074
	(0.043)	(0.048)	(0.077)
	[0.393]	[0.026]	[0.345]
trend break $\times$ avg. emissions		-0.004	-0.002
		(0.002)	(0.003)
		[0.017]	[0.455]
trend break $\times$ avg. emissions^2			-0.000
			(0.000)
			[0.438]

NOTES: Estimates of pre-C&T differential emissions trend (i.e., κ_1^p from equation (1)) and and post-C&T differential emissions trend break (i.e., κ_2^p from equation (1)) for GHG, PM$_{2.5}$, PM$_{10}$, NO$_x$, and SO$_x$ across panels. Columns 1 shows benchmark model. Column 2 (3) further interacts post C&T differential trend break with a linear (quadratic) function of sample average annual emissions. All models include facility-specific and year-specific dummy variables. Standard errors clustered at the county-level in parentheses, p-value in brackets.

Table S4: Emissions effect robustness: restricting treatment spillovers

	(1) Benchmark	(2) Nonattainment	(3) Single facilities
	Outcome is asinh(GHG) emissions		
κ_1^p	0.187	-	0.210
	(0.052)	-	(0.053)
	[0.001]	-	[0.000]
κ_2^p	-0.297	-	-0.322
	(0.077)	-	(0.078)
	[0.000]	-	[0.000]
Facilities	316	-	310
Observations	2,054	-	2,029
	Outcome is asinh(PM$_{2.5}$) emissions		
κ_1^p	0.058	0.085	0.066
	(0.043)	(0.049)	(0.043)
	[0.183]	[0.092]	[0.137]
κ_2^p	-0.097	-0.119	-0.101
	(0.048)	(0.052)	(0.049)
	[0.053]	[0.029]	[0.046]
Facilities	302	260	299
Observations	1,968	1,729	1,952
	Outcome is asinh(PM$_{10}$) emissions		
κ_1^p	0.083	0.101	0.091
	(0.033)	(0.034)	(0.033)
	[0.016]	[0.006]	[0.008]
κ_2^p	-0.117	-0.145	-0.121
	(0.039)	(0.054)	(0.040)
	[0.005]	[0.012]	[0.004]
Facilities	302	140	299
Observations	1,968	1,080	1,952
	Outcome is asinh(NO$_x$) emissions		
κ_1^p	0.075	0.057	0.065
	(0.039)	(0.041)	(0.039)
	[0.061]	[0.173]	[0.101]
κ_2^p	-0.104	-0.090	-0.098
	(0.050)	(0.054)	(0.050)
	[0.042]	[0.102]	[0.060]
Facilities	303	287	300
Observations	1,970	1,879	1,954
	Outcome is asinh(SO$_x$) emissions		
κ_1^p	0.006	-	0.005
	(0.035)	-	(0.036)
	[0.875]	-	[0.892]
κ_2^p	-0.037	-	-0.036
	(0.043)	-	(0.044)
	[0.393]	-	[0.423]
Facilities	303	-	300
Observations	1,965	-	1,950

NOTES: Estimates of pre-C&T differential emissions trend (i.e., κ_1^p from equation (1)) and and post-C&T differential emissions trend break (i.e., κ_2^p from equation (1)) for GHG, PM$_{2.5}$, PM$_{10}$, NO$_x$, and SO$_x$ across panels. Columns 1 shows benchmark model. Column 2 restricts unregulated facilities to those in counties under Clear Air Act nonattainment for pollutant of interest. Nonattainment does not apply for GHG emissions and there were no counties under SO$_x$ nonattainment during our sample period. For NO$_x$, we use nonattainment in the one-hour ozone standard, for which NO$_x$ is a precursor pollutant. Column 3 restricts unregulated facilities to those whose parent company only operates a single facility. All models include facility-specific and year-specific dummy variables. Standard errors clustered at the county-level in parentheses, p-value in brackets.

Table S5: Correlation between HYSPLIT-driven and ambient pollution concentrations

	(1)	(2)	(3)	(4)
	Outcome is ambient asinh(concentration)			
	$PM_{2.5}$	PM_{10}	NO_x	SO_x
HYSPLIT-driven asinh(concentration)	0.860	0.625	0.436	0.231
	(0.154)	(0.137)	(0.148)	(0.207)
	[0.000]	[0.000]	[0.004]	[0.272]
Zip codes	133	160	121	39

NOTES: Linear coefficient from zip code-level regressions of asinh daily HYSPLIT-driven pollution concentrations (in $\mu g/m^3$/day) averaged across 2008-2017 on asinh daily pollution concentrations from ambient pollution monitors (in $\mu g/m^3$/day) averaged across 2008-2017. We employ a asinh-asinh specification because ambient pollution readings, which capture the average daily instantaneous stock of pollution, are not directly comparable to our concentration measure, which capture average daily pollution flow from C&T-driven emissions. Ambient pollution are assumed to be uniformly distributed within a monitor's zip code. Standard errors clustered at the county-level in parentheses, p-value in brackets.

