Do Environmental Markets Cause Environmental Injustice? Evidence from California's Carbon Market

Danae Hernandez-Cortes and Kyle C. Meng*

November 2022

Abstract

Market-based environmental policies are widely adopted on the basis of 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 examine how this "environmental justice" (EJ) gap changed following the 2013 introduction of California's carbon market, the world's second largest and the most subjected to EJ critiques. We estimate that the program lowered GHG, $PM_{2.5}$, PM_{10}, and NO_x emissions by 3-9% annually between 2012-2017 for sample industrial facilities regulated only by the carbon market. Using a pollution dispersal model to characterize resulting spatial changes in pollution concentrations, we find the program caused EJ gaps in $PM_{2.5}$, PM_{10}, and NO_x from these facilities to narrow by 6-10% annually. We demonstrate that explicit modeling of pollution dispersal is critical for detecting these results.

JEL Codes: I14, Q51, Q53, Q54, D63

*Hernandez-Cortes: School for the Future of Innovation in Society and the School of Sustainability, Arizona State University (email: Danae.Hernandez-Cortes@asu.edu). Meng: Bren School, Dept. of Economics, and emLab, UC Santa Barbara and NBER (email: kmeng@bren.ucsb.edu) This paper has benefited from comments by Maximilian Auffhammer, Spencer Banzhaf, Youssef Benzarti, Severin Borenstein, Jim Bushnell, Kelly Caylor, Marc Conte, Chris Costello, Olivier Deschenes, Meredith Fowlie, Corbett Grainger, Larry Goulder, Kelsey Jack, Arturo Keller, Gary Libecap, Emily Maynard, Andrew Plantinga, David Pellow, Ed Rubin, Jim Salzman, Sam Stevenson, Alisa Tazhitdinova, Chris Tessum, and Paige Weber. We are also grateful for feedback received at various seminars and conferences. Kent Strauss and Vincent Thivierge provided excellent research assistance. Use was made of computational facilities purchased with funds from the National Science Foundation (CNS-1725797) and administered by the Center for Scientific Computing (CSC). The CSC is supported by the California NanoSystems Institute and the Materials Research Science and Engineering Center (MRSEC; NSF DMR 1720256) at UC Santa Barbara.

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 could 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). At the same time, the 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] While all environmental policies can generate pollution inequities, market-based policies, by emphasizing allocative efficiency, have raised questions about whether they present a distinct equity-efficiency trade-off (Shonkoff et al., 2011; Farber, 2012; Boyce and Pastor, 2013).

This concern has been particularly prominent for California's economy-wide greenhouse gas (GHG) cap-and-trade (C&T) program, which was introduced in 2013 to meet the state's GHG target. It created the world's second largest carbon market. While GHGs are a globally-mixed pollutant and thus not subject to local pollution concerns, GHGs are often co-emitted with local air pollutants such that a GHG C&T program could alter local air pollution disparities. The possibility that the California's GHG C&T program could widen the state's existing EJ gaps in local air pollution has, among other critiques, led to political opposition that temporarily paused initial program development in 2011 and nearly halted renewal efforts in 2017.[2] However, to date, there has been limited evidence on whether emissions reallocation due to the program has caused EJ gaps to widen.

[1] EJ gaps in the U.S. have been demonstrated at both local (Bullard, 2000; Bowen, 2002; Ringquist, 2005; Mohai, Pellow and Roberts, 2009; Banzhaf, Ma and Timmins, 2019) and national (Tessum et al., 2019; Colmer et al., 2020; Currie, Voorheis and Walker, 2020) scales.

[2] Similar concerns have appeared elsewhere: recent efforts to introduce state-level U.S. carbon pricing and to renew the European carbon market were opposed on EJ grounds (Leber, 2016; Herron, 2019; Transnational Institute, 2013).

Whether a market-based policy widens or narrows the EJ gap depends on how emitting facilities, their abatement costs, and disadvantaged communities are distributed across space, as well as what policy, if any, it replaces. Predicting the EJ gap consequence of a market-based policy ex-ante, however, is generally difficult because a key determinant, facility-level pollution abatement cost curves, is not typically observed. Under this data constraint, we show that what happens to the EJ gap is ambiguous regardless of whether a market-based policy replaces a no-policy scenario or an existing command-and-control policy , as in California's case. This underscores the need for ex-post empirical approaches, for which prior studies have found mixed EJ gap effects across different settings (Shadbegian, Gray and Morgan, 2007; Fowlie, Holland and Mansur, 2012; Grainger and Ruangmas, 2018; Shapiro and Walker, 2021; Sheriff, 2022).

Quantifying the EJ gap consequences of a market-based climate policy requires overcoming two empirical challenges. First, one must isolate how the policy altered emissions of regulated facilities apart from potentially confounding macroeconomic conditions and other "overlapping" climate policies that may interfere with market-based allocation. Second, policy effects on pollution emissions at the facility-level must be converted to location-level changes in pollution concentrations. We develop a statistical framework that jointly addresses these two challenges for California's GHG C&T program.

To overcome the first challenge, we compare emissions between C&T-regulated and -unregulated facilities which remove any common influence of macroeconomic conditions. By design, C&T eligibility is based on whether a facility's historic emissions is above a threshold. As a consequence, emissions from C&T-regulated and -unregulated facilities differ not only in pre-C&T levels but possibly also in pre-C&T trends. We use a differential trend-break model to control for differential pre-trends, focusing on how differential trends change after C&T's introduction. To inform the central distributional concern that market-based emissions allocation may have altered EJ gaps, we remove facilities additionally regulated by overlapping climate programs in California, such as renewable portfolio and low carbon fuel standards, as their abatement may be induced by those programs and not by the carbon market. As a result, our main analysis sample consists of industrial facilities subject only to the GHG C&T program as climate policy (hereafter "sample facilities"). While this comes with the benefit of isolating the EJ gap consequences of the carbon market, the downside is that our main sample covers only 5% of reported California GHG emissions. In sensitivity analyses, we consider broader samples, including the full set of C&T regulated facilities.

We find that C&T reduced emissions annually between 2012-2017 at an average rate of 9%, 5%, 4%, and 3% for GHG, $PM_{2.5}$, PM_{10}, and NO_x, respectively, across sample facilities. These emissions effects are robust to various model specification; to concerns about

differential trends across sectors; and to the possibility of regulatory spillover effects. In a placebo test that systematically imposes trend breaks across sample years, we generally detect the largest trend break in the year when the program was actually introduced. In a heterogeneity analysis, we find that larger emitting facilities undergo more abatement under C&T.

To address the second challenge, we explicitly embed an atmospheric dispersal model to determine how program-driven changes in facility-level pollution *emissions* alter the spatial distribution of pollution *concentrations*. The canonical framework for evaluating the pollution consequences of environmental policies requires characterizing the link between pollution "source" and "receptors" (Baumol and Oates, 1988). Failure to accurately account for pollution dispersal can lead to biased estimates even in otherwise valid quasi-experimental settings (Deschenes and Meng, 2018). To address this, recent studies use atmospheric dispersal models to examine how source-level emissions vary as a function of the demographic characteristics of downwind receptors (e.g., Grainger and Ruangmas (2018); Mansur and Sheriff (2019)). However, one may still need to convert source-level emissions effects onto the resulting spatial changes in pollution concentrations in order to obtain unbiased EJ gap effects. We build on this literature by combining estimates of source-level emissions effects (and its uncertainty) with an analysis of resulting EJ gap changes at the receptor level, as determined by the atmospheric dispersal model, a computationally-intensive procedure involving over two million pollution trajectories.

Employing a definition of a "disadvantaged" zip code that serves as a basis for California's EJ policies, we report three EJ gap findings from our sample facilities. 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 slowed down these previously widening EJ gaps so much that they have been narrowing since 2013. Between 2012-2017, the program reduced California's EJ gap by 7%, 6%, and 10% annually for $PM_{2.5}$, PM_{10}, and NO_x, respectively. Third, while EJ gaps have narrowed, they have not been eliminated: by 2017, the C&T program returned EJ gaps roughly to 2008 levels.

We subject our benchmark EJ gap effects to a variety of robustness checks. We detect similar EJ gap consequences when we augment our sample with facilities that are additionally regulated by overlapping climate policies, though we note the difficulty with disentangling the role of C&T from that of overlapping policies in these estimates. We find that allowing for heterogeneous emissions effects as a function of a facility's average emissions does not meaningfully alter EJ gap effects. We further demonstrate similar EJ gap effects when employing an alternative atmospheric dispersal model that generates secondary $PM_{2.5}$ con-

centrations as well as when using a finer scale definition of disadvantaged communities at the census tract-level. An analysis of spatial heterogeneity reveals that EJ gaps narrowed most for disadvantaged areas in California's Central Valley, while a few disadvantaged areas 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, such as restricting affected areas to within a facility's zip code or within distance circles (of varying radii) of a facility. We posit that our empirical approach may have broader applicability as many environmental policy settings requires researchers 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).

Finally, we discuss how to interpret our results in terms of a suitable counterfactual. Because California was legally required to meet a GHG target, an alternative economy-wide climate policy would likely have been adopted if not C&T, suggesting a continuation of pre-trends may not serve as a suitable counterfactual. At the same time, it is unclear which alternative policy would have been adopted. Given this ambiguity, we instead focus on the 2012-2017 estimated change, as done in other policy settings with pre-trends that differ in direction with post-trends (Greenstone and Hanna, 2014; Lovenheim and Willén, 2019). This implicitly assumes that the 2012 emissions distribution is held fixed in the absence of C&T and is more conservative than estimates based on extending pre-trends.

The paper is structured as follows: Section 2 introduces 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 with an overview of cap-and-trade (C&T) programs and how they can either widen or narrow existing pollution concentration gaps between disadvantaged and other communities. We then discuss 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.[3] Regulated facilities are then either given, or must purchase through trade, permits to cover their emissions. Traditionally, two consequences of C&T are emphasized. First, by placing a price on pollution, a C&T program requires polluting facilities to internalize (some of) the social costs of their emissions.[4] Second, by creating a market for pollution, a C&T program could equate marginal abatement costs across facilities, inducing relatively less abatement from facilities with steeper MAC curves and more abatement from facilities with flatter MAC curves. The resulting allocation of abatement could, in theory, achieve the aggregate emissions cap at the lowest total abatement cost across regulated facilities (Montgomery, 1972).

What is less clear is how emissions reallocation under C&T alters the spatial distribution of pollution concentration. In particular, there is growing concern that the same market forces intended for improvements in allocative efficiency may also be widening the difference in pollution concentrations experienced between disadvantaged and other communities, which we call the "environmental justice gap" (or EJ gap).

The introduction of C&T can either widen or narrow an existing EJ gap. The direction of this effect may also be hard to anticipate ex-ante. We demonstrate this using a simple economic framework in Figure 1.[5] There are two polluting facilities with emissions (e) on the horizontal axis and permit prices (τ) on the vertical axis. The first facility is upwind of a disadvantaged community with a marginal abatement curve labeled in solid orange (labeled 'D'). The second facility is upwind of a non-disadvantaged community and has a marginal abatement curve in dashed gray (labeled 'A').

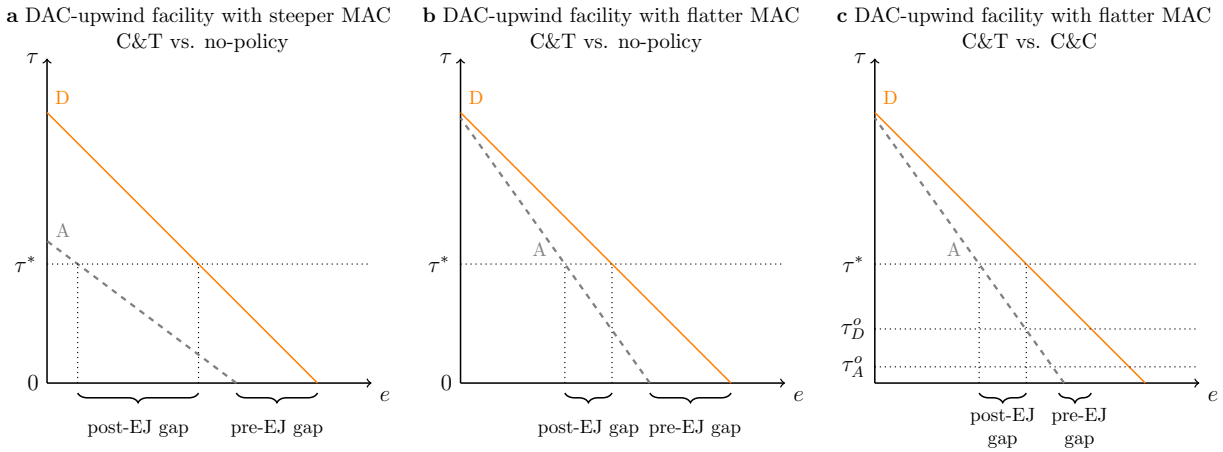
First, consider the case where C&T replaces a no-policy scenario, or when $\tau_D^o = \tau_A^o = 0$. When C&T is introduced, each facility's MAC is equated to the equilibrium permit price τ^*. Air quality improves for both disadvantaged and other communities. But because the improvement may not be the same for both communities, C&T could widen or narrow the EJ gap. If the DAC facility has a steeper MAC curve (panel a of Figure 1), then the DAC facility will abate less than the non-DAC facility with C&T, and the EJ gap will widen. If, however, the DAC facility has a flatter MAC curve (panel b of Figure 1), then the DAC facility abates more than the non-DAC facility, narrowing the EJ gap. Unfortunately, facility-level MAC curves are rarely observed, making it challenging to forecast what would happen to the EJ

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

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

[5] A richer theoretical framework would also incorporate various political economy considerations, such as procedural justice concerns regarding obstacles disadvantaged communities face in environmental decision making.

Figure 1: EJ gap under cap-and-trade



NOTES: Panels illustrate how the introduction of C&T can widen or narrow 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 is shown as solid orange line (labeled 'D'). The marginal abatement cost curve for facility upwind of a non-disadvantaged community is shown as dashed gray line (labeled 'A'). τ^* indicates the permit price under C&T. Panel a: C&T vs. no-policy ($\tau_D^o = \tau_A^o = \tau^o$) with DAC-upwind facility having a relatively steeper MAC curve. Panel b: C&T vs. no-policy ($\tau_D^o = \tau_A^o = \tau^o$) with DAC-upwind facility having a relatively flatter MAC curve. Panel c: C&T vs. command-and-control (C&C) policy with heterogeneous shadow prices for DAC and non-DAC facilities, τ_D^o and τ_A^o, respectively with DAC-upwind facility having a relatively flatter MAC curve.

gap ex-ante.

This difficulty is compounded in the case when C&T is overlaid on existing environmental regulation. Consider, for example, a prior command-and-control regulation that generates heterogeneous shadow pollution prices across facilities (panel c of Figure 1). Because such prices are not observed, the introduction of C&T could widen or narrow the EJ gap even for a given set of marginal abatement cost curves.

Figure 1 underscores the need for ex-post analyses. In practice, such studies must further account for several empirical features. First, the introduction of C&T may coincide with changing macroeconomic conditions, requiring one to compare C&T effects for regulated facilities relative to unregulated facilities. Second, when evaluating a GHG C&T program, the EJ gap effect depends on the extent in which GHG and local air pollutants are co-produced. Third, in practice, the spatial relationships between polluting facilities, their abatement costs, and where disadvantaged communities reside are far more complex than is illustrated in Figure 1. In particular, when multiple emitting facilities can affect air quality in multiple locations, one needs to explicitly characterize how emitted pollution disperses spatially. Section 4 details how we overcome these challenges.

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, creating the first economy-wide GHG target in the U.S. which required state-wide GHG emissions to return to a 1990 level by 2020.[6] In 2016, California met its 2020 GHG target four years early. That same year, the state extended its GHG target to 40% below the 1990 level by 2030.

A central program for achieving these GHG targets is cap-and-trade, introduced in 2013 and administered by the California Air Resources Board (CARB).[7] The program requires participation by 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.[8] The program becomes more stringent over time, mandating a declining cap on aggregate emissions across eligible facilities. Today, California has the world's second largest carbon market by permit value, following the emissions trading system in Europe.

