What Caused Racial Disparities in Particulate Exposure to Fall? New Evidence from the Clean Air Act and Satellite-Based Measures of Air Quality[1]

Janet Currie
Princeton University

John Voorheis
U.S. Census Bureau

Reed Walker
UC Berkeley

December 2019

Abstract

Racial differences in exposure to ambient air pollution have declined significantly in the United States over the past 20 years. This project links administrative Census microdata to newly available, spatially continuous high resolution measures of ambient particulate pollution (PM2.5) to examine the underlying causes and consequences of differences in black-white pollution exposures. We begin by decomposing differences in pollution exposure into components explained by observable population characteristics (e.g., income) versus those that remain unexplained. We then use quantile regression methods to show that a significant portion of the "unexplained" convergence in black-white pollution exposure can be attributed to differential impacts of the Clean Air Act (CAA) in African American and non-Hispanic white communities. Areas with larger black populations saw greater CAA-related declines in PM2.5 exposure. We show that the CAA has been the single largest contributor to racial convergence in PM2.5 pollution exposure in the U.S. since 2000 accounting for over 60 percent of the reduction.

[1] This paper is released to inform interested parties of research and to encourage discussion. The views expressed are those of the authors and not necessarily those of the U.S. Census Bureau. All results have been reviewed to ensure that no confidential information is disclosed, release authorization numbers CBDRB-FY19-CMS-7029, CBDRB-FY19-CMS-7227, CBDRB-FY19-CMS-7328 and CBDRB-FY19-CMS-7735. We would like to thank Abhay Aneja, Spencer Banzhaf, David Card, Conrad Miller, Joe Shapiro, and seminar participants at the Chicago Federal Reserve, the Environmental Defense Fund, Gothenberg University, LISER, Rotterdam University, Stanford University, Tufts, UC Berkeley, the University of Chicago, the University of Hawaii, and the University of Illinois for helpful comments. Ellen Lin and Matthew Tarduno provided exceptionally helpful research assistance.

I. Introduction

Landmark studies in the 1980s (see for example U.S. General Accounting Office, 1983; Chavkin and Lee, 1987) demonstrated that low income and/or racial minorities in the U.S. are disproportionately exposed to environmental burdens. This issue had become so politically important by the 1990s that President Clinton issued Executive Order 12898 in 1994, which ordered the U.S. Environmental Protection Agency (EPA) to explicitly study this "environmental justice" issue.[2] However, despite its large volume, the existing evidence about racial disparities in pollution exposure is largely piecemeal and indirect.

The evidence is *piecemeal* because pollution monitoring networks are sparse. For example, fewer than 20 percent of U.S. counties contain a regulatory grade device capable of monitoring small particulates (Fowlie, Rubin, and Walker 2019).[3] The evidence remains somewhat *indirect* because researchers have been forced to use proxies for potential exposure such as distance to a polluting facility.[4] Distance to a facility is a crude substitute for ambient air pollution exposure, both for reasons related to air transport and because mobile sources of pollution are also important contributors to local air quality. Hence, while we know that there are racial differences in the proximity to toxic facilities and hazardous waste sites, it is less clear how these differences translate to differences in measured exposures. Moreover, we know very little about why racial gaps in pollution exposure may have changed over time. In particular, there are two unanswered questions: To what extent are observed gaps shaped by differential movement of people across areas, vs. differential improvement within areas where African-Americans are concentrated? And what role has environmental regulation played, especially the provisions of the Clean Air Act (CAA) that regulate certain airborne pollutants?

This paper addresses these gaps in our knowledge using newly available national data on ambient particulate matter (PM2.5) exposure from 2000 to 2015. Advances in remote sensing technology combined with machine learning prediction tools have allowed researchers to combine data from satellite imagery, pollution monitors, land use characteristics, chemical air transport models to generate fine-grained (1km grid) measures of ambient air pollution levels for the entire United States (Di et al., 2016, von Donkelaar et al 2016). We merge these data to individual survey

[2] Banzhaf, Ma, and Timmins (2019) have an excellent recent review of the economics literature on this subject.

[3] Similarly, Hsiang, Oliva and Walker (2018) point out that out of 3144 counties, only 1289 have monitors for any "criteria" air pollutant (i.e. pollutants regulated under the Clean Air Act) at any point between 1990-2015.