Table S6: Pollution concentration difference between disadvantaged and other zip codes in 2008

	(1) Disadvantaged	(2) Other	(3) Difference
$PM_{2.5}$	0.256 (0.888)	0.093 (0.572)	0.163 (0.038) [0.000]
PM_{10}	0.322 (1.066)	0.109 (0.532)	0.214 (0.043) [0.000]
NO_x	0.451 (2.842)	0.387 (6.856)	0.064 (0.243) [0.792]
SO_x	0.364 (1.092)	0.091 (0.217)	0.273 (0.041) [0.000]
Zip codes	722	984	1,706

NOTES: Column 1 shows average 2008 pollution concentration (μg/m^3/day) across disadvantaged zip codes, with standard deviation in parentheses. Column 2 shows average 2008 pollution concentration (μg/m^3/day) across other zip codes, with standard deviation in parentheses. Column 3 shows the average difference in 2008 pollution concentrations between disadvantaged and other zip codes, with standard error in parentheses and p-value in brackets. All pollution concentrations generated by HYS-PLIT from facilities that would eventually be regulated by the GHG C&T program.

Table S7: EJ gap effect robustness: step 1

	(1) Year-specific effects	(2) Sector-year FEs	(3) Drop switchers	(4) GHG cutoff: 70%	(5) GHG cutoff: 80%	(6) Hetero by avg. emissions	(7) SUTVA NA	(8) SUTVA Single fac.
				Panel a: $PM_{2.5}$				
β_1^p	0.040	0.041	0.040	0.025	0.043	0.041	0.048	0.043
	(0.011)	(0.011)	(0.013)	(0.006)	(0.010)	(0.012)	(0.012)	(0.012)
	[0.001]	[0.001]	[0.003]	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]
β_2^p	-0.061	-0.061	-0.058	-0.031	-0.063	-0.075	-0.067	-0.064
	(0.019)	(0.019)	(0.021)	(0.008)	(0.019)	(0.021)	(0.021)	(0.020)
	[0.003]	[0.002]	[0.007]	[0.000]	[0.001]	[0.001]	[0.002]	[0.002]
$(\beta_2^p/\beta_1^p)*100$	-152.583	-148.723	-144.528	-125.385	-146.581	-182.096	-141.543	-146.262
Observations	16,416	16,416	16,413	16,387	16,426	16,416	16,416	16,416
				Panel b: PM_{10}				
β_1^p	0.062	0.063	0.059	0.038	0.069	0.064	0.074	0.066
	(0.014)	(0.014)	(0.017)	(0.008)	(0.013)	(0.014)	(0.016)	(0.015)
	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
β_2^p	-0.089	-0.089	-0.079	-0.046	-0.093	-0.100	-0.105	-0.091
	(0.027)	(0.026)	(0.028)	(0.009)	(0.026)	(0.028)	(0.030)	(0.028)
	[0.002]	[0.001]	[0.007]	[0.000]	[0.001]	[0.001]	[0.001]	[0.002]
$(\beta_2^p/\beta_1^p)*100$	-142.447	-140.455	-134.110	-121.310	-134.948	-156.155	-141.932	-136.905
Observations	16,416	16,416	16,413	16,387	16,426	16,416	16,416	16,416
				Panel c: NO_x				
β_1^p	0.079	0.091	0.077	0.043	0.087	0.079	0.079	0.082
	(0.033)	(0.037)	(0.032)	(0.026)	(0.031)	(0.033)	(0.032)	(0.034)
	[0.019]	[0.015]	[0.021]	[0.108]	[0.006]	[0.021]	[0.018]	[0.018]
β_2^p	-0.132	-0.145	-0.132	-0.055	-0.149	-0.142	-0.138	-0.141
	(0.066)	(0.070)	(0.069)	(0.031)	(0.069)	(0.075)	(0.070)	(0.071)
	[0.051]	[0.043]	[0.061]	[0.084]	[0.035]	[0.062]	[0.052]	[0.053]
$(\beta_2^p/\beta_1^p)*100$	-167.212	-158.315	-170.919	-128.157	-170.878	-180.760	-175.096	-172.408
Observations	16,416	16,416	16,413	16,387	16,426	16,416	16,416	16,416
				Panel d: SO_x				
β_1^p	0.036	0.036	0.038	0.023	0.037	0.011	-	0.037
	(0.022)	(0.022)	(0.023)	(0.015)	(0.020)	(0.012)	-	(0.023)
	[0.108]	[0.109]	[0.104]	[0.141]	[0.077]	[0.349]	-	[0.108]
β_2^p	-0.103	-0.099	-0.102	-0.080	-0.099	-0.114	-	-0.100
	(0.050)	(0.049)	(0.050)	(0.045)	(0.046)	(0.058)	-	(0.049)
	[0.045]	[0.047]	[0.046]	[0.084]	[0.037]	[0.054]	-	[0.047]
$(\beta_2^p/\beta_1^p)*100$	-284.826	-275.129	-268.202	-353.380	-267.896	-1003.707	-	-272.630
Observations	16,416	16,416	16,413	16,387	16,426	16,416	-	16,416