California's GHG C&T program was adopted under a particular context. First, local air pollution in California is subject to a combination of local, state, and federal air quality regulations, many of which are command-and-control policies. Thus, it is likely that facilities faced heterogeneous shadow prices on local pollution prior to the introduction of the GHG C&T program, as captured by panel c of Figure 1. Second, cap-and-trade is not California's only climate program. The adoption of C&T occurred against a backdrop of various sector-specific climate programs that overlap with the economy-wide C&T, including a renewable portfolio standard for electricity generators and a low carbon fuel standard for refiners, both of which were adopted before 2013. Indeed, an ex-ante analysis of California's GHG C&T program demonstrated a potentially large role played by these overlapping programs (Borenstein et al., 2019). When overlapping policies bind – i.e., induce sector-specific abatement that the economy-wide C&T would not on its own – it becomes hard to attribute emission changes in these facilities to market-based allocation under C&T. To better isolate the role of C&T, our main analysis sample focuses on GHG-emitting facilities that were not subject to these overlapping policies.

Finally, there is the issue of what is the appropriate counterfactual had California not adopted C&T. Because program eligibility is based on whether a facility's historic emissions

[6] Covered greenhouse gases include CO_2, CH_4, N_2O, HFCs, PFCs, SF6, NF3 and other fluorinated GHGs.

[7] 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.

[8] Note that the emissions allocation under C&T includes the use of offset credits which allow a regulated facility to meet up to 8% of its emissions with credits purchased from GHG abatement projects originating in sectors not covered by the cap. Our estimated C&T effects are therefore inclusive of the use of offsets.

is above a threshold, emissions from C&T-regulated and -unregulated facilities differ not only in pre-C&T levels but possibly also in pre-C&T trends, an empirical feature we indeed detect in Section 5.1. While we employ a differential trend-break model (discussed in Section 4) to control for differential pre-trends, a question arises as to whether continuation of these pre-trends serve as a suitable counterfactual. We do not think so. California's 2020 GHG target under Assembly Bill 32 was legally binding, implying that the state likely would have adopted some form of economy-wide climate policy if not for the C&T program. In contrast, pre-2013 trends are estimated in a period without an economy-wide climate policy. At the same, it is unclear which particular C&T alternative would have been adopted instead: policy proposals at the time ranged from local regulations targeting facilities located near disadvantaged communities to uniform abatement requirements across all facilities. Given this ambiguity about a suitable counterfactual, we instead emphasize the estimated 2012-2017 change: how much did pollution emissions and the resulting EJ gap change for sample C&T-regulated facilities after C&T's adoption. This is similar to the estimated policy-induced change emphasized by Greenstone and Hanna (2014), who also detect differential pollution pre-trends in the context of Indian environmental regulations, and in other applied settings where pre-trends are detected (Lovenheim and Willén, 2019; Rambachan and Roth, 2022). Implicitly, the estimated 2012-2017 change assumes a counterfactual holding the 2012 distribution of emissions across C&T-regulated facilities fixed over time in the absence of C&T. As we show in Section 5.1, with rising differential emissions pre-trends and falling post-trends, this implies a more conservative EJ gap effect than one based on extending pre-trends.

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.[9] We observe GHG as well as criteria air pollution emissions for both C&T-regulated and non-regulated stationary facilities, before and after the introduction of the C&T program.[10]

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

[10] 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

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 pollution exposure and socioeconomic indicators developed by the California Environmental Protection Agency.[11] In our benchmark analysis, we use CalEnviroScreen v1.1 which assigns disadvantaged status at the zip code-level and is constructed using pre-2013 indicators, mitigating concerns that C&T may directly impact disadvantaged community designation. 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. As a robustness check, we also use a later CalEnviroScreen v3.0 which has the benefit of defining a disadvantaged community at a finer census tract-level but at the cost of being constructed using post-2013 indicators. We further augment our zip code level data with average 2008-2012 population obtained from the U.S. Census Bureau.

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. With the exception of very large emitters with continuous emissions monitoring equipment that are not in our sample, CEIDARS emissions is primarily estimated using data on facility-level inputs, outputs, and emissions factors, akin to how emissions are estimated for the US EPA's National Emissions Inventory. For CEIDARS, the added regulatory structure between CARB and the state's 35 Air Districts provides several additional layers of data quality checks as well as the use of locally-relevant emissions factors.

[11] The socioeconomic indicators used to construct CalEnviroScreen do not explicitly include a location's racial and/or ethnic composition, though many included indicators such as poverty levels, educational attainment, unemployment rate, are correlated with racial and/or ethnic composition (CalEPA, 2018). In a robustness check, we use a location's minority share of population as a measure of a disadvantaged community.

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 The C&T program regulates stationary GHG-emitting facilities producing at least 25,000 metric tons of annual CO_2e during any year between 2009-2012. We exploit this facility-level eligibility criteria and the program's 2013 introduction to estimate its effects on GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x facility-level emissions during 2008-2017.[12] As noted in above, 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 emissions trends. Our empirical test therefore examines whether differential emission trends exhibit a break after 2013.[13]

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.[14] For facility j in year t, Y_{jt}^p is annual emissions of pollutant

[12]Data availability for facility GHG and criteria air pollution emissions for both regulated and unregulated facilities and for periods before and after the program's introduction is not common across cap-and-trade programs. For example, facility-level pre-program emissions are not directly observed for the European 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 are able to compare regulated and unregulated units within the same jurisdiction. Note that 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.

[13]A differential trend-break model is similar to that of a standard difference-in-difference model with the distinction that one is interested in the change in differential outcome *trends* after the policy rather than on the change in differential outcome *means*.

[14]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.

$p \in \{GHG, PM_{2.5}, PM_{10}, NO_x, SO_x\}$. Because emissions exhibit a skewed distribution and contain zero values, we apply an inverse hyperbolic sine transformation, which like a log 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:

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

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. ν_{jt}^p is clustered at the county-level to allow for arbitrary forms of heteroskedasticity and serial correlation within a county.

The coefficient 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, in annual percentage point changes. The coefficient κ_2^p is the change, or break, in the differential emission trend after the program's introduction during 2013-2017, such that $\kappa_1^p + \kappa_2^p$ is the post-C&T differential emission trend, also in percentage point changes.

The functional form in equation (1) imposes linear differential pre- and post-trends that go through the same point in 2012, implicitly assuming that there is no mean shift in differential emissions after 2012. To examine whether this functional form is justified, we also estimate a more flexible version of equation (1) with annual C&T coefficients and compare the shape of these coefficients to a linear trend break function.[15]

We employ two sample restrictions. First, despite the C&T program's unique eligibility criteria and timing, the presence of overlapping climate programs, such as the renewable portfolio standard for electricity generators and the low carbon fuel standard for refineries, imply that emission changes from facilities further regulated by these programs may not reflect abatement induced by the carbon market. Since our central interest is to understand how the carbon market altered EJ gaps, we remove electricity generators and refineries

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.

[15] Specifically, the more flexible version of equation (1) is:

$$asinh(Y_{jt}^p) = \sum_{\substack{2008 \leq \tau \leq 2017 \\ \tau \neq 2012}} \kappa_\tau^p[D_i \times \mathbf{1}(t = \tau)] + \phi_j^p + \gamma_t^p + \nu_{jt}^p \quad (1')$$

from our sample, which constitutes 70% of reported California 2008-2012 GHG emissions.[16] 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.[17] These sample restrictions, while reasonable for identification purposes, may limit the external validity of our results: our benchmark sample of industrial facilities contributed only 5% of reported California GHG emissions during 2008-2012. In a robustness check, we consider facilities emitting the other 95% of reported GHG emissions.

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.

C&T generates heterogeneous emission abatement levels across regulated facilities. We recover facility-level heterogeneity in abatement levels by applying a hyperbolic sine transformation to the first two terms of equation (1) and the estimated facility-level fixed effect.[18] Because facilities differ by average emission levels, this allows us to translate a common percentage effect into heterogeneous C&T-driven abatement levels, as shown in Figure S4.[19] Whether a common percentage effect is appropriate in our setting is an empirical matter. To explore this, we estimate variants of equation (1) that interact the post-C&T trend break with linear and quadratic functions of each facility's average annual emissions, allowing the

[16] 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).

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

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

$$\widehat{Y}^p_{jt} = sinh\left(\widehat{\kappa}^p_1[C_j \times t] + \widehat{\kappa}^p_2[C_j \times \mathbf{1}(t \geq 2013) \times t] + \widehat{\phi}^p_j\right) e^{RMSE^2/2}$$

where hat notation indicates estimated parameters and RMSE is the root mean squared error from equation (1). The $e^{RMSE^2/2}$ adjustment is akin to that required for a log-transformed variable. To see this, if $X \sim \mathcal{LN}(\mu, \sigma^2)$, then $E[sinh(X)] = E[(1/2)(e^X - e^{-X})] = (1/2)(e^{\hat{\mu}+\frac{\hat{\sigma}^2}{2}} - e^{-\hat{\mu}+\frac{\hat{\sigma}^2}{2}}) = sinh(\hat{\mu})e^{\frac{\hat{\sigma}^2}{2}}$ (MacKinnon and Magee, 1990). 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.

[19] 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.

data to determine whether the C&T percentage effect varies with emission levels. These models consistently reveal that larger emitting facilities abate more in levels under C&T, slightly more so than is assumed with a common percentage effect (see Table S3). However, predicted emissions that allow for this heterogeneity do not meaningfully alter pollution disparity results (see col. 6 of Table S7).[20]

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 that employs topography and real-time meteorological conditions to generate spatial distributions of criteria pollution concentrations driven by C&T.

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 real-time 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 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). At the same time, it is less 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. To see if secondary $PM_{2.5}$ concentra-

[20]We note that these results are consistent with a recent literature building on Melitz (2003)'s heterogeneous firm model that emphasizes heterogeneity in abatement levels as a function of a facility's baseline emissions level, with more productive firms having both higher baseline emission levels and lower pollution intensities (Forslid, Okubo and Ulltveit-Moe, 2018; Shapiro and Walker, 2018).

tions 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 that models secondary $PM_{2.5}$ concentrations (Tessum, Hill and Marshall, 2017). InMAP, however, has a major disadvantage: while HYSPLIT's underlying meteorological inputs cover our 2008-2017 sample period, InMAP only models dispersal patterns in 2005.

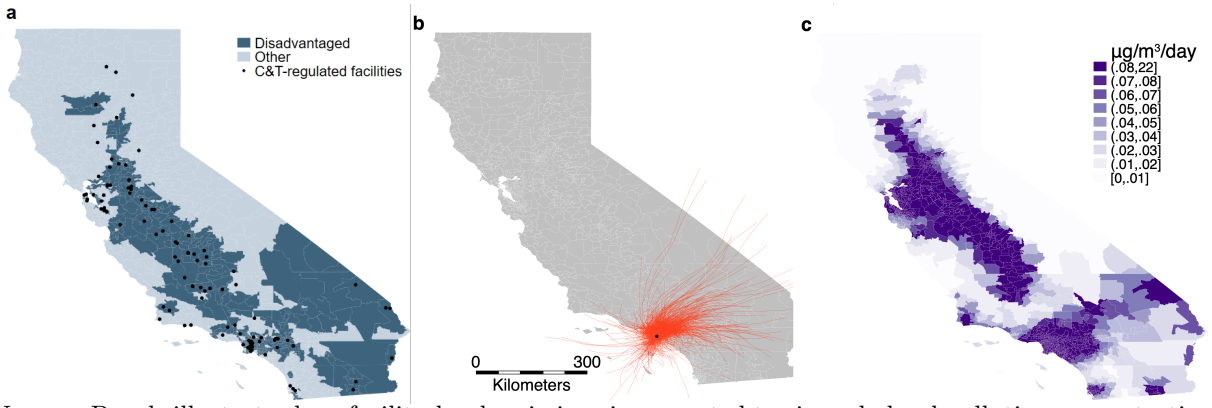
We note several features of our HYSPLIT implementation. First, to account for high-frequency variation in meteorological conditions, we run forward particle trajectories at four hour intervals, implicitly assuming that annual emissions are distributed uniformly within the year. In a robustness check, we relax this assumption by distributing C&T-driven annual emissions at the monthly level according to sector-specific monthly output shares.[21] Each trajectory runs for 24 hours, a duration long enough to ensure most emitted particles leave California.[22] 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 distribution of 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.

Following common HYSPLIT practice, we convert the distribution of trajectories generated by HYSPLIT into concentration units by summing 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

[21] Specifically, using Federal Reserve Economic Data, we obtain U.S. aggregate monthly output for manufacturing (NAICS 31-33) and 2-digit NAICS level monthly output for extraction (NAICS 21) during the 2008-2017 period. We then distribute C&T-driven annual emissions according to these monthly output shares when feeding emissions into HYSPLIT.

[22] 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.

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 the spatial distribution of 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.

above each zip code in units of μg/m^3/*day*.[23] Figure 2c shows our benchmark HYSPLIT-generated daily concentration (in μg/m^3/day) for each zip code, averaged across 2008-2017 for $PM_{2.5}$. Figure S2 similarly shows average 2008-2017 zip-code concentrations for PM_{10}, NO_x, and SO_x.[24] Note that pollution concentration levels in Figure S2 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.

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^p_{it}, for criteria air pollutant $p \in$

[23] 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.

[24] Figure 2, Figure S2, 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.

$\{PM_{2.5}, PM_{10}, NO_x, SO_x\}$, and estimate the following specification:

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

where ψ_i^p are zip code-specific dummies and δ_t^p are year-specific dummies. The coefficient 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. The coefficient β_2^p 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 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.[25] 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.[26] Appendix A.1 provides more details on this bootstrap procedure.

[25] 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 $\beta_1^p + \beta_2^p < 0$.

[26] As with prior literature, we omit uncertainty associated with atmospheric dispersal, or the mapping

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.[27]

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 dispersal whereby concentrations in each location 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 is not generally informative of the EJ gap effect. Not only can this coefficient differ in value from the true EJ gap effect, but it is also not necessarily positively correlated or have the same sign as the EJ gap effect. The reason is that a facility-level analysis fails to account for the full complexity of pollution dispersal patterns by overlooking the possibility that a facility can alter pollution concentrations in both disadvantaged and other locations. For example, a facility-level analysis may recognize that emissions from facilities in Los Angeles and San Francisco affect air quality in both

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.

[27] 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 to emissions from these plants, but cannot speak to the policies that are affecting coal-fired power plant emissions.

disadvantaged and other receptor locations, but it does not undergo the additional exercise of tallying up pollution from Los Angeles and San Francisco at each receptor location.

To illustrate this, Appendix A.2 shows the restrictions on pollution dispersal patterns required in order for estimates from a facility-level analysis to equal the true EJ gap effect: emissions from each facility can *only* affect disadvantaged communities or *only* affect non-disadvantaged communities. This assumption is readily rejected in our setting. For each sample C&T regulated facility-year observation, we calculate the share of downwind affected locations containing disadvantaged communities. Figure S3 shows that this share is always within the unit interval. That is, emissions from every sample C&T regulated facilities alters pollution concentrations in both disadvantaged and non-disadvantaged communities.

In summary, our approach combines both facility- and location-level analyses. This enables us 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.