[4] For example, several case studies on residential proximity to polluting industrial facilities find that racial and ethnicity minority groups and/or lower socioeconomic status groups experienced closer average proximity to industrial facilities compared with other groups, and that this pattern persists over time (e.g., Abel and White (2011) who study Seattle, 1990 to 2007; Hipp and Lakon (2010) who study southern California, 1990 to 2000; Pais et al. (2014) who examine a national cohort from 1990 to 2007). There are challenges to drawing causal inferences from this literature ranging from from ecological fallacy (Depro et al. 2015; Hsiang et al. 2019) to problems associated with assuming that people in geographic areas that do not contain hazards are not exposed to pollutants, even when the hazards in question may lie close to geographic boundaries (Banzhaf et al., 2019; Mohai and Saha, 2006; Mohai, Pellow and Roberts, 2009).

responses from restricted versions of the 2000 Census and the 2001-2015 American Community Survey (ACS) at the Census Block level.

The paper proceeds in four parts. We first use these data to document gaps in ambient exposure to PM2.5 between African Americans and non-Hispanic whites and to show how these gaps changed over time from 2000 to 2015. Next, we explore whether these cross-sectional gaps in pollution exposure can be explained by differences in individual and/or neighborhood characteristics, as reported in the Census or ACS. Third, we explore the extent to which changes in relative mobility versus relative improvements in neighborhood air quality have contributed to the changes in pollution gaps in pollution exposure over this time period. Lastly, we use quantile regression methods proposed by Firpo et al. (2009) to explore the extent to which the spatially targeted nature of the Clean Air Act, and associated introduction of the PM2.5 National Ambient Air Quality Standards (NAAQS), has affected different parts of the national pollution distribution and, in turn, the observed black-white pollution gap in the United States.

The analysis confirms that African Americans tend to live in the most polluted areas nationally. However, this black-white gap in mean pollution exposure has closed substantially since the turn of the century. The mean gap in pollution exposure has converged from 1.5ug/m3 in 2000 to only 0.5ug/m3 in 2015. This convergence alone could potentially explain 7% of the improvement in relative life expectancy between blacks/whites over this time period.[5] We then explore the underlying cross-sectional correlates of the observed pollution gaps by leveraging the individual microdata in the Census and ACS. We begin by comparing the unconditional mean gap in pollution exposure between African Americans and non-Hispanic whites to the conditional mean pollution gap after controlling for individual characteristics (e.g., income, education, household structure). We then examine whether individual characteristics are able to explain gaps in exposure at other quantiles of the pollution distribution, in the spirit of Dinardo, Fortin, and Lemieux (1996). We find that virtually none of the racial difference in exposure can be explained by differences in individual or household-level characteristics such as income, suggesting that only a small portion of the observed convergence in pollution levels can be explained by relative changes in these characteristics over time.

Mechanically, there are two remaining ways this narrowing of the pollution gap could have occurred: Areas that were disproportionately African American may have become cleaner faster than other areas; or relative population shares could have shifted in ways that benefited African Americans relative to the non-Hispanic white population. We use a simple decomposition to show that relative mobility differences or changes in black-white population shares are not able to explain the observed convergence in pollution exposure. While the white population is gradually shifting to more urban and hence more polluted areas, this phenomenon explains very little of the

[5] See Section II below for a more complete description of this calculation.

observed convergence when compared to the significant air quality improvements in the average black neighborhood over this time period.

The convergence in the racial gap in pollution exposure between 2000 and 2015 can be entirely accounted for by the fact that areas that were disproportionately African American showed the greatest improvement in air quality over this time period. Why is this the case? We show that much of this improvement is driven by the introduction of the PM2.5 National Ambient Air Quality Standards, which disproportionately improved air quality in newly regulated areas --- areas with higher concentrations of African Americans. More specifically, we use unconditional quantile regression (Firpo et al., 2009) in a difference-in-differences setting to show how the introduction of the PM2.5 NAAQS affected different quantiles of the national pollution distribution. We then combine these quantile regression estimates with the black and white population shares in the respective pollution quantiles to calculate that over 60 percent of the observed convergence in mean PM2.5 differences between blacks and whites over this time period can be traced back to the spatially targeted nature of the CAA regulations.

A small but growing literature has begun using remote sensing data to measure the distribution of environmental hazards (see e.g., Fowlie, Rubin, and Walker, 2019 for a recent overview and application). A subset of this literature studies environmental inequality by merging remote-sensing data to Census tract or Census block group demographic data. For example, Clark, Millet, Marshall (2014, 2017) examine racial differences in exposure to NO_2 using 2006 Census Block-Group NO_2 measurements linked to the 2000 Decennial Census data, and Voorheis (2015) examine differences in exposure to NO_2 and PM2.5 using satellite data combined with tract-level aggregates from the American Community Survey.[6]

This is the first paper, to our knowledge, that links nationally representative, individual-level survey data to a national surface grid of PM2.5 pollution measurements to explore cross-sectional differences and trends in environmental inequality between racial groups. Moreover, we are not aware of any papers that have explored the causal determinants of narrowing pollution gaps between racial groups over time. Our data not only has useful features from a measurement