NOTES: Estimates of the pre-C&T EJ gap trend (i.e., β_1^p from equation (2)), post-C&T EJ gap trend break (i.e., β_2^p from equation (2)), and the percentage change in the EJ gap trend following the introduction of the C&T program (i.e., $\frac{\beta_2^p}{\beta_1^p}*100$) for $PM_{2.5}$, PM_{10}, NO_x, and SO_x down panels. All models include zip code-specific and year-specific dummy variables. Observations weighted by zip code-level average population during 2008-2012. Column 1 uses year-specific effects to estimate C&T-driven emissions. Column 2 estimates C&T-driven emissions with sector-by-year fixed effects (see column 1 of Table S2). Column 3 estimates C&T-driven emissions after dropping facilities that switched regulatory status in 2017 (see column 2 of Table S2). Columns 4 and 5 restrict facilities to those with sample average GHG emissions below the 70th and 80th percentile, respectively to estimate C&T-driven emissions (see columns 3 and 5 of Table S2). Column 6 uses C&T-driven emissions that allow the C&T differential trend break to vary as a linear function of sample average emissions (see column 2 of Table S3). Column 7 restricts unregulated facilities to those in counties under Clear Air Act nonattainment for pollutant of interest (see column 2 of Table S4). Column 8 restricts unregulated facilities to those whose parent company only operates a single facility (see column 3 of Table S4). Standard errors, in parentheses, cluster ϵ_{it}^p from equation (2) at the county-level but are not adjusted for statistical uncertainty from equation (1). P-value in brackets. Observations apply to all panels.

Table S8: EJ gap effect robustness: steps 2 and 3

	(1) Slower decay	(2) Faster decay	(3) Lower boundary	(4) Higher boundary	(5) Spatial corr. err.	(6) Pollution corr. err.
			Panel a: $PM_{2.5}$			
β_1^p	0.043	0.041	0.037	0.043	0.042	0.042
	(0.011)	(0.011)	(0.010)	(0.011)	(0.004)	(0.006)
	[0.000]	[0.000]	[0.001]	[0.000]	[0.000]	[0.000]
β_2^p	-0.064	-0.062	-0.055	-0.064	-0.063	-0.063
	(0.020)	(0.020)	(0.019)	(0.020)	(0.009)	(0.010)
	[0.002]	[0.003]	[0.007]	[0.002]	[0.000]	[0.000]
$(\beta_2^p/\beta_1^p)*100$	-149.007	-150.533	-148.764	-149.992	-149.699	-149.699
			Panel b: PM_{10}			
β_1^p	0.066	0.063	0.057	0.066	0.065	0.065
	(0.015)	(0.014)	(0.013)	(0.014)	(0.006)	(0.008)
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
β_2^p	-0.092	-0.089	-0.079	-0.092	-0.090	-0.090
	(0.027)	(0.027)	(0.027)	(0.027)	(0.011)	(0.013)
	[0.001]	[0.002]	[0.005]	[0.001]	[0.000]	[0.000]
$(\beta_2^p/\beta_1^p)*100$	-139.150	-140.448	-137.785	-140.161	-139.739	-139.739
			Panel c: NO_x			
β_1^p	0.089	0.081	0.083	0.085	0.085	0.085
	(0.036)	(0.034)	(0.035)	(0.035)	(0.039)	(0.021)
	[0.018]	[0.020]	[0.020]	[0.018]	[0.030]	[0.000]
β_2^p	-0.148	-0.139	-0.140	-0.144	-0.143	-0.143
	(0.073)	(0.073)	(0.073)	(0.073)	(0.050)	(0.033)
	[0.048]	[0.063]	[0.060]	[0.054]	[0.004]	[0.000]
$(\beta_2^p/\beta_1^p)*100$	-166.117	-170.804	-168.674	-168.261	-168.282	-168.282
			Panel d: SO_x			
β_1^p	0.037	0.037	0.030	0.038	0.037	0.037
	(0.023)	(0.022)	(0.019)	(0.023)	(0.007)	(0.006)
	[0.109]	[0.107]	[0.133]	[0.103]	[0.000]	[0.000]
β_2^p	-0.102	-0.100	-0.087	-0.102	-0.101	-0.101
	(0.050)	(0.049)	(0.044)	(0.050)	(0.012)	(0.010)
	[0.047]	[0.047]	[0.053]	[0.045]	[0.000]	[0.000]
$(\beta_2^p/\beta_1^p)*100$	-271.967	-272.688	-295.166	-270.107	-272.291	-272.291
Observations	16,416	16,416	16,359	16,430	16,417	16,417