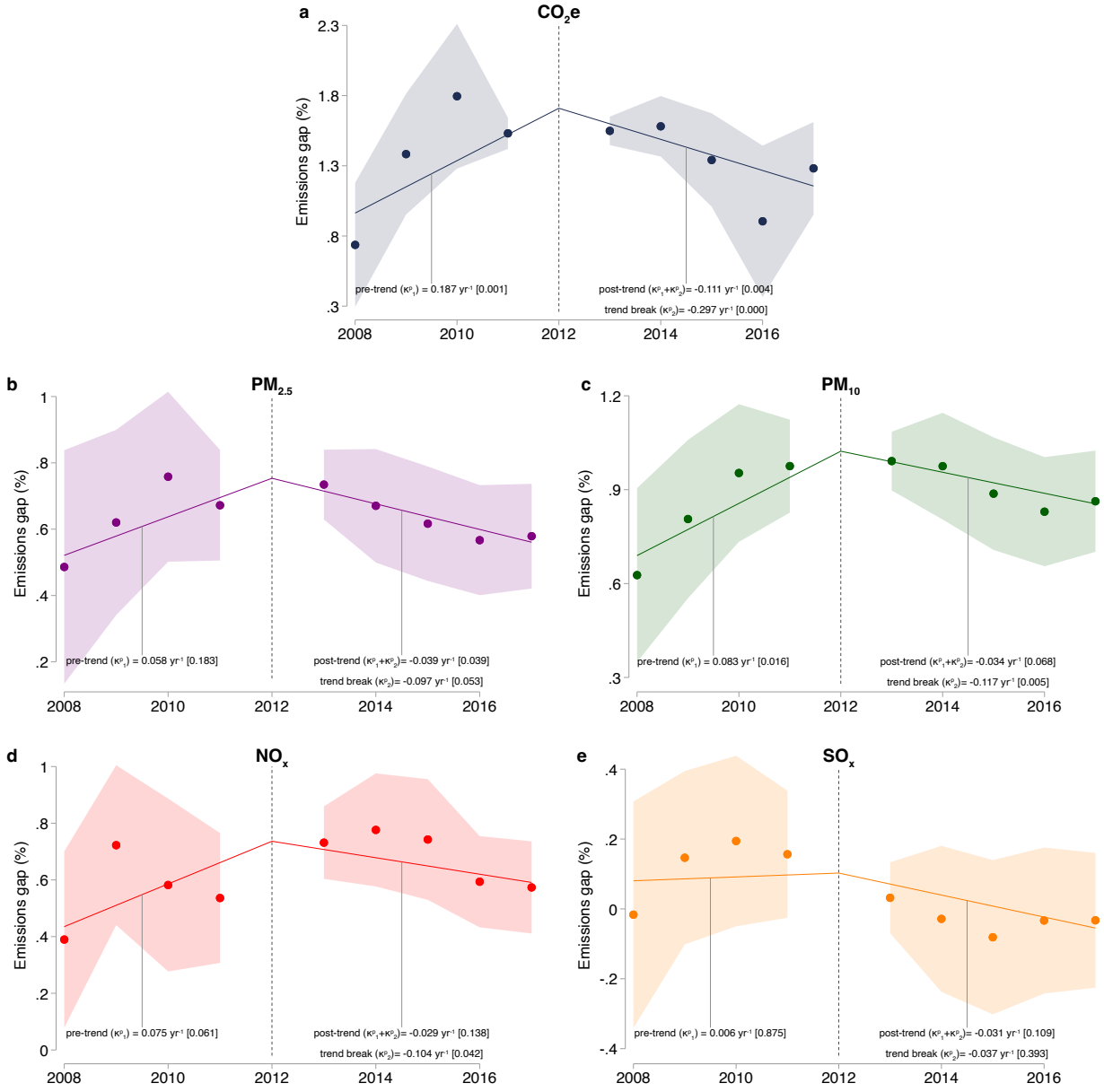
5.1 Cap-and-trade effects on emissions

Main results To verify whether the linear trend break functional form in equation (1) is appropriate, we begin with estimating a more flexible variant of that model with annual C&T effect coefficients, as shown in equation (1'). Those annual coefficients, plotted for CO_2e, $PM_{2.5}$, PM_{10}, NO_x, and SO_x in Figure 3, appear approximately linear before and after C&T's introduction and do not exhibit a mean shift after C&T, consistently with the functional form in equation (1).

Table 1 reports the trend coefficients before and after the introduction of C&T for GHG and the four criteria air pollutants.[28] Prior to the program, the gap in GHG emissions between regulated and unregulated facilities increased at an annual rate of 19 percentage points (i.e., κ_1^p). Following the introduction of the program, this trend slowed down dramatically leading the gap in GHG emissions to fall at an annual rate of 11 percentage points (i.e.,

[28]The estimating sample sizes differ across the columns of Table 1 because reporting thresholds differ by pollutant, as noted in Footnote 10.

Figure 3: Emissions effect of cap-and-trade program



NOTES: Panels show the estimated emissions gap (in %) between C&T-regulated and unregulated facilities for CO_2e, $PM_{2.5}$, PM_{10}, NO_x, and SO_x, respectively. Dots show year-specific effects from equation (1'). Solid lines show linear fits from equation (1). Associated text indicates point estimates and p-values (in brackets) for the pre-C&T linear trend (κ_1^p), post-C&T trend break (κ_2^p), and post-C&T linear trend ($\kappa_1^p + \kappa_2^p$), as reported in Table 1. Estimates centered at the 2008 emissions difference.

$\kappa_1^p + \kappa_2^p$) between 2012-2017. Columns 2-4 exhibit similar patterns for $PM_{2.5}$, PM_{10}, and NO_x. For SO_x, the trend break is negative but not statistically significant; we henceforth will not emphasize SO_x results.

Percentage point trend effects can be hard to interpret. Instead, one can translate the

Table 1: Trend break in emissions

	Outcome is (asinh) emissions				
	(1)	(2)	(3)	(4)	(5)
	CO_2e	$PM_{2.5}$	PM_{10}	NO_x	SO_x
pre-trend (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]
trend break (κ_2^p)	-0.297	-0.097	-0.117	-0.104	-0.037
	(0.077)	(0.048)	(0.040)	(0.050)	(0.043)
	[<0.001]	[0.053]	[0.005]	[0.042]	[0.393]
post-trend ($\kappa_1^p + \kappa_2^p$)	-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]	[0.109]
2012-2017 annual abatement pct	-8.51	-4.68	-4.15	-2.94	-9.35
2012-2017 total abatement (tons)	-3.2e+06	-97.89	-140.66	-519.94	-62.10
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 (κ_1^p), post-C&T differential emissions trend break (κ_2^p), and post-C&T differential emissions trend ($\kappa_1^p + \kappa_2^p$) from equation (1) for GHG, $PM_{2.5}$, PM_{10}, NO_x, and SO_x across columns. Average annual abatement percentage and total abatement level (in metric tons) for 2012-2017 shown. All models include facility-specific and year-specific dummy variables. Standard errors clustered at the county-level in parentheses, p-value in brackets.

estimates in Table 1 to the average 2012-2017 annual percent change across sample regulated facilities.[29] During 2012-2017, the C&T program reduced emissions annually at a rate of 9%, 5%, 4%, and 3% for GHG, $PM_{2.5}$, PM_{10}, and NO_x, respectively, for the average sample regulated facility. Figure S4 shows the implied heterogeneous facility-level C&T-driven abatement between 2012-2017 for GHG and criteria air pollutants, as defined in Footnote 18. Altogether, sample regulated facilities reduced 3.2 million tons of CO_2e between 2012-2017.[30]

[29] This is calculated by averaging $(\frac{\widehat{Y}^p_{j,2017} - \widehat{Y}^p_{j,2012}}{\widehat{Y}^p_{j,2012}})/5$, as defined in footnote 18, across regulated sample facilities for each pollutant p.

[30] This amounts to an average abatement of 6,740 tons per facility per year during a period with permit prices between $12-15 per ton. While this may be a surprising amount of abatement at such permit prices, Colmer et al. (2020) find that the European carbon market lowered emissions by 28,830 tons per firm per year under a $21 per ton price.

Robustness checks We subject these emission effects to several robustness checks. First, Figure S5 considers placebo program start years, plotting κ_2^p for GHG and criteria pollution emissions from variants of equation (1) that impose alternative C&T start years across 2009-2016. Generally, 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, with column 1 reproducing our benchmark results. In Column 2 we add a dummy variable for C&T-regulated facilities after C&T's introduction. We do not detect a statistically significant post C&T mean shift in differential emissions, consistent with the annual C&T coefficient shown in Figure 3. 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 3 replaces year fixed effects with sector-by-year fixed effects. Column 4 drops the handful of facilities whose treatment status switched only in 2017. Columns 5 and 6 change the 75th percentile average GHG emissions cutoff to the 70th and 80th percentiles.[31] 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 specification confirms that large emitting facilities are indeed abating more in levels and that our benchmark model, which estimates an average trend break coefficient across facilities (regardless of size) is conservatively understating this dimension of emissions abatement heterogeneity. 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) violation 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. Our first

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

test observes that firms with multiple facilities could more readily reallocate pollution across their facilities. In column 2, we restrict the control group of unregulated facilities to those whose parent company only operates a single plant.[32] Our second test notes that a facility located in a county under U.S. Clean Air Act nonattainment for a particular pollutant may be more constrained from increasing pollution levels. This idea is implemented in column 3, which restricts the sample of unregulated facilities to those located in nonattainment counties for that pollutant under the Clean Air Act.[33] 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.[34] 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.[35]

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

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

[33] 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.

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

[35] 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.

(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.

Table 2: Trend break in the environmental justice gap

	(1) $PM_{2.5}$	(2) PM_{10}	(3) NO_x	(4) SO_x
pre-trend (β_1^p)	0.042	0.065	0.085	0.037
	(0.015)	(0.017)	(0.037)	(0.025)
	[0.006]	[<0.001]	[0.026]	[0.151]
trend break (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]
post-trend ($\beta_1^p + \beta_2^p$)	-0.021	-0.026	-0.058	-0.064
	(0.015)	(0.020)	(0.050)	(0.027)
	[0.159]	[0.203]	[0.252]	[0.024]
Trend pct change ($100 * \beta_2^p/\beta_1^p$)	-149.699	-139.739	-168.282	-272.291
	(36.369)	(29.971)	(53.377)	(66.044)
	[<0.001]	[<0.001]	[0.002]	[<0.001]
2012-2017 annual EJ gap change pct	-6.57	-5.78	-10.15	-17.01
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 (β_1^p), the post-C&T EJ gap trend break (β_2^p), the post-C&T EJ gap trend ($\beta_1^p + \beta_2^p$), the percentage change in the EJ gap trend following the introduction of the C&T program ($\frac{\beta_2^p}{\beta_1^p} * 100$), and the 2012-2017 average annual EJ gap percentage change 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 shown in Figure S6, 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 4. Across criteria pollutants, 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 by 150%, 140%, and 170% for $PM_{2.5}$, PM_{10}, NO_x, respectively, after the program's introduction. Figure 4 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 with estimates centered at the 2008 EJ gap.[36] Between 2012-2017, the program reduced California's EJ gap by 7%, 6%, and 10% annually for $PM_{2.5}$, PM_{10}, and NO_x, respectively (also shown in Table 2). Figure 4 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.

Spatial heterogeneity Estimates from equation (2) shown in Table 2 and Figure 4 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.[37] Figure 5 shows the percentage change in the EJ gap trend following the introduction of C&T for each disadvantaged zip code. Across pollutants, post-C&T EJ gaps narrowed the most for disadvantaged zip codes in California's Central Valley. Figure 5 also shows a cluster of zip codes in Los Angeles County that experienced widening post-C&T EJ gaps. Figure S8 shows histograms for the distribution of percentage changes in EJ gap trends across disadvantaged zip codes. Figure S7 replicates Figure 5 but in grayscale.

[36] Specifically, the annual coefficients in Figure 4 are β_τ^p from

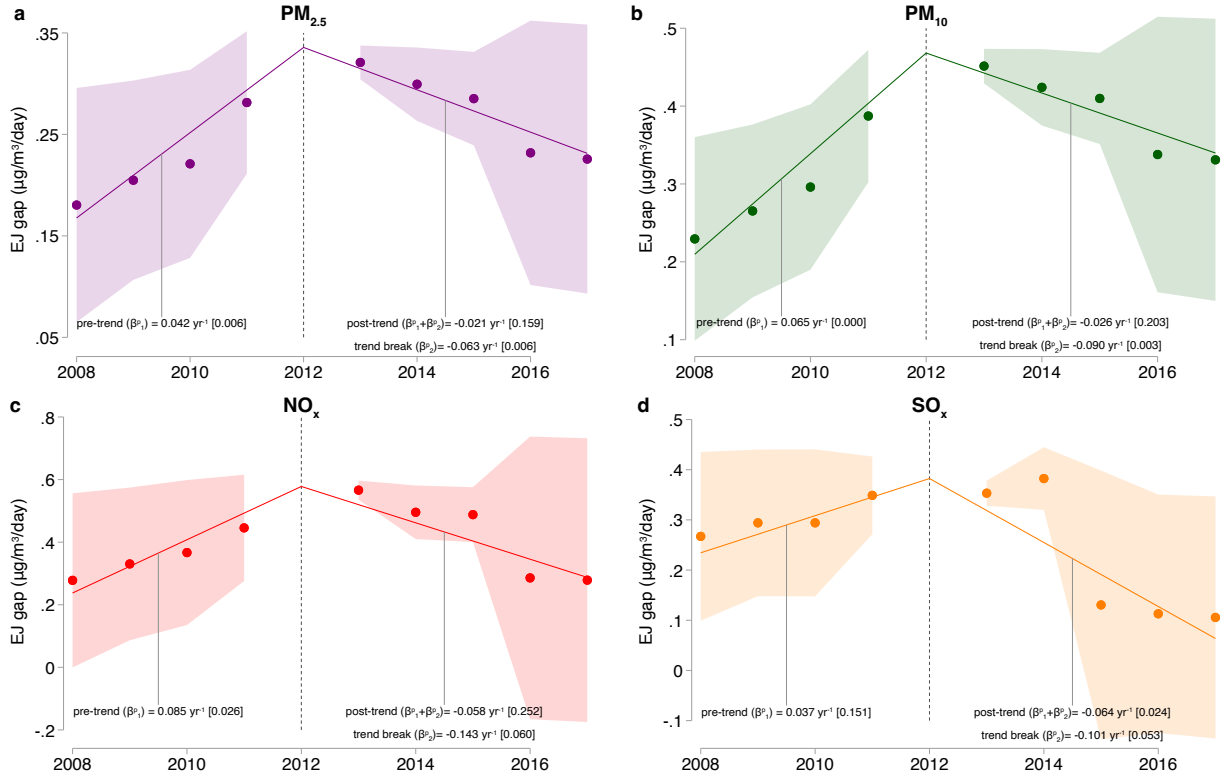
$$E_{it}^p = \sum_{\substack{2008 \leq \tau \leq 2017 \\ \tau \neq 2012}} \beta_\tau^p [D_i \times \mathbf{1}(t = \tau)] + \psi_i^p + \delta_t^p + \epsilon_{it}^p$$

[37] 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 5 and S8 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.

Figure 4: Environmental justice gap effect of the cap-and-trade program



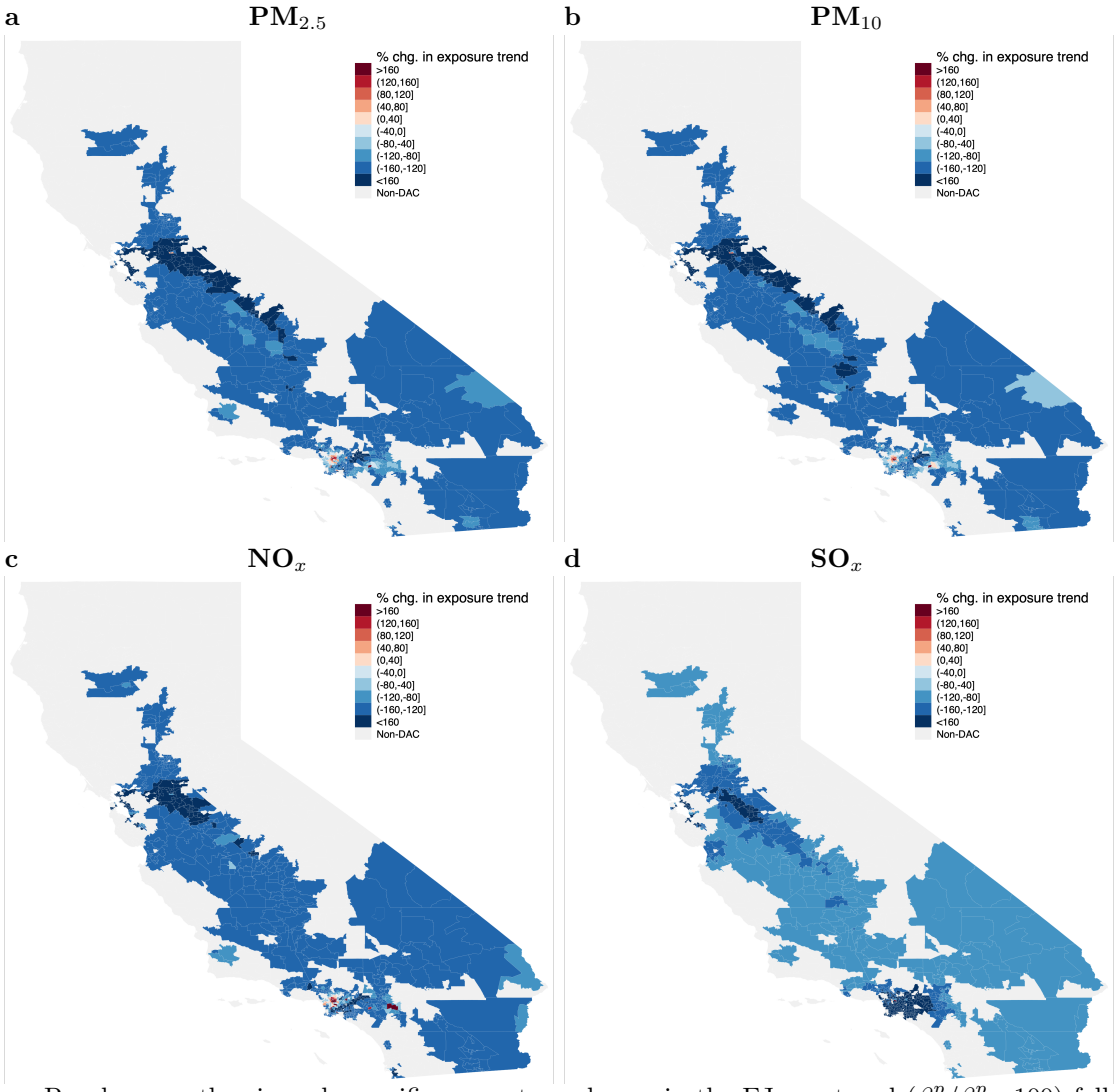
NOTES: Panels show the estimated average pollution concentration gap (in $\mu g/m^3$/day) between disadvantaged and other zip codes (i.e., "EJ gap") during 2008-2017 for (a) $PM_{2.5}$, (b) PM_{10}, (c) NO_x, and (d) SO_x, respectively. Dots show year-specific EJ gap with 95% confidence interval. Solid lines show linear fits from equation (2). Associated text indicates point estimates and p-values (in brackets) for the pre-C&T linear trend (β_1^p), post-C&T trend break (β_2^p), and post-C&T linear trend ($\beta_1^p + \beta_2^p$), as reported in Table 2. Estimates centered at the 2008 EJ gap shown in Table S6. Confidence intervals and p-values account for uncertainty in C&T predicted emissions (ν_{it}^p from equation (1) via the bootstrap procedure in Appendix A.1) and shown in Figure S6, and county-level heterogeneity in EJ gap effects of arbitrary form (ϵ_{it}^p from equation (2)).