[6] Relatively few U.S. studies have explored temporal trends in racial gaps in ambient air pollution or transportation-related air pollution. Brajer and Hall (2005), studied ozone and coarse particulate matter in southern California from 1990 to 1999, and found that on average, as air pollution decreased over time, Asians and Hispanics experienced larger reductions in ozone concentrations but smaller reductions in coarse particulate matter concentrations, compared with other groups. Kravitz-Wirtz et al. (2016), studied nitrogen dioxide and particulate matter exposures in the United States for a cohort of 9,000 families from 1990 to 2009, and found that though exposures decreased over time, they remained higher for blacks and Hispanics than for whites. Ard et al. (2015) look at racial trends in exposures to toxic releases from 1995 to 2004, using data on releases of 415 chemicals from 17,604 facilities reporting in the Toxic Release Inventory. They find that potential exposure to toxics declined for all, but that African Americans are still more exposed than Whites or Hispanics, even conditional on income and education. Voorheis (2018) uses administrative records and satellite data to compare within-person changes in pollution exposure between blacks and whites.

perspective, allowing for higher quality measurement of the distribution of pollution exposure than has been possible before, but it also allows us to answer questions that cannot be addressed with public-use Census data. For example, the restricted versions of the Census and ACS data (which provide block-level geographic information) allow us to distinguish between differences in pollution exposure that can be explained by differences in individual characteristics (e.g., income) or differences in neighborhood characteristics (e.g., racial mix). While these findings are descriptive, we are able to explore, for the first time, how much variation in pollution exposure might plausibly be explained by individual endowments, and how much may instead be explained by aggregate, neighborhood-level characteristics. Lastly, the spatially continuous PM2.5 measurements allow us to create these statistics and perform these decompositions for the entire continental U.S., as opposed to focusing on a single city or on communities that are proximate to a toxic plant or a pollution monitor.

The second contribution of this paper is to explore the distributional effects of environmental policy and the Clean Air Act more specifically. While a substantial literature examines the average effects of the Clean Air Act on pollution exposure and the harms that it causes (e.g., Chay and Greenstone, 2003, Isen et al., 2017), we know of no other work that explores the impact of the Clean Air Act on different empirical moments of the nationwide pollution distribution. For example, we are able to address the question of how much the CAA NAAQS have compressed the national pollution distribution, and to ask which quantiles show the largest effects? The advent of unconditional quantile regression and related methods (Firpo et al., 2009), combined with the new availability of spatially continuous PM2.5 measurements, make such analysis possible.

In summary, our study uses high-resolution PM2.5 data and restricted-access Census data to measure gaps in racial exposure to pollution for a nationally representative sample of the U.S. population. Our main innovation is to explore the reasons why the racial gap has declined by asking what part of the decline can be attributed to convergence in the individual and household-level characteristics of African American and other households; how much can be explained by the relative mobility of different racial groups (e.g., blacks moving away from polluted areas); how much can be explained by more rapid clean-up of historically black neighborhoods; and how much of the more rapid clean-up of historically black neighborhoods that we observe is due to the spatially targeted enforcement of the CAA.

The rest of the paper proceeds as follows. Section 2 introduces the linked Census and satellite derived pollution data and provides initial descriptive statistics on environmental inequality between racial groups. Section 3 explores whether these mean and quantile differences can be explained by differences in individual characteristics and changes to those characteristics over time. Section 4 tests for how the spatially targeted nature of the Clean Air Act affects the national distribution of PM2.5 exposure, and how these quantile changes map into the observed black-white gap in pollution exposure. Section 5 concludes.

II. Data and Descriptive Statistics on Racial Gaps in Pollution Exposure

Environmental science has seen a recent explosion of research combining spatially-continuous satellite measurements of pollution correlates (e.g., aerosol optical depth) with other observable pollution correlates such as emissions inventories, chemical transport models, land use characteristics, and weather patterns (see e.g., Di et al. (2016), von Donkelaar et al (2016)). The basic idea is to build a predictive model of a pollutant of interest (e.g., PM2.5) by correlating in-situ EPA monitor data with the observable predictors of air pollution measures using modern model selection techniques (e.g., cross-validated neural nets). Researchers then use these models to predict air pollution "out of sample" for the large segments of the United States without existing pollution monitors but with satellite measurements. We use data from Di et al. (2016), who produce daily PM2.5 concentrations at a 0.01 degree by 0.01 degree resolution (1km by 1km at the equator) for the contiguous U.S. from 2000-2015.[7] We spatially intersect this gridded, raster data with Census block boundary files from the 2000 and 2010 Census, which we use to merge the pollution data to the individual demographic data in the 2000 Census and 2001-2015 ACS.