NOTES: Estimates of the pre-C&T EJ gap trend (i.e., β_1^p from equation (2)), post-C&T EJ gap trend break (i.e., β_2^p from equation (2)), and the percentage change in the EJ gap trend following the introduction of the C&T program (i.e., $\frac{\beta_2^p}{\beta_1^p}*100$) for $PM_{2.5}$, PM_{10}, NO_x, and SO_x down panels. All models include zip code-specific and year-specific dummy variables. Observations weighted by zip code-level average population during 2008-2012. Column 1 applies a slower pollution decay to HYSPLIT pollution trajectories (i.e., 10% larger half-life parameter). Column 2 applies a faster pollution decay to HYSPLIT pollution trajectories (i.e., 10% smaller half-life parameter). Column 3 applies a lower planetary boundary layer set at 0.5 km. Column 4 applies a higher planetary boundary layer set at 2 km. Column 5 adjusts standard errors for spatial (500 km uniform kernel) and serial correlation (5 years). Column 6 adjusts standard errors allowing correlation across pollutants using a Seemingly Unrelated Regression (SUR) procedure. Standard errors, in parentheses, cluster ϵ_{it}^p from equation (2) at the county-level but are not adjusted for statistical uncertainty from equation (1). P-value in brackets. Observations apply to all panels.

Table S9: EJ gap effect robustness: asinh concentration

	(1)	(2)	(3)	(4)
	Outcome is (asinh) concentration			
	$PM_{2.5}$	PM_{10}	NO_x	SO_x
β_1^p	0.027	0.037	0.032	0.017
	(0.013)	(0.014)	(0.021)	(0.017)
	[0.045]	[0.009]	[0.137]	[0.336]
β_2^p	-0.032	-0.042	-0.038	-0.051
	(0.014)	(0.015)	(0.023)	(0.030)
	[0.026]	[0.009]	[0.102]	[0.095]
$\beta_1^p + \beta_2^p$	-0.006	-0.004	-0.005	-0.034
	(0.005)	(0.007)	(0.008)	(0.015)
	[0.302]	[0.551]	[0.487]	[0.029]
Zip codes	1649	1649	1649	1649
Counties	58	58	58	58
Observations	16,416	16,416	16,416	16,416

NOTES: Estimates of the pre-C&T EJ gap trend (i.e., β_1^p from equation (2)), the post-C&T EJ gap trend break (i.e., β_2^p from equation (2)), and the post-C&T EJ gap trend (i.e, $\beta_1^p + \beta_2^p$) for asinh($PM_{2.5}$), asinh(PM_{10}), asinh(NO_x), and asinh(SO_x), across columns. All models include zip code-specific and year-specific dummy variables. Observations weighted by zip code-level average population during 2008-2012. Parentheses indicate standard errors that account for statistical uncertainty in C&T predicted emissions (ν_{it}^p from equation (1) via the bootstrap procedure in Appendix A.1) and county-level heterogeneity in EJ gap effects of arbitrary form (ϵ_{it}^p from equation (2)). P-value in brackets.