Robustness checks We subject our EJ gap trend break to several robustness checks, shown in Figure 6. Because of the computational demands of our bootstrap procedure across steps 1-3 (detailed in Appendix A.1), Figure 6 presents only point estimates for 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, drawing a comparison with the point estimate and 95% confidence interval of our benchmark result which does account for statistical uncertainty from equation (1) via our bootstrap procedure.[38]

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 be-

[38]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).

Figure 5: Spatial heterogeneity in EJ gap effects



NOTES: Panels maps the zip code-specific percentage change in the EJ gap trend ($\beta_2^p/\beta_1^p * 100$) following the introduction of the C&T program for disadvantaged zip codes across (a) $PM_{2.5}$, (b) PM_{10}, (c) NO_x, and (d) 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. Grayscale version separating zipcodes with negative and positive change shown in S7.

tween 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 6 and column 1 of Table S7); when we replace year fixed effects with sector-by-year fixed effects in equation (1) (M3 of Figure 6 and column 2 of Table S7); and when we drop facilities that switched regulatory status in 2017 (M4 of Figure 6 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 6 and

columns 4-5 of Table S7). These alternative facility sample restrictions do not alter EJ gap trend effects.

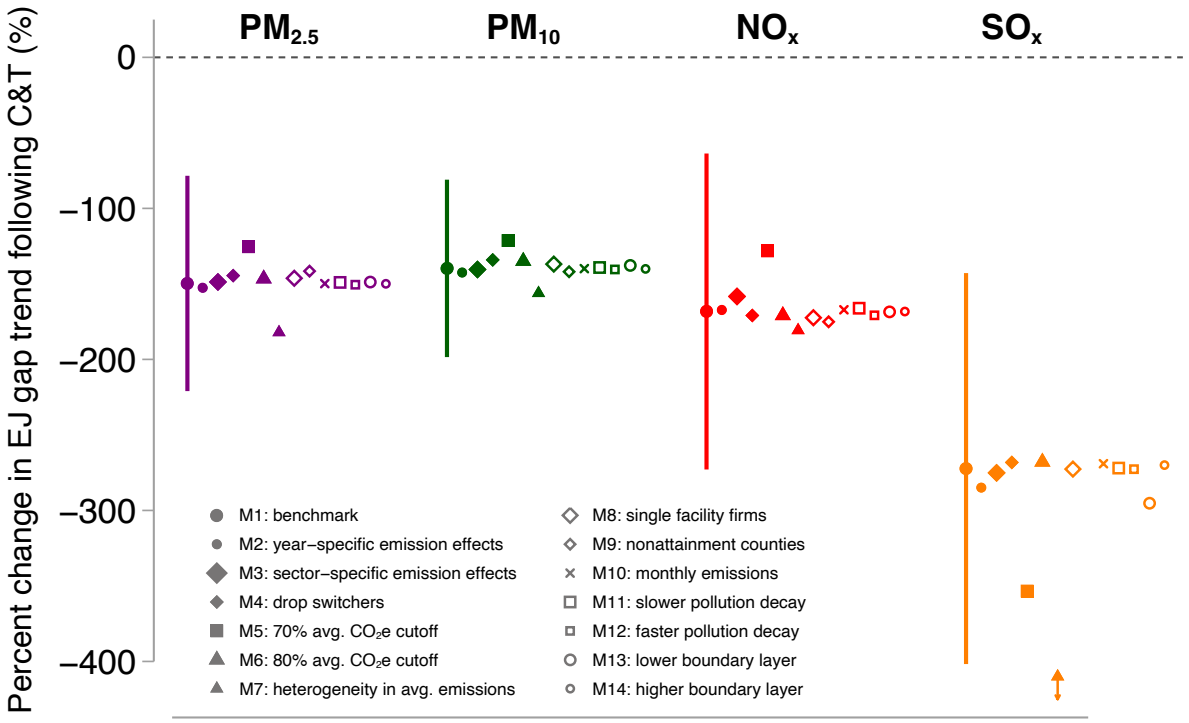
We further allow the post C&T emissions trend break to vary as a linear function of sample average emissions. Recall from the heterogeneous emissions effects shown in column 2 of Table S3 that large-emitting facilities are abating more than is assumed in our baseline model which assumes a common percentage 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 6 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 whose parent company only operates a single facility and those in counties under Clean Air Act nonattainment (M8-9 of Figure 6 and columns 7 and 8 of Table S7). SUTVA concerns do not alter EJ gap trend effects.

We conduct five robustness checks within step 2. For sample facilities in some sectors, we are able to obtain monthly output data during our 2008-2017 sample period. For those facilities, we distribute C&T-driven annual emissions using the monthly share of annual output when feeding emissions into HYSPLIT and show that accounting for monthly variability in these sectors does not alter EJ gap results (M10 of Figure 6 and column 1 Table S8). 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 (M11 of Figure 6 and column 2 Table S8) or a 10% smaller half-life parameter which implies a faster decay rate (M12 of Figure 6 and column 3 of Table S8). Likewise our results are little affected when we lower the height of the planetary boundary layer to 0.5 km (M13 of Figure 6 and column 4 Table S8) or raise it to 2 km (M14 of Figure 6 and column 5 Table S8).

We also conduct multiple robustness checks within step 3. The first set of checks consider alternative error structures for ϵ^p_{it}. We find that precision increases when we allow ϵ^p_{it} 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 6 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 7 of Table S8).

Equation (2) examines the EJ gap in daily pollution levels of $\mu g/m^3/day$, the concen-

Figure 6: 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 whose parent company only operates a single plant. M9: restricting unregulated facilities to those in counties under Clean Air Act nonattainment. M10: distributes annual C&T-driven emissions by monthly output. M11: applying a slower pollution decay (i.e., 10% larger half-life parameter). M12: applying a faster pollution decay (i.e., 10% smaller half-life parameter). M13: applying a planetary boundary layer set at 0.5 km. M14: applying planetary boundary layer set at 2 km. Point estimates also reported in Tables S7-S8.

tration 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.

In Table S10, we employ CARB's CalEnviroScreen v3.0, an updated designation for dis-

advantaged communities at the census tract-level, the finest spatial scale for which California defines disadvantaged status. A census tract better captures local variation in socio-economic characteristics that underlie EJ concerns. However, this and all census tract-level versions of the CalEnviroScreen present endogeneity concerns as some underlying indicators were observed after the start of the C&T program and thus may itself be affected by C&T. Table S10 shows similar EJ gap effects across criteria pollutants when conducing census tract-level analysis using CalEnviroScreen v3.0 as compared with our benchmark zip code-level analysis.

The socioeconomic indicators used to construct CalEnviroScreen does not include an explicit measure of racial and/or ethnic composition, though many included indicators are correlated with racial and/or ethnic composition. In Table S11, we replace our CalEnviroScreen-based indicator for a disadvantaged community with a dummy for whether a zip code's average 2008-2012 minority share of population according to the U.S. Census (defined as the sum of Black and Hispanic share of population) is above the median value across California zip codes. We find similar EJ gap effects when using this alternative definition of a disadvantaged community.

In our benchmark sample, we exclude electricity generators and refineries from our sample because the presence of overlapping climate programs applied to those facilities imply that their emission abatement are likely driven by these programs and not by the carbon market. We also excluded facilities with sample average annual GHG emissions above the 75th percentile to ensure comparability between treated and control facilities. Table S12 shows EJ gap effects when we add these facilities into our sample. Their inclusion does not qualitatively change our main findings.

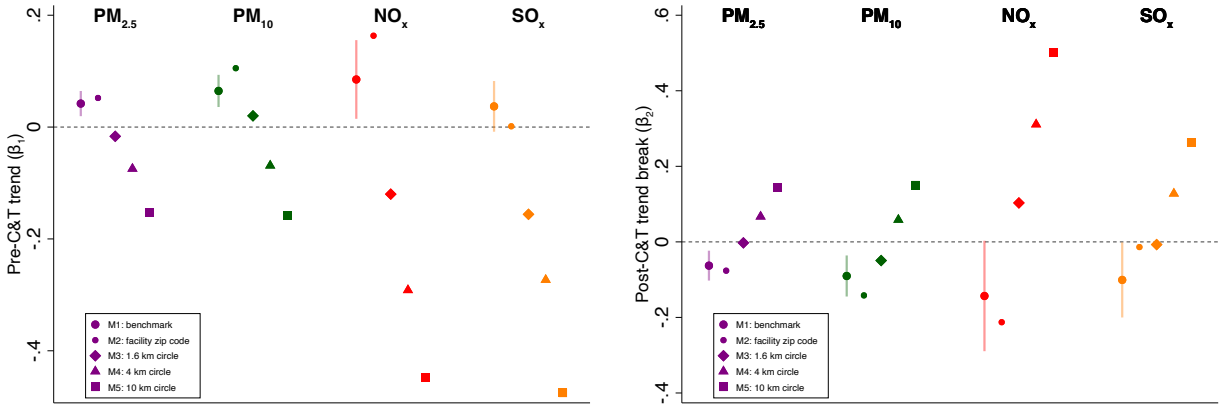
Finally, to examine the potential role of secondary $PM_{2.5}$, we replace HYSPLIT in step 2 of our procedure with InMAP, a reduced-complexity dispersal model based on WRF-Chem, which incorporates atmospheric chemistry to model total (i.e., primary plus secondary) $PM_{2.5}$ concentrations from C&T-driven facility-level $PM_{2.5}$, NO_x, and SO_x, emissions (Tessum, Hill and Marshall, 2017).[39] InMAP, however, has one major limitation: it only models dispersal patterns in 2005. Because InMAP does not model dispersal patterns during our sample period (unlike HYSPLIT), we are unable to directly compare estimates using InMAP- and HYSPLIT-generated concentrations. Instead, we compare within InMAP by looking at how EJ gap effects differ for InMAP-generated primary and total $PM_{2.5}$ concentrations. If these two InMAP estimates are similar, it becomes more plausible that the true EJ gap effect on total $PM_{2.5}$ is similar to that of our estimated EJ gap effect on HYSPLIT-generated primary

[39] 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.

$PM_{2.5}$. Table S13 replicates the structure of Table 2. Note that the units differ between Tables S13 and 2 as HYSPLIT provides the stock of pollution concentration for a given period whereas InMAP provides the flow. Column 1 examines InMAP-generated primary $PM_{2.5}$ concentrations while column 2 examines InMAP-generated total $PM_{2.5}$ concentrations. These two EJ gap effects are indeed similar.

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 7: 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 S14.

Figure 7 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.[40] In M1, we show our benchmark estimate where pollution dispersal is modeled by HYSPLIT. In M2 (and column 1 of Table S14), 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 S14), 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.

6 Discussion

Many market settings are characterized by efficiency-equity trade-offs. We find that disparities in local air pollution concentrations from industrial facilities subject only to California's carbon market fell following its introduction. 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 policy 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 from sample facilities 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, including changes in energy prices and wages. A comprehensive understanding of welfare inequality would also examine other pollutants emphasized in the environmental justice literature. For example, Sheriff (2022) finds that the concentration gap in toxics between white and minority communities did not increase as a consequence of California's carbon market. One 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 entry decisions by polluters (Weber, 2020). Third, a broader notion of

[40] Unlike Figure 6, Figure 7 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.

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, 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.

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.

Boyce, James K, and Manuel Pastor. 2013. "Clearing the air: incorporating air quality and environmental justice into climate policy." *Climatic change*, 120(4): 801–814.

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

CalEPA. 2018. "Analysis of Race/Ethnicity, Age, and CalEnviroScreen 3.0 Scores."

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.

Farber, Daniel A. 2012. "Pollution markets and social equity: Analyzing the fairness of cap and trade." *Ecology LQ*, 39: 1.

Forslid, Rikard, Toshihiro Okubo, and Karen Helene Ulltveit-Moe. 2018. "Why are firms that export cleaner? International trade, abatement and environmental emissions." *Journal of Environmental Economics and Management*, 91: 166–183.

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 Rema Hanna. 2014. "Environmental Regulations, Air and Water Pollution, and Infant Mortality in India." *American Economic Review*, 104(10): 3038–72.

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.

Lovenheim, Michael F., and Alexander Willén. 2019. "The Long-Run Effects of Teacher Collective Bargaining." *American Economic Journal: Economic Policy*, 11(3): 292–324.

MacKinnon, James G, and Lonnie Magee. 1990. "Transforming the dependent variable in regression models." *International Economic Review*, 315–339.

Mansur, Erin T., and Glenn Sheriff. 2019. "Do Pollution Markets Harm Low Income and Minority Communities? Ranking Emissions Distributions Generated by California's RECLAIM 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.

Melitz, Marc J. 2003. "The impact of trade on intra-industry reallocations and aggregate industry productivity." *econometrica*, 71(6): 1695–1725.

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.

Rambachan, Ashesh, and Jonathan Roth. 2022. "A More Credible Approach to Parallel Trends." *Review of Economic Studies*, forthcoming.

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.

Shadbegian, Ronald J, Wayne Gray, and Cynthia Morgan. 2007. "Benefits and costs from sulfur dioxide trading: a distributional analysis." In *Acid in the Environment*. 241–259. Springer.

Shapiro, Joseph S., and Reed Walker. 2018. "Why Is Pollution from US Manufacturing Declining? The Roles of Environmental Regulation, Productivity, and Trade." *American Economic Review*, 108(12): 3814–54.

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

Sheriff, Glenn. 2022. "California's GHG trading program and the equity of air toxic releases."

Shonkoff, Seth B, Rachel Morello-Frosch, Manuel Pastor, and James Sadd. 2011. "The climate gap: environmental health and equity implications of climate change and mitigation policies in California—a review of the literature." *Climatic Change*, 109(1): 485–503.

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 S6 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 4.

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

In Section 4, we detailed a multi-step procedure to estimate the change in the EJ gap as a consequence of the the C&T policy. In this section we discuss how our procedure differs from an alternative approach found in the literature.