It is important to note that these data are estimates of ground-level pollution concentrations. These estimates perform well – on average, they match the "ground truth" as measured by EPA monitors, with very high, in-sample measures of fit. However, there is some evidence that these satellite-derived measures may deviate from the ground truth at the tails of the pollution exposure distribution. Fowlie, Rubin, and Walker (2019) show that two of the most commonly used measures, Di et al. (2016) and von Donkelaar et al (2016), are biased downward for high PM2.5 levels. It seems likely in our setting that these prediction errors will attenuate measured disparities since African Americans are disproportionately located in the most highly polluted places. We will proceed with our analyses treating the satellite data as if it were the truth, keeping in mind the caveat that our results may be a lower bound on true racial gaps in exposure.

We merge the pollution data with individual-level data from the 2000 Census long form (1 in 6 U.S. households) and from the 2001-2015 American Community Surveys using household Census block locations. Our primary comparisons focus on the non-Hispanic white and African American populations. We focus on gaps between African Americans and non-Hispanic whites because historically these have been the largest and most well-documented gaps. Additionally, there are potential measurement issues in studying differences between Hispanics and non-Hispanic whites over time. For example, studies which have linked individual responses to the 2000 and 2010 Decennial Censuses find evidence that Hispanic identity is more fluid over time than white or black racial identities (Liebler, et al, 2017).

[7] We have replicated most of the results in this project using similar data from von Donkelaar et al. (2016), and the qualitative conclusions are very similar.

Figure 1 plots mean pollution exposure for both the African American (dotted red line) and non-Hispanic white population (solid blue line) from 2000 to 2015. The observed black-white gap in mean pollution exposure is 1.6 micrograms per cubic meter ($\mu g/m^3$) in 2000, narrowing to 0.54 $\mu g/m^3$ in 2015. One way to interpret these differences is to translate the gaps into racial differences in life-expectancy through the lens of a PM2.5 concentration-mortality response function. Ebenstein et al. (2017) estimate that life expectancy is reduced by 0.98 years for each 10 $\mu g/m^3$ increase in sustained exposure to PM2.5. Over this time period, the black-white gap in life expectancy fell from about 5 years to 3.5 years (Arias and Xu, 2019). Thus, the observed 1 $\mu g/m^3$ improvement in the black-white pollution gap could potentially explain 7% of this improvement in the black-white mortality gap.[8]

While Figure 1 presents the mean black-white gap over time, the data also allow us to explore the entire distribution of pollution exposure for each race group and how these distributions have changed over time. Figure 2 plots the pollution densities, separately for the African American and non-Hispanic white population in both 2000 and 2015.[9] When comparing the x-axis of both the 2000 and 2015 figures, it becomes clear that PM2.5 pollution levels have fallen dramatically for both groups. For example, in 2000, the bulk of both distributions lies above 10 $\mu g/m^3$, while in 2015 the bulk of both distributions lies below that threshold. The other salient feature of these densities is that most of the improvements for blacks relative to whites between 2000 and 2015 come from compressing the upper portion of the pollution distribution, a point to which we will return.

III. Decomposing Differences in Pollution Exposure

What explains the observed differences in pollution exposure? Perhaps lower income individuals live in more polluted places, and differences disappear or are greatly attenuated when we condition on income. Or perhaps differences in education and/or knowledge about the potential harms of PM2.5 exposure lead to observed differences in pollution exposure? The Census demographic data allow us to explore the cross-sectional and time-series differences in pollution exposure between racial groups. The demographic data also allow us to begin to explore the extent to which individual endowments can explain the observed differences and changes. Conditional on income differences, does the observed black-white gap in pollution exposure remain? We focus on the following individual-level characteristics measured in the Census long-form sand ACS surveys: race, age, gender, income, education, number of children, and home ownership. We also explore the role of neighborhood characteristics in explaining pollution disparities, using census tract

[8] This is calculated as 0.98*(1/10)/1.5 years. Note that the Arias and Xu (2019) only reports black-white gaps in life expectancy as far back as 2006. The 2000-2015 change in the black-white life expectancy was likely larger than 1.5, reducing the percentage contribution of PM2.5 in explaining this improvement.

[9] Due to Census disclosure avoidance review, we were forced to trim the upper 97th and lower 3rd percentiles of each pollution distribution.

characteristics constructed from the underlying microdata, including: the population share of African Americans in a tract, mean public assistance income, the teen pregnancy rate, average years of schooling, the population share living in single family residences, and the home ownership rate. Appendix Table 1 presents summary statistics of these variables, separately for non-Hispanic white and African American populations in our sample.

Conditional versus Unconditional Differences in Pollution Exposure

The data show a large gap in pollution exposure in 2000, which fell considerably over the next fifteen years