Table S10: EJ gap effect robustness: total PM$_{2.5}$ concentration using InMAP

	(1) Primary PM$_{2.5}$	(2) Total PM$_{2.5}$
β_1^p	0.002	0.003
	(0.001)	(0.001)
	[0.001]	[0.001]
β_2^p	-0.003	-0.004
	(0.001)	(0.001)
	[0.000]	[0.000]
$\beta_1^p + \beta_2^p$	-0.001	-0.002
	(0.000)	(0.001)
	[0.004]	[0.001]
$(\beta_2^p/\beta_1^p) * 100$	-150.559	-172.948
	(16.261)	(16.415)
	[0.000]	[0.000]
Zip codes	1648	1648
Counties	58	58
Observations	16,480	16,480

NOTES: Estimates of the pre-C&T EJ gap trend (i.e., β_1^p from equation (2)), the post-C&T EJ gap trend break (i.e., β_2^p from equation (2)), the post-C&T EJ gap trend (i.e, $\beta_1^p + \beta_2^p$), and the percentage change in the EJ gap trend following the introduction of the C&T program (i.e., $\frac{\beta_2^p}{\beta_1^p} * 100$) for InMAP-modeled primary PM$_{2.5}$ concentration (column 1) and for InMAP-modeled total (i.e., primary and secondary) PM$_{2.5}$ concentration (column 2). InMAP employs dispersal patterns for 2005 and not for the 2008-2017 sample period. All models include zip code-specific and year-specific dummy variables. Observations weighted by zip code-level average population during 2008-2012. Standard errors, in parentheses, cluster ϵ_{it}^p from equation (2) at the county-level but are not adjusted for statistical uncertainty from equation (1). P-value in brackets.

Table S11: Importance of modeling pollution dispersal

	(1) Facility zip code	(2) 1.6 km circle	(3) 4 km circle	(4) 10 km circle
	Panel a: PM$_{2.5}$			
β_1^p	0.052	-0.017	-0.075	-0.140
	(0.036)	(0.026)	(0.040)	(0.079)
	[0.157]	[0.527]	[0.075]	[0.084]
β_2^p	-0.076	-0.003	0.067	0.132
	(0.049)	(0.023)	(0.036)	(0.072)
	[0.134]	[0.912]	[0.075]	[0.076]
Observations	785	1,831	3,573	7,545
	Panel b: PM$_{10}$			
β_1^p	0.105	0.020	-0.069	-0.143
	(0.070)	(0.030)	(0.047)	(0.089)
	[0.142]	[0.509]	[0.155]	[0.116]
β_2^p	-0.142	-0.049	0.059	0.137
	(0.091)	(0.036)	(0.055)	(0.095)
	[0.132]	[0.177]	[0.294]	[0.157]
Observations	785	1,831	3,573	7,545
	Panel c: NO$_x$			
β_1^p	0.163	-0.120	-0.292	-0.417
	(0.188)	(0.110)	(0.096)	(0.175)
	[0.391]	[0.285]	[0.005]	[0.022]
β_2^p	-0.213	0.103	0.311	0.480
	(0.247)	(0.132)	(0.110)	(0.179)
	[0.396]	[0.442]	[0.008]	[0.011]
Observations	785	1,831	3,573	7,545
	Panel d: SO$_x$			
β_1^p	0.001	-0.156	-0.273	-0.433
	(0.004)	(0.122)	(0.183)	(0.250)
	[0.688]	[0.210]	[0.145]	[0.091]
β_2^p	-0.014	-0.007	0.128	0.253
	(0.009)	(0.030)	(0.103)	(0.143)
	[0.125]	[0.813]	[0.223]	[0.085]
Observations	783	1,823	3,553	7,535

NOTES: Estimates of the pre-C&T EJ gap trend (i.e., β_1^p from equation (2)) and the post-C&T EJ gap trend break (i.e., β_2^p from equation (2)) for PM$_{2.5}$, PM$_{10}$, NO$_x$, and SO$_x$ down panels. All models include zip code-specific and year-specific dummy variables. Observations weighted by zip code-level average population during 2008-2012. Column 1 assigns pollution concentration to only the zip code of the emitting facility. Columns 2-4 assign pollution concentration to zip codes with centroid within a 1.6, 4 km and 10 km circle of emitting facility, respectively. Standard errors in parentheses cluster ϵ_{it}^p from equation (2) at the county-level but are not adjusted for statistical uncertainty from equation (1). P-value in brackets.