A.2.1 Estimand

We begin by first formalizing our estimand. Let j index polluting facilities. The (heterogeneous) change in emissions due to C&T is ΔY_j for regulated facility j, with regulation status denoted by $C_j = 1$ and $C_j = 0$ otherwise. Locations are indexed by i, with $D_i = 1$ indicating if location i is a disadvantaged community and $D_i = 0$ otherwise. For simplicity, we assume the same number of disadvantaged and non-disadvantaged locations, $N = \sum_{i:D_i=1} D_i = \sum_{i:D_i=0}(1 - D_i)$ and that population is constant across locations. Let $\mathbf{W}$ be the source-receptor matrix obtained by the atmospheric dispersal model, with element $w_{ji} = 1$ if emissions from facility j disperses to location i and $w_{ji} = 0$ otherwise.[41]

The EJ gap effect, our estimand, is the difference in the average change in pollution concentration between disadvantaged and other communities due to C&T-driven emission changes from regulated facilities. Formally, this is

$$\begin{aligned}
\theta &= \underbrace{\frac{1}{N}\sum_{i:D_i=1}\sum_{j:C_j=1}\Delta Y_j w_{ji}}_{\substack{\text{avg. DAC concentration}\\ \text{from C\&T-driven}\\ \text{emissions}}} - \underbrace{\frac{1}{N}\sum_{i:D_i=0}\sum_{j:C_j=1}\Delta Y_j w_{ji}}_{\substack{\text{avg. non-DAC concentration}\\ \text{from C\&T-driven}\\ \text{emissions}}} \\
&= \frac{1}{N}\sum_{j:C_j=1}\Delta Y_j\left(\sum_{i:D_i=1} w_{ji}\right) - \frac{1}{N}\sum_{j:C_j=1}\Delta Y_j\left(\sum_{i:D_i=0} w_{ji}\right) \\
&= \frac{1}{N}\sum_{j:C_j=1}\Delta Y_j\left(\sum_{i:D_i=1} w_{ji} - \sum_{i:D_i=0} w_{ji}\right) \qquad \text{(A.1)}
\end{aligned}$$

Note that our procedure in Section 4 is designed to recover θ: equation (1) estimates $\widehat{\Delta Y_j}$, the pollution dispersal model provides the source-receptor matrix $\mathbf{W}$, and equation (2) takes these inputs to obtain the estimate $\widehat{\theta}$.

A.2.2 Alternative approach: dispersal-augmented facility-level regressions

Previous studies argue that θ can be estimated with only a facility-level analysis, augmented with information from the dispersal model. This approach typically examines how a facility's emission effect from the policy varies with the share of its downwind locations that are disadvantaged, as provided by the dispersal model (Grainger and Ruangmas, 2018; Mansur

[41] We assume w_{ji} to be dichotomous in order to match the literature's definition of the share of affected locations that are disadvantaged, or s_j below. In our actual implementation, we apply spatial decay rates to pollution dispersal patterns which allows w_{ji} to vary continuously within the unit interval.

and Sheriff, 2019). Specifically, this approach estimates

$$\Delta Y_j = \phi_0 C_j + \phi_1 C_j s_j + \phi_2 s_j + \varepsilon_j \tag{A.2}$$

where $s_j = N_j^D/(N_j^D+N_j^A)$ is the share of affected downwind locations that is disadvantaged, with $N_j^D = \sum_{i:D_i=1} w_{ji}$ and $N_j^A = \sum_{i:D_i=0} w_{ji}$ being the total number of disadvantaged and other locations affected by facility j, respectively. The coefficient of interest is typically $\widehat{\phi}_1$, the additional emissions effect for facilities that disproportionately affect disadvantaged locations.

Unfortunately, $\widehat{\phi}_1$ is not generally informative of θ. First, it may be that equation (A.2) is misspecified in the sense that s_j is not the relevant dimension of heterogeneity for ΔY_j. In that case, estimates from equation (A.2) cannot be inserted into equation (A.1). Even when s_j is the relevant dimension of heterogeneity, recovery of θ may still not be possible. To see why, note that when s_j is the relevant dimension of heterogeneity, we have $\widehat{\Delta Y}_j = \widehat{\phi}_0 + \widehat{\phi}_1 s_j$. Inserting into equation (A.1) yields

$$\widehat{\theta} = \frac{1}{N} \sum_{j:C_j=1} (\widehat{\phi}_0 + \widehat{\phi}_1 s_j)(2s_j - 1)(N_j^D + N_j^A) \tag{A.3}$$

showing that $\widehat{\phi}_1$ in general does not equal $\widehat{\theta}$.[42]

But are $\widehat{\theta}$ and $\widehat{\phi}_1$ systematically related? Consider first whether $\widehat{\theta}$ and $\widehat{\phi}_1$ move in the same direction. The derivative of $\widehat{\theta}$ with respect to $\widehat{\phi}_1$ is

$$\frac{d\widehat{\theta}}{d\widehat{\phi}_1} = \frac{1}{N} \sum_{j:C_j=1} \underbrace{(2s_j - 1)}_{\gtrless 0} \underbrace{s_j(N_j^D + N_j^A)}_{\geq 0} \gtrless 0$$

which shows no systematic correlation. Next, consider whether $\widehat{\phi}_1$ and $\widehat{\theta}$ consistently have the same sign. We can reject this with the following example. Assume $\widehat{\phi}_0 = 0$ and the normalization $N_j^D + N_j^A = N \ \forall j$. Next, suppose $s_1 = 1$ for the first facility and $s_{j>1} < 0.5$ for all other facilities. Equation (A.3) becomes

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

[42] Observe also that having $\widehat{\phi}_0$, $\widehat{\phi}_1$, and s_j does not recover $\widehat{\theta}$. One also needs either N_j^D or N_j^A, which is typically missing from equation (A.2).

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

A.2.3 Source of bias

Why do $\widehat{\theta}$ and $\widehat{\phi}_1$ differ? The issue is that a facility-level analysis does not account for the full complexity of pollution dispersal patterns. In particular, it overlooks the possibility that each facility can alter pollution concentrations in both disadvantaged and other locations. Equation (A.2) may recognize, for example, that emissions from facilities in Los Angeles and San Francisco affect air quality in both disadvantaged and other receptor locations, but it does not undergo the additional exercise of tallying up pollution from Los Angeles and San Francisco at each receptor location.

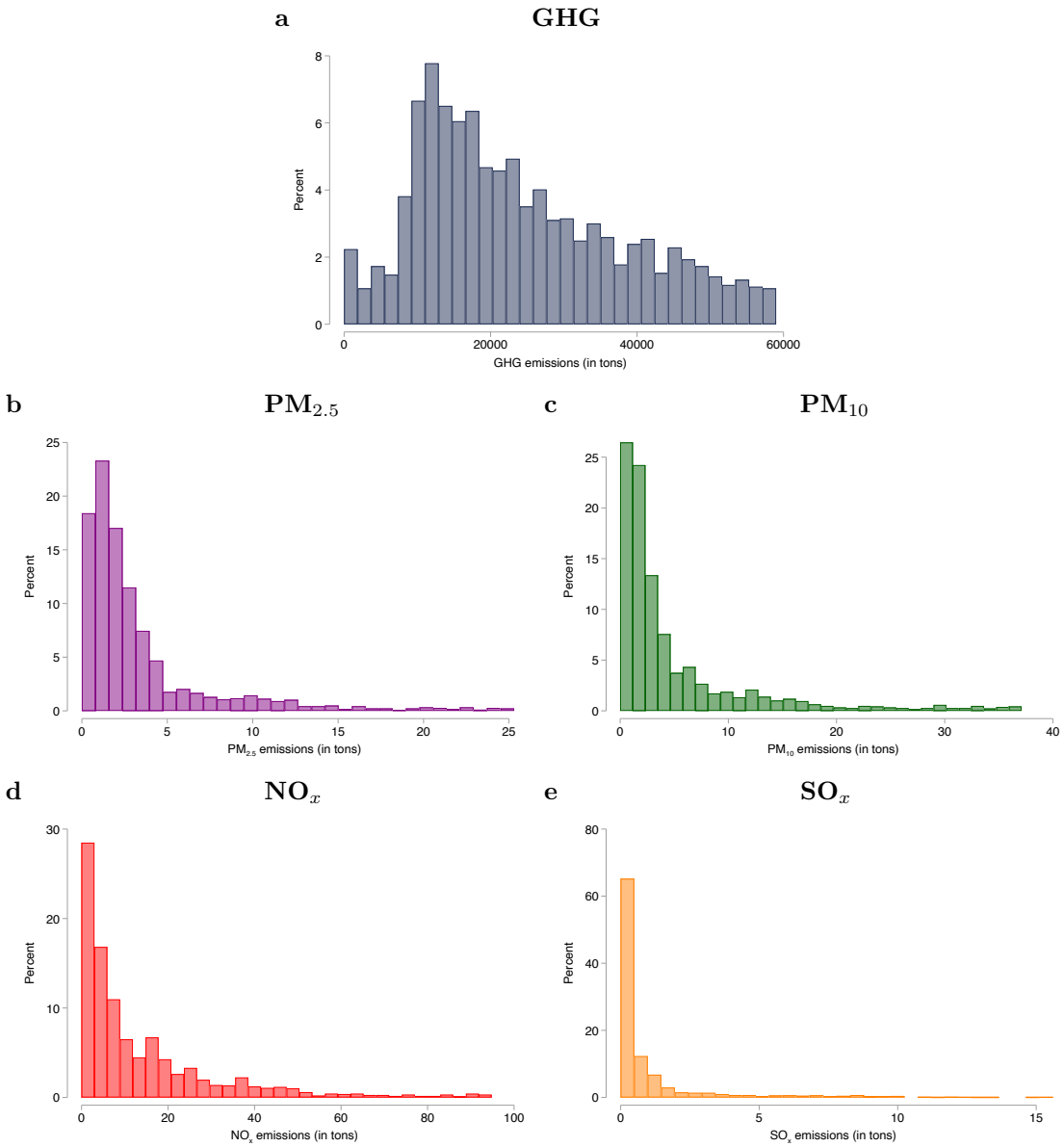
To see this formally, we show that $\widehat{\phi}_1$ can equal $\widehat{\theta}$ if one assumes pollution from each facility *only* affects one type of location. Let J facilities only affect disadvantaged communities (i.e., $s_j = 0$) and another J facilities only affect non-disadvantaged communities (i.e., $s_j = 1$). Further imposing the normalization $(N_j^D + N_j^A) = \frac{N}{J}\ \forall j$, we have

$$\begin{aligned}\theta &= \frac{1}{N}\left(\sum_{j:C_j=1,s_j=0} -\widehat{\phi}_0(N_j^D + N_j^A) + \sum_{j:C_j=1,s_j=1} (\widehat{\phi}_0 + \widehat{\phi}_1)(N_j^D + N_j^A)\right) \\ &= \frac{N}{NJ}(-J\widehat{\phi}_0 + J\widehat{\phi}_0 + J\widehat{\phi}_1) \\ &= \widehat{\phi}_1\end{aligned}$$

Observe the restrictions on pollution dispersal considered here: facilities can *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. Our California setting violates this assumption. For each sample C&T regulated facility-year observation, we calculate the share of downwind affected locations containing disadvantaged communities. Figure S3 shows that this share is always within the unit interval. That is, emissions from sample C&T regulated facilities always alter pollution concentrations in both disadvantaged and non-disadvantaged communities. Recovery of θ therefore requires the extra step of converting facility-level emissions changes onto location-level concentration changes, as is done with our approach, detailed in Section 4.

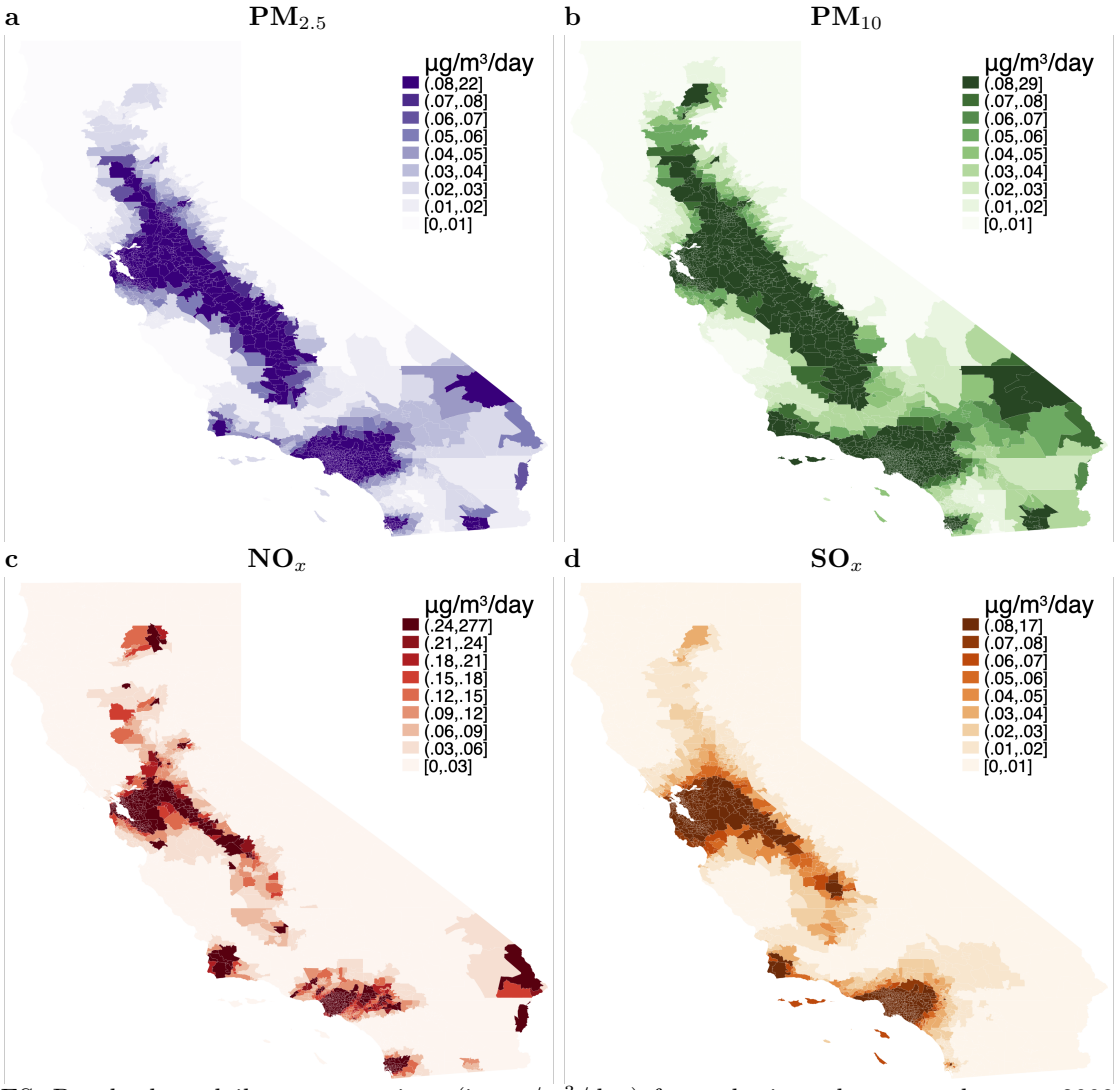
Appendix Figures

Figure S1: Distribution of sample facility-year emissions



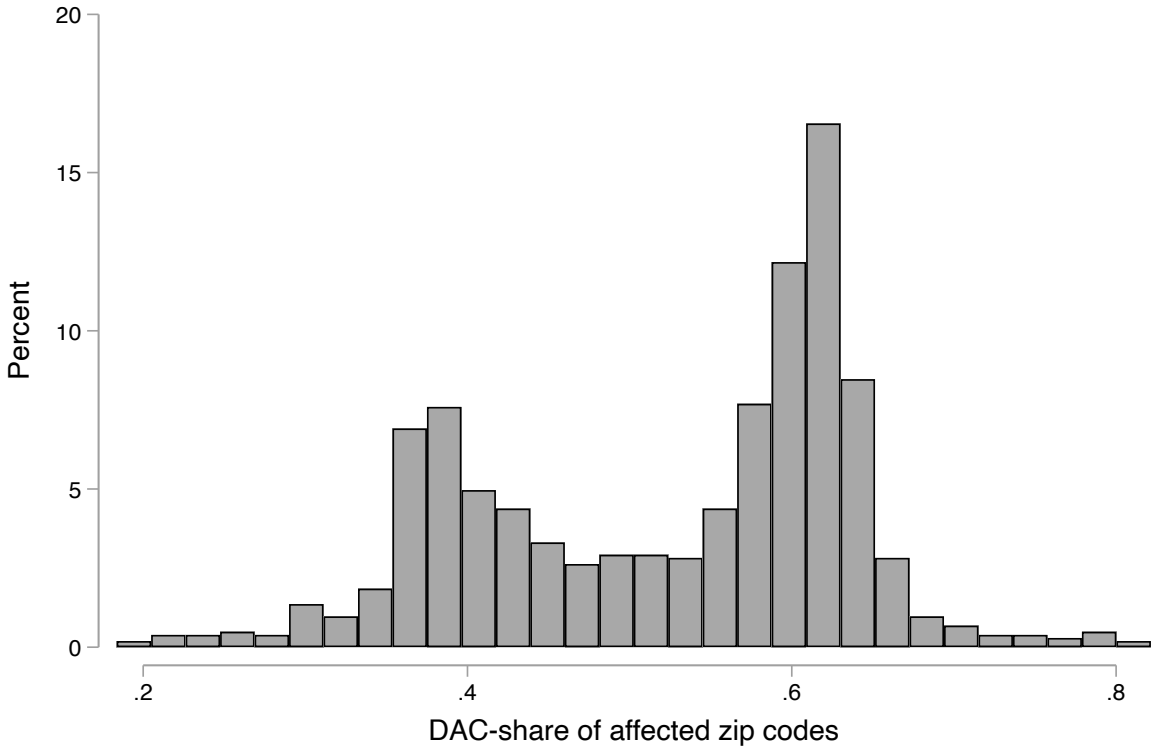
NOTES: Panels show the distribution of facility-year (a) GHG, (b) $PM_{2.5}$, (c) PM_{10}, (d) NO_x, and (e) SO_x emissions for sample observations. Observations above the 95th percentile are truncated.

Figure S2: Average pollution concentrations driven by C&T regulated facilities



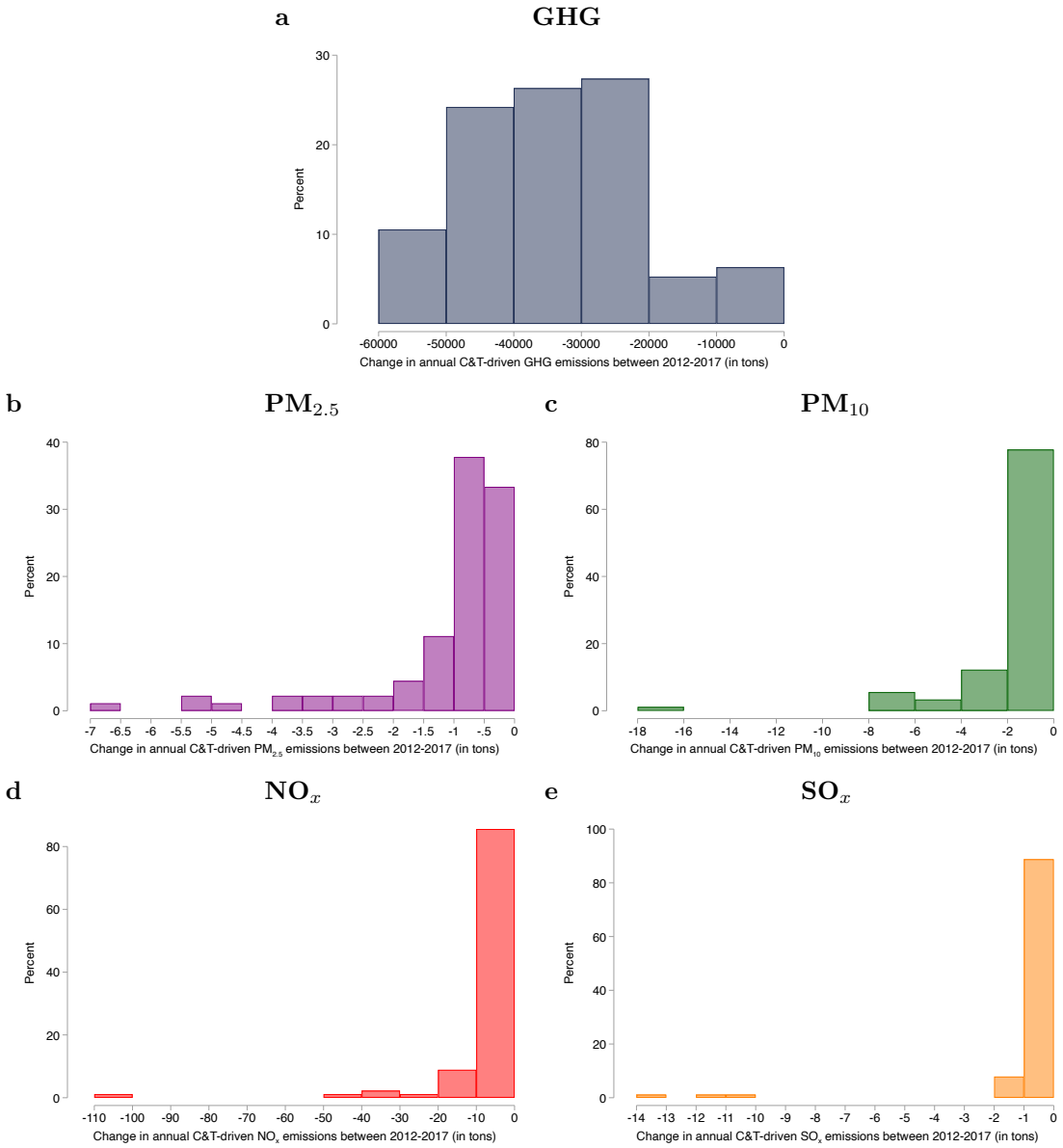
NOTES: Panels show daily concentrations (in $\mu g/m^3$/day) for each zip code averaged across 2008-2017 from GHG C&T-regulated facilities as modeled in step 2 by HYSPLIT for (a) $PM_{2.5}$, (b) PM_{10}, (c) NO_x, and (d) SO_x, respectively.

Figure S3: Share of zip codes affected by sample regulated facilities that contains DACs



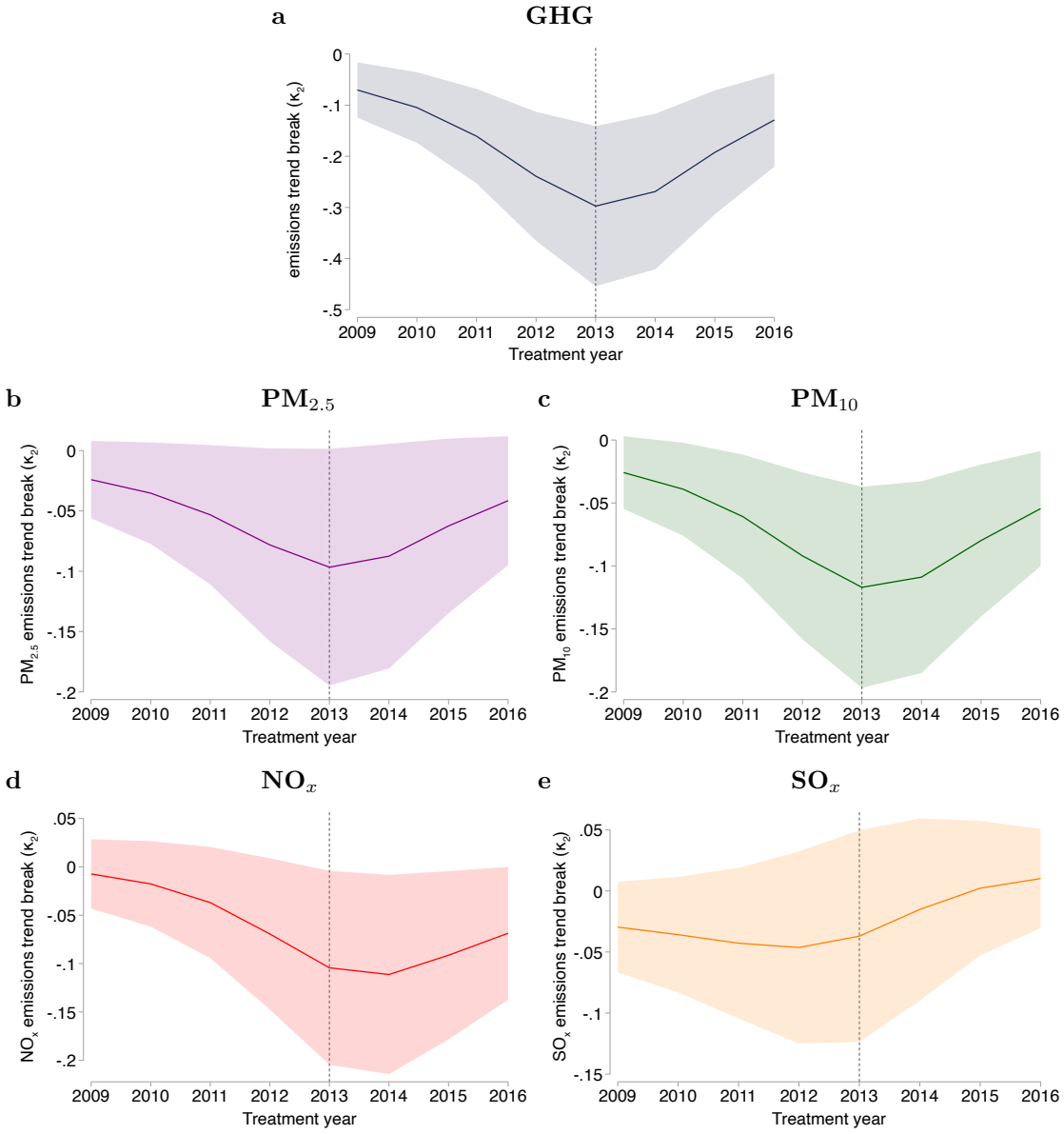
NOTES: Figure plots the distribution of the share of downwind affected zip codes that is designated as DAC across sample C&T regulated facility-year observations.

Figure S4: Facility-level C&T-driven abatement between 2012-2017



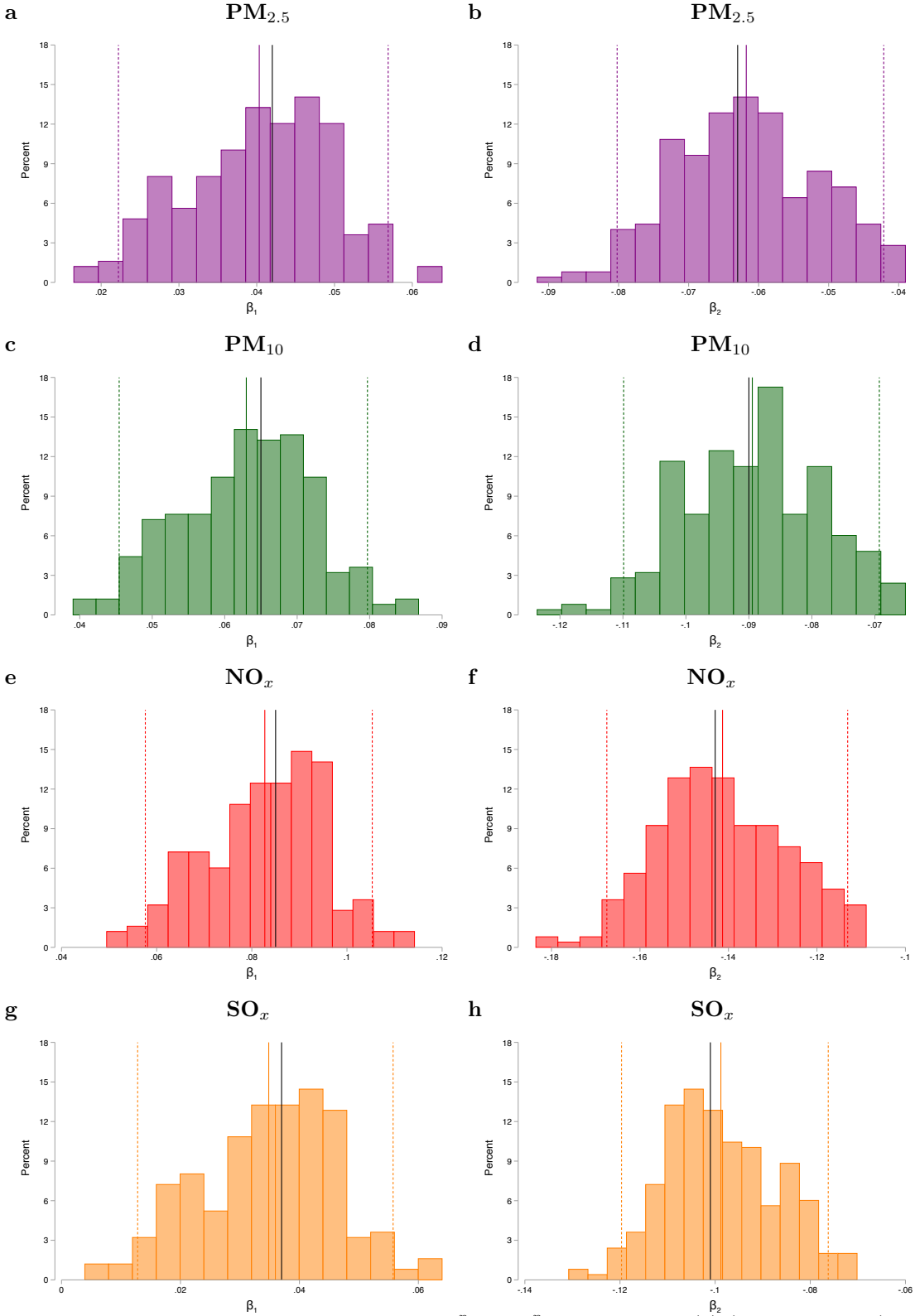
NOTES: Panels show the distribution of facility-level change in C&T-driven pollution between 2012-2017 (or abatement) predicted from step 1 for (a) GHG, (b) $PM_{2.5}$, (c) PM_{10}, (d) NO_x, and (e) SO_x emissions, respectively.

Figure S5: Emissions robustness: placebo C&T program timing



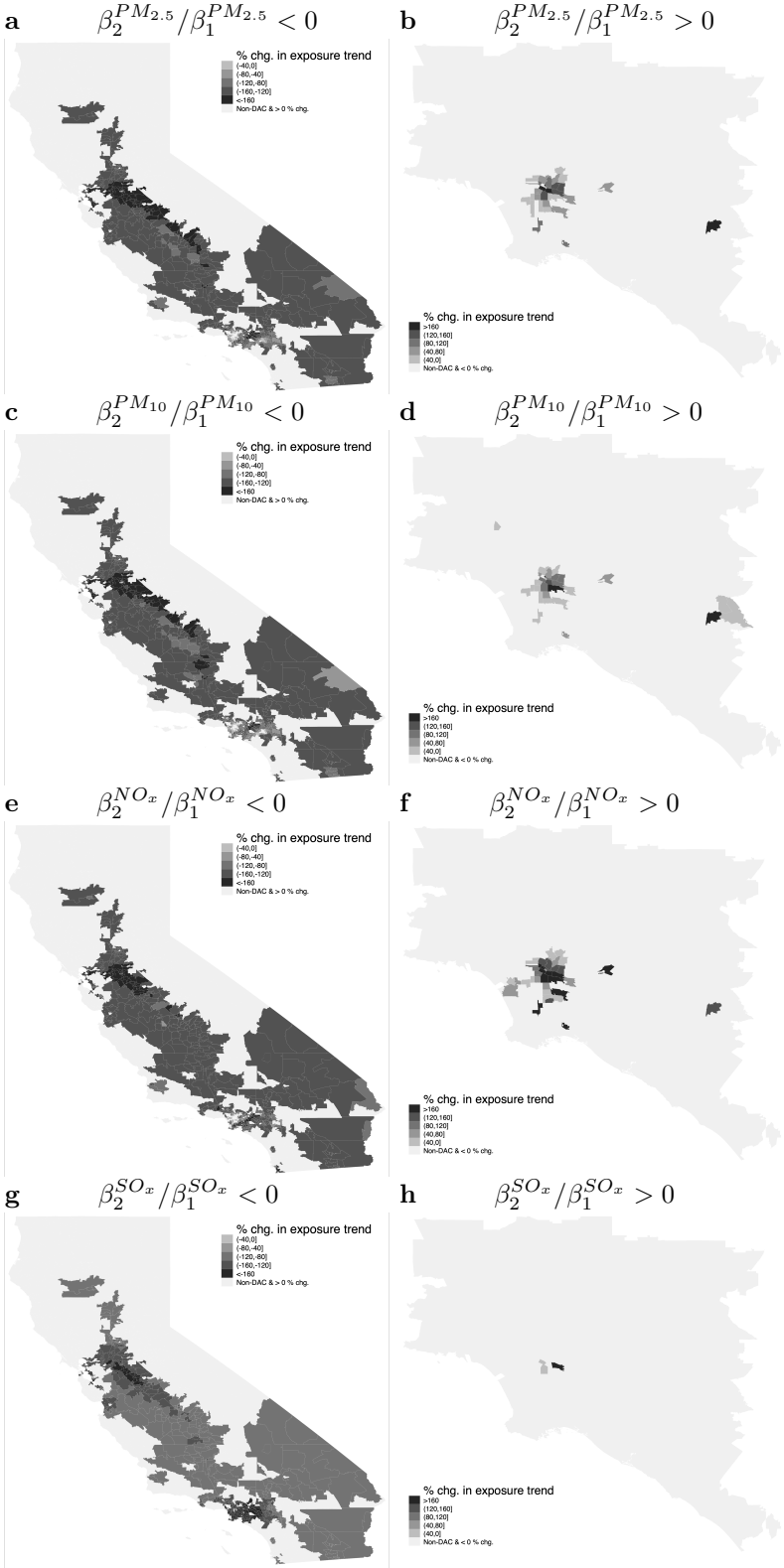
NOTES: Panels show estimated (true and placebo) emissions trend break coefficients (i.e., κ_2 from eq. (1)) for (a) GHG, (b) $PM_{2.5}$, (c) PM_{10}, (d) NO_x, and (e) SO_x emissions from varying the start year of the C&T program. Vertical line at 2013 indicates actual introduction of the program. Shaded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

Figure S6: Empirical distribution of β_1^p and β_2^p from bootstrapping step 1



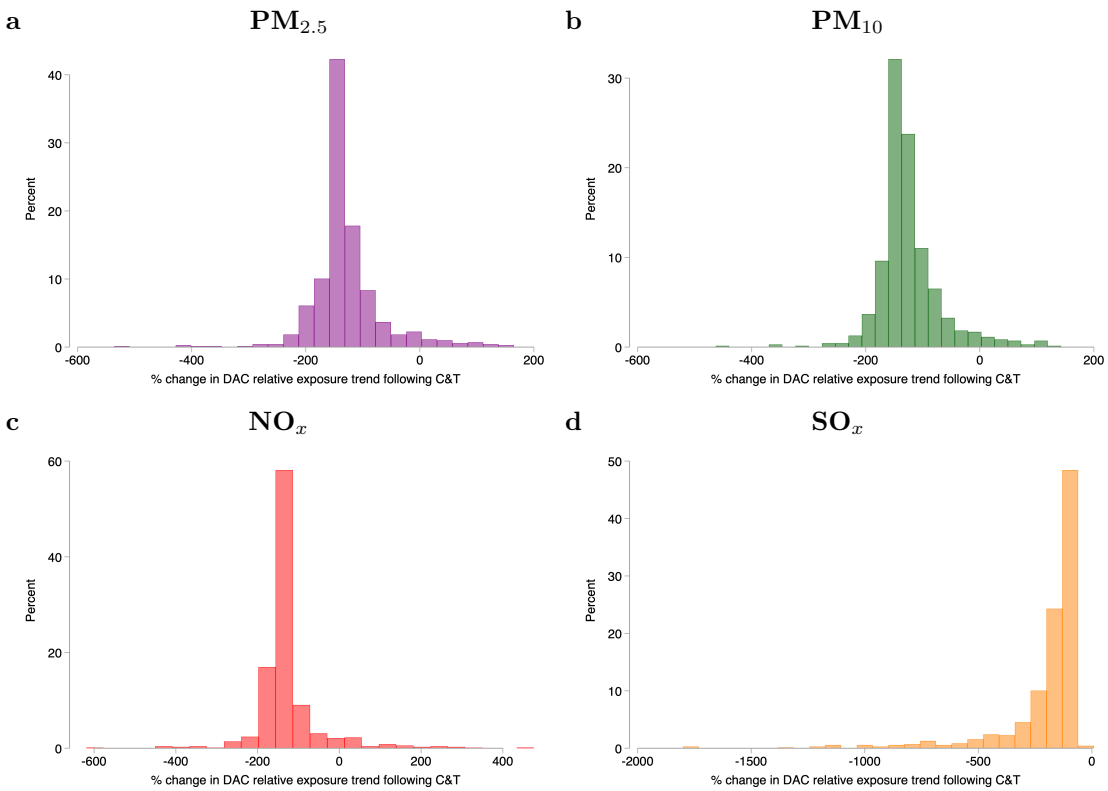
NOTES: Panels show the empirical distribution of β_1^p and β_2^p from equation (1) (across columns) for (a,b) $PM_{2.5}$, (c,d) PM_{10}, (e,f) NO_x, and (g,h) SO_x (across rows) using the bootstrap procedure detailed in Section A.1 with 250 draws. Solid black line shows parameter from directly estimating 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 S7: Spatial heterogeneity in EJ gap effects: grayscale



NOTES: Replication of Figure 5 in grayscale. Panels maps the zip code-specific percentage change in the EJ gap trend ($\beta_2^p/\beta_1^p * 100$) following the introduction of the C&T program for disadvantaged zip codes across (a,b) $PM_{2.5}$, (c,d) PM_{10}, (e,f) NO_x, and (g,h) SO_x. Gray scale in left panel shows zipcodes with negative change with lightest gray shading showing non-disadvantaged and positively changing zipcodes. Gray scale in right panel shows zipcodes with positive change with lightest gray shading showing non-disadvantaged and negatively changing zipcodes.

Figure S8: 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 (a) $PM_{2.5}$, (b) PM_{10}, (c) NO_x, and (d) 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. Observations above the 99.5th percentile and below the 0.5th percentile dropped.

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_{10}	14.48	6.25
$PM_{2.5}$	8.02	3.74
NO_x	53.42	16.03
SO_x	10.86	2.8
Shares by sector:		
Agriculture	0	.018
Manufacturing	.623	.504
Mining, Oil and Gas extraction	.16	.097
Services	.066	.23
Transportation	.066	.053
Utilities	.075	.093
Wholesale	.009	.004

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)	(2)	(3)	(4)	(5)	(6)
	Benchmark	post mean	sector-year FEs	Drop switchers	Baseline CO_2e cutoff (%) 70	80
			Outcome is asinh(CO_2e) emissions			
pre-trend (κ_1^p)	0.187	0.178	0.163	0.223	0.194	0.174
	(0.052)	(0.051)	(0.053)	(0.054)	(0.055)	(0.050)
	[0.001]	[0.001]	[0.004]	[< 0.001]	[0.001]	[0.001]
trend break (κ_2^p)	-0.297	-0.299	-0.257	-0.327	-0.307	-0.260
	(0.077)	(0.079)	(0.082)	(0.081)	(0.085)	(0.072)
	[<0.001]	[<0.001]	[0.003]	[<0.001]	[0.001]	[0.001]
post-mean		0.069				
		(0.124)				
		[0.578]				
Facilities	316	316	316	294	294	337
Observations	2,054	2,054	2,054	1,887	1,863	2,234
			Outcome is asinh($PM_{2.5}$) emissions			
pre-trend (κ_1^p)	0.058	0.056	0.059	0.067	0.071	0.046
	(0.043)	(0.041)	(0.050)	(0.045)	(0.043)	(0.043)
	[0.183]	[0.183]	[0.241]	[0.149]	[0.111]	[0.298]
trend break (κ_2^p)	-0.097	-0.097	-0.095	-0.100	-0.105	-0.079
	(0.048)	(0.049)	(0.056)	(0.051)	(0.050)	(0.050)
	[0.053]	[0.054]	[0.101]	[0.059]	[0.044]	[0.121]
post-mean		0.019				
		(0.067)				
		[0.774]				
Facilities	302	302	302	281	281	323
Observations	1,968	1,968	1,968	1,810	1,780	2,147
			Outcome is asinh(PM_{10}) emissions			
pre-trend (κ_1^p)	0.083	0.078	0.088	0.091	0.097	0.075
	(0.033)	(0.032)	(0.037)	(0.035)	(0.034)	(0.035)
	[0.016]	[0.020]	[0.024]	[0.014]	[0.008]	[0.039]
trend break (κ_2^p)	-0.117	-0.118	-0.115	-0.117	-0.129	-0.103
	(0.040)	(0.040)	(0.045)	(0.043)	(0.043)	(0.042)
	[0.005]	[0.005]	[0.015]	[0.010]	[0.005]	[0.018]
post-mean		0.043				
		(0.064)				
		[0.505]				
Facilities	302	302	302	281	281	323
Observations	1,968	1,968	1,968	1,810	1,780	2,147
			Outcome is asinh(NO_x) emissions			
pre-trend (κ_1^p)	0.075	0.059	0.057	0.060	0.085	0.058
	(0.039)	(0.035)	(0.039)	(0.040)	(0.033)	(0.037)
	[0.061]	[0.106]	[0.154]	[0.143]	[0.015]	[0.128]
trend break (κ_2^p)	-0.104	-0.109	-0.082	-0.083	-0.126	-0.091
	(0.050)	(0.051)	(0.047)	(0.051)	(0.047)	(0.048)
	[0.042]	[0.039]	[0.086]	[0.111]	[0.010]	[0.066]
post-mean		0.131				
		(0.101)				
		[0.204]				
Facilities	303	303	303	282	282	324
Observations	1,970	1,970	1,970	1,812	1,782	2,149
			Outcome is asinh(SO_x) emissions			
pre-trend (κ_1^p)	0.006	0.020	0.010	0.009	-0.005	-0.004
	(0.035)	(0.036)	(0.039)	(0.036)	(0.038)	(0.035)
	[0.875]	[0.580]	[0.791]	[0.803]	[0.890]	[0.912]
trend break (κ_2^p)	-0.037	-0.033	-0.040	-0.043	-0.025	-0.020
	(0.043)	(0.042)	(0.050)	(0.046)	(0.048)	(0.045)
	[0.393]	[0.437]	[0.431]	[0.354]	[0.600]	[0.657]
post-mean		-0.112				
		(0.068)				
		[0.108]				
Facilities	303	303	303	282	282	324
Observations	1,965	1,965	1,965	1,810	1,777	2,142

NOTES: Estimates of pre-C&T differential emissions trend (κ_1^p) and post-C&T differential emissions trend break (κ_2^p) from equation (1) across panels. All models include facility-specific and year-specific dummy variables (except for column 3). Columns 1 shows benchmark model. Column 2 adds a dummy for C&T-regulated facilities after C&T's introduction. post-C&T capturing the post C&T mean shift in differential emissions. Column 3 replaces year fixed effects with sector-by-year fixed effects with sectors shown in Table S1. Column 4 drops facilities that switched C&T regulatory status in 2017. Columns 5 and 6 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(CO_2e) emissions		
pre-trend (κ_1^p)	0.187	0.176	0.172
	(0.052)	(0.052)	(0.052)
	[0.001]	[0.002]	[0.002]
trend break (κ_2^p)	-0.297	-0.361	-0.354
	(0.077)	(0.092)	(0.097)
	[<0.001]	[<0.001]	[0.001]
trend break × avg. emissions		0.000	0.000
		(0.000)	(0.000)
		[0.053]	[0.090]
trend break × avg. emissions^2			-0.000
			(0.000)
			[0.158]
	Outcome is asinh($PM_{2.5}$) emissions		
pre-trend (κ_1^p)	0.058	0.060	0.060
	(0.043)	(0.042)	(0.043)
	[0.183]	[0.167]	[0.165]
trend break (κ_2^p)	-0.097	-0.133	-0.146
	(0.048)	(0.051)	(0.068)
	[0.053]	[0.012]	[0.040]
trend break × avg. emissions		-0.004	-0.005
		(0.003)	(0.004)
		[0.197]	[0.249]
trend break × avg. emissions^2			0.000
			(0.000)
			[0.661]
	Outcome is asinh(PM_{10}) emissions		
pre-trend (κ_1^p)	0.083	0.084	0.086
	(0.033)	(0.033)	(0.033)
	[0.016]	[0.015]	[0.012]
trend break (κ_2^p)	-0.117	-0.143	-0.172
	(0.040)	(0.042)	(0.048)
	[0.005]	[0.002]	[0.001]
trend break × avg. emissions		-0.002	-0.003
		(0.001)	(0.002)
		[0.081]	[0.073]
trend break × avg. emissions^2			0.000
			(0.000)
			[0.197]
	Outcome is asinh(NO_x) emissions		
pre-trend (κ_1^p)	0.075	0.079	0.080
	(0.039)	(0.038)	(0.039)
	[0.061]	[0.046]	[0.046]
trend break (κ_2^p)	-0.104	-0.143	-0.157
	(0.050)	(0.045)	(0.079)
	[0.042]	[0.003]	[0.054]
trend break × avg. emissions		-0.001	-0.001
		(0.000)	(0.001)
		[0.002]	[0.294]
trend break × avg. emissions^2			0.000
			(0.000)
			[0.793]
	Outcome is asinh(SO_x) emissions		
pre-trend (κ_1^p)	0.006	0.013	0.013
	(0.035)	(0.035)	(0.035)
	[0.875]	[0.715]	[0.705]
trend break (κ_2^p)	-0.037	-0.110	-0.074
	(0.043)	(0.048)	(0.077)
	[0.393]	[0.026]	[0.345]
trend break × avg. emissions		-0.004	-0.002
		(0.002)	(0.003)
		[0.017]	[0.455]
trend break × avg. emissions^2			-0.000
			(0.000)
			[0.438]

NOTES: Estimates of pre-C&T differential emissions trend (κ_1^p) and post-C&T differential emissions trend break (κ_2^p) from equation (1) 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) Single facilities	(3) Nonattainment
	Outcome is asinh(CO_2e) emissions		
pre-trend (κ_1^p)	0.187	0.210	
	(0.052)	(0.053)	
	[0.001]	[< 0.001]	
trend break (κ_2^p)	-0.297	-0.322	
	(0.077)	(0.078)	
	[<0.001]	[<0.001]	
Facilities	316	310	
Observations	2,054	2,029	
	Outcome is asinh($PM_{2.5}$) emissions		
pre-trend (κ_1^p)	0.058	0.066	0.085
	(0.043)	(0.043)	(0.049)
	[0.183]	[0.137]	[0.092]
trend break (κ_2^p)	-0.097	-0.101	-0.119
	(0.048)	(0.049)	(0.052)
	[0.053]	[0.046]	[0.029]
Facilities	302	299	260
Observations	1,968	1,952	1,729
	Outcome is asinh(PM_{10}) emissions		
pre-trend (κ_1^p)	0.083	0.091	0.101
	(0.033)	(0.033)	(0.034)
	[0.016]	[0.008]	[0.006]
trend break (κ_2^p)	-0.117	-0.121	-0.145
	(0.040)	(0.040)	(0.054)
	[0.005]	[0.004]	[0.012]
Facilities	302	299	140
Observations	1,968	1,952	1,080
	Outcome is asinh(NO_x) emissions		
pre-trend (κ_1^p)	0.075	0.065	0.057
	(0.039)	(0.039)	(0.041)
	[0.061]	[0.101]	[0.173]
trend break (κ_2^p)	-0.104	-0.098	-0.090
	(0.050)	(0.050)	(0.054)
	[0.042]	[0.060]	[0.102]
Facilities	303	300	287
Observations	1,970	1,954	1,879
	Outcome is asinh(SO_x) emissions		
pre-trend (κ_1^p)	0.006	0.005	
	(0.035)	(0.036)	
	[0.875]	[0.892]	
trend break (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 (κ_1^p) and post-C&T differential emissions trend break (κ_2^p) from equation (1) across panels. Columns 1 shows benchmark model. Column 2 restricts unregulated facilities to those whose parent company only operates a single facility. Column 3 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. 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)
	[<0.001]	[<0.001]	[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 in zip codes with 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	0.093	0.163
	(0.888)	(0.572)	(0.038)
			[<0.001]
PM_{10}	0.322	0.109	0.214
	(1.066)	(0.532)	(0.043)
			[<0.001]
NO_x	0.451	0.387	0.064
	(2.842)	(6.856)	(0.243)
			[0.792]
SO_x	0.364	0.091	0.273
	(1.092)	(0.217)	(0.041)
			[<0.001]
Zip codes	722	984	1,706

NOTES: Column 1 shows average 2008 pollution concentration ($\mu g/m^3/day$) across disadvantaged zip codes, with standard deviation in parentheses. Column 2 shows average 2008 pollution concentration ($\mu g/m^3/day$) 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 HYSPLIT 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-specific effects	(3) Drop switchers	(4) GHG cutoff: 70%	(5) GHG cutoff: 80%	(6) Hetero by emissions	(7) SUTVA Single fac.	(8) SUTVA NA
				Panel a: $PM_{2.5}$				
β_1^p	0.040	0.041	0.040	0.025	0.043	0.041	0.043	0.048
	(0.011)	(0.011)	(0.013)	(0.006)	(0.010)	(0.012)	(0.012)	(0.012)
	[0.001]	[0.001]	[0.003]	[<0.001]	[<0.001]	[0.001]	[<0.001]	[<0.001]
β_2^p	-0.061	-0.061	-0.058	-0.031	-0.063	-0.075	-0.064	-0.067
	(0.019)	(0.019)	(0.021)	(0.008)	(0.019)	(0.021)	(0.020)	(0.021)
	[0.003]	[0.002]	[0.007]	[<0.001]	[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	-146.262	-141.543
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	0.066	0.074
	(0.014)	(0.014)	(0.017)	(0.008)	(0.013)	(0.014)	(0.015)	(0.016)
	[<0.001]	[<0.001]	[0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
β_2^p	-0.089	-0.089	-0.079	-0.046	-0.093	-0.100	-0.091	-0.105
	(0.027)	(0.026)	(0.028)	(0.009)	(0.026)	(0.028)	(0.028)	(0.030)
	[0.002]	[0.001]	[0.007]	[<0.001]	[0.001]	[0.001]	[0.002]	[0.001]
$(\beta_2^p/\beta_1^p) * 100$	-142.447	-140.455	-134.110	-121.310	-134.948	-156.155	-136.905	-141.932
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	0.082	0.079
	(0.033)	(0.037)	(0.032)	(0.026)	(0.031)	(0.033)	(0.034)	(0.032)
	[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	-0.141	-0.138
	(0.066)	(0.070)	(0.069)	(0.031)	(0.069)	(0.075)	(0.071)	(0.070)
	[0.051]	[0.043]	[0.061]	[0.084]	[0.035]	[0.062]	[0.053]	[0.052]
$(\beta_2^p/\beta_1^p) * 100$	-167.212	-158.315	-170.919	-128.157	-170.878	-180.760	-172.408	-175.096
Observations	16,416	16,416	16,413	16,387	16,426	16,416	16,416	16,416
				Panel c: 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 (β_1^p), post-C&T EJ gap trend break (β_2^p), and the percentage change in the EJ gap trend following the introduction of the C&T program ($\frac{\beta_2^p}{\beta_1^p} * 100$) 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) Monthly emissions	(2) Slower decay	(3) Faster decay	(4) Lower boundary	(5) Higher boundary	(6) Spatial corr. err.	(7) Pollution corr. err.
				Panel a: $PM_{2.5}$			
β_1^p	0.043	0.043	0.041	0.037	0.043	0.042	0.042
	(0.011)	(0.011)	(0.011)	(0.010)	(0.011)	(0.004)	(0.006)
	[<0.001]	[<0.001]	[<0.001]	[0.001]	[<0.001]	[<0.001]	[<0.001]
β_2^p	-0.064	-0.064	-0.062	-0.055	-0.064	-0.063	-0.063
	(0.020)	(0.020)	(0.020)	(0.019)	(0.020)	(0.009)	(0.010)
	[0.002]	[0.002]	[0.003]	[0.007]	[0.002]	[<0.001]	[<0.001]
$(\beta_2^p/\beta_1^p) * 100$	-149.824	-149.007	-150.533	-148.764	-149.992	-149.699	-149.699
Observations	16,416	16,416	16,416	16,359	16,430	16,417	16,417
				Panel b: PM_{10}			
β_1^p	0.066	0.066	0.063	0.057	0.066	0.065	0.065
	(0.015)	(0.015)	(0.014)	(0.013)	(0.014)	(0.006)	(0.008)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
β_2^p	-0.092	-0.092	-0.089	-0.079	-0.092	-0.090	-0.090
	(0.028)	(0.027)	(0.027)	(0.027)	(0.027)	(0.011)	(0.013)
	[0.002]	[0.001]	[0.002]	[0.005]	[0.001]	[<0.001]	[<0.001]
$(\beta_2^p/\beta_1^p) * 100$	-139.896	-139.150	-140.448	-137.785	-140.161	-139.739	-139.739
Observations	16,416	16,416	16,416	16,359	16,430	16,417	16,417
				Panel c: NO_x			
β_1^p	0.088	0.089	0.081	0.083	0.085	0.085	0.085
	(0.037)	(0.036)	(0.034)	(0.035)	(0.035)	(0.039)	(0.021)
	[0.019]	[0.018]	[0.020]	[0.020]	[0.018]	[0.030]	[<0.001]
β_2^p	-0.147	-0.148	-0.139	-0.140	-0.144	-0.143	-0.143
	(0.075)	(0.073)	(0.073)	(0.073)	(0.073)	(0.050)	(0.033)
	[0.056]	[0.048]	[0.063]	[0.060]	[0.054]	[0.004]	[<0.001]
$(\beta_2^p/\beta_1^p) * 100$	-167.136	-166.117	-170.804	-168.674	-168.261	-168.282	-168.282
Observations	16,416	16,416	16,416	16,359	16,430	16,417	16,417
				Panel d: SO_x			
β_1^p	0.038	0.037	0.037	0.030	0.038	0.037	0.037
	(0.023)	(0.023)	(0.022)	(0.019)	(0.023)	(0.007)	(0.006)
	[0.104]	[0.109]	[0.107]	[0.133]	[0.103]	[<0.001]	[<0.001]
β_2^p	-0.102	-0.102	-0.100	-0.087	-0.102	-0.101	-0.101
	(0.050)	(0.050)	(0.049)	(0.044)	(0.050)	(0.012)	(0.010)
	[0.047]	[0.047]	[0.047]	[0.053]	[0.045]	[<0.001]	[<0.001]
$(\beta_2^p/\beta_1^p) * 100$	-269.062	-271.967	-272.688	-295.166	-270.107	-272.291	-272.291
Observations	16,416	16,416	16,416	16,359	16,430	16,417	16,417

NOTES: Estimates of the pre-C&T EJ gap trend (β_1^p), post-C&T EJ gap trend break (β_2^p), and the percentage change in the EJ gap trend following the introduction of the C&T program ($\frac{\beta_2^p}{\beta_1^p} * 100$) 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 distributes annual facility-level C&T-driven emissions according to monthly share of annual sectoral output. Column 2 applies a slower pollution decay to HYSPLIT pollution trajectories (i.e., 10% larger half-life parameter). Column 3 applies a faster pollution decay to HYSPLIT pollution trajectories (i.e., 10% smaller half-life parameter). Column 4 applies a lower planetary boundary layer set at 0.5 km. Column 5 applies a higher planetary boundary layer set at 2 km. Column 6 adjusts standard errors for spatial (500 km uniform kernel) and serial correlation (5 years). Column 7 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) $PM_{2.5}$	(2) PM_{10}	(3) NO_x	(4) SO_x
pre-trend (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]
trend break (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]
post-trend ($\beta_1^p + \beta_2^p$)	-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 (β_1^p), the post-C&T EJ gap trend break (β_2^p), and the post-C&T EJ gap trend ($\beta_1^p + \beta_2^p$) 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: census tract-level DAC definition

	(1) $PM_{2.5}$	(2) PM_{10}	(3) NO_x	(4) SO_x
pre-trend (β_1^p)	0.057	0.082	0.152	0.063
	(0.016)	(0.019)	(0.043)	(0.029)
	[0.001]	[< 0.001]	[0.001]	[0.036]
trend break (β_2^p)	-0.083	-0.113	-0.245	-0.131
	(0.026)	(0.037)	(0.104)	(0.047)
	[0.002]	[0.003]	[0.022]	[0.007]
post-trend ($\beta_1^p + \beta_2^p$)	-0.026	-0.031	-0.094	-0.068
	(0.019)	(0.026)	(0.075)	(0.020)
	[0.180]	[0.231]	[0.220]	[0.001]
Trend pct change ($100 * \beta_2^p/\beta_1^p$)	-144.708	-137.866	-161.746	-207.392
	(33.612)	(29.323)	(43.279)	(31.638)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Census tracts	8023	8023	8023	8023
Counties	58	58	58	58
Observations	80,230	80,230	80,230	80,230

NOTES: Data at the census tract-year level. Disadvantaged community defined at the census tract using CalEnviroScreen v3.0. Estimates of the pre-C&T EJ gap trend (β_1^p), the post-C&T EJ gap trend break (β_2^p), the post-C&T EJ gap trend ($\beta_1^p + \beta_2^p$), 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 census tract-specific and year-specific dummy variables. Observations weighted by census tract-level population from the 2010 census. 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 shown in Figure S6, and county-level heterogeneity in EJ gap effects of arbitrary form (ϵ_{it}^p from equation (2)). P-value in brackets.

Table S11: EJ gap effect robustness: minority-share DAC definition

	(1) $PM_{2.5}$	(2) PM_{10}	(3) NO_x	(4) SO_x
pre-trend (β_1^p)	0.032	0.049	0.084	0.029
	(0.013)	(0.016)	(0.034)	(0.022)
	[0.020]	[0.003]	[0.016]	[0.190]
trend break (β_2^p)	-0.049	-0.069	-0.139	-0.073
	(0.022)	(0.028)	(0.070)	(0.043)
	[0.027]	[0.017]	[0.053]	[0.093]
post-trend ($\beta_1^p + \beta_2^p$)	-0.017	-0.020	-0.055	-0.044
	(0.014)	(0.019)	(0.046)	(0.023)
	[0.230]	[0.278]	[0.238]	[0.059]
Trend pct change ($100 * \beta_2^p/\beta_1^p$)	-152.201	-141.658	-165.554	-248.045
	(42.970)	(35.481)	(46.534)	(64.826)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Zip codes	1649	1649	1649	1649
Counties	58	58	58	58
Observations	16,416	16,416	16,416	16,416

NOTES: Disadvantaged community defined at whether a zip code has 2008-2012 average minority share of population above the median across California zip codes. Estimates of the pre-C&T EJ gap trend (β_1^p), the post-C&T EJ gap trend break (β_2^p), the post-C&T EJ gap trend ($\beta_1^p + \beta_2^p$), and the percentage change in the EJ gap trend following the introduction of the C&T program ($\frac{\beta_2^p}{\beta_1^p} * 100$) for $PM_{2.5}$, PM_{10}, NO_x, and SO_x, across columns. All models include census tract-specific and year-specific dummy variables. Observations weighted by census tract-level population from the 2010 census. 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 shown in Figure S6, and county-level heterogeneity in EJ gap effects of arbitrary form (ϵ_{it}^p from equation (2)). P-value in brackets.

Table S12: EJ gap effect robustness: including large emitters, electricity generators, and refineries

	(1) $PM_{2.5}$	(2) PM_{10}	(3) NO_x	(4) SO_x
pre-trend (β_1^p)	0.209	0.223	0.084	0.139
	(0.137)	(0.147)	(0.092)	(0.096)
	[0.132]	[0.134]	[0.366]	[0.154]
trend break (β_2^p)	-0.431	-0.477	-0.348	-0.199
	(0.188)	(0.201)	(0.134)	(0.094)
	[0.026]	[0.021]	[0.012]	[0.039]
post-trend ($\beta_1^p + \beta_2^p$)	-0.222	-0.254	-0.264	-0.060
	(0.059)	(0.063)	(0.062)	(0.021)
	[<0.001]	[<0.001]	[<0.001]	[0.007]
Trend pct change ($100 * \beta_2^p/\beta_1^p$)	-206.402	-213.745	-414.288	-143.527
	(48.897)	(54.661)	(314.152)	(36.385)
	[<0.001]	[<0.001]	[0.187]	[<0.001]
Zip codes	1650	1650	1650	1650
Counties	58	58	58	58
Observations	16,488	16,488	16,488	16,488

NOTES: Sample augments benchmark sample with large emitters (i.e., >75th percentile average annual emissions), electricity generators, and refineries. Estimates of the pre-C&T EJ gap trend (β_1^p), the post-C&T EJ gap trend break (β_2^p), the post-C&T EJ gap trend ($\beta_1^p+\beta_2^p$), 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 county-level heterogeneity in EJ gap effects of arbitrary form (ϵ_{it}^p from equation (2)). P-value in brackets.

Table S13: EJ gap effect robustness: total $PM_{2.5}$ concentration using InMAP

	(1) Primary $PM_{2.5}$	(2) Total $PM_{2.5}$
pre-trend (β_1^p)	0.002	0.003
	(0.001)	(0.001)
	[0.001]	[0.001]
trend break (β_2^p)	-0.003	-0.004
	(0.001)	(0.001)
	[<0.001]	[<0.001]
post-trend ($\beta_1^p + \beta_2^p$)	-0.001	-0.002
	(0.000)	(0.001)
	[0.004]	[0.001]
Trend pct change ($100 * \beta_2^p/\beta_1^p$)	-150.559	-172.948
	(16.261)	(16.415)
	[<0.001]	[<0.001]
Zip codes	1648	1648
Counties	58	58
Observations	16,480	16,480

NOTES: Estimates of the pre-C&T EJ gap trend (β_1^p), the post-C&T EJ gap trend break (β_2^p), the post-C&T EJ gap trend ($\beta_1^p + \beta_2^p$), and the percentage change in the EJ gap trend following the introduction of the C&T program ($\frac{\beta_2^p}{\beta_1^p} * 100$) 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 S14: Importance of modeling pollution dispersal

	(1) Facility zip code	(2) 1.6 km circle	(3) 4 km circle	(4) 10 km circle
		Panel a: $PM_{2.5}$		
pre-trend (β_1^p)	0.052	-0.017	-0.075	-0.153
	(0.036)	(0.026)	(0.040)	(0.079)
	[0.157]	[0.527]	[0.075]	[0.059]
trend break (β_2^p)	-0.076	-0.003	0.067	0.144
	(0.049)	(0.023)	(0.036)	(0.070)
	[0.134]	[0.912]	[0.075]	[0.047]
post-trend ($\beta_1^p + \beta_2^p$)	-0.024	-0.019	-0.008	-0.009
	(0.014)	(0.013)	(0.014)	(0.026)
	[0.089]	[0.162]	[0.601]	[0.736]
		Panel b: PM_{10}		
pre-trend (β_1^p)	0.105	0.020	-0.069	-0.159
	(0.070)	(0.030)	(0.047)	(0.088)
	[0.143]	[0.509]	[0.155]	[0.079]
trend break (β_2^p)	-0.142	-0.049	0.059	0.150
	(0.091)	(0.036)	(0.055)	(0.093)
	[0.132]	[0.177]	[0.294]	[0.114]
post-trend ($\beta_1^p + \beta_2^p$)	-0.036	-0.029	-0.010	-0.008
	(0.021)	(0.015)	(0.019)	(0.034)
	[0.101]	[0.060]	[0.595]	[0.809]
		Panel c: NO_x		
pre-trend (β_1^p)	0.163	-0.120	-0.292	-0.447
	(0.188)	(0.110)	(0.096)	(0.178)
	[0.391]	[0.285]	[0.005]	[0.016]
trend break (β_2^p)	-0.213	0.103	0.311	0.501
	(0.247)	(0.133)	(0.110)	(0.191)
	[0.397]	[0.442]	[0.008]	[0.012]
post-trend ($\beta_1^p + \beta_2^p$)	-0.049	-0.017	0.019	0.054
	(0.060)	(0.046)	(0.047)	(0.077)
	[0.416]	[0.715]	[0.680]	[0.488]
		Panel d: SO_x		
pre-trend (β_1^p)	0.001	-0.156	-0.273	-0.475
	(0.004)	(0.122)	(0.183)	(0.265)
	[0.688]	[0.210]	[0.145]	[0.081]
trend break (β_2^p)	-0.014	-0.007	0.128	0.263
	(0.009)	(0.030)	(0.103)	(0.155)
	[0.125]	[0.813]	[0.223]	[0.098]
post-trend ($\beta_1^p + \beta_2^p$)	-0.012	-0.163	-0.145	-0.213
	(0.009)	(0.097)	(0.082)	(0.115)
	[0.196]	[0.104]	[0.086]	[0.073]
Observations	783	1,823	3,553	7,026

NOTES: Estimates of the pre-C&T EJ gap trend (β_1^p) and the post-C&T EJ gap trend break (β_2^p) 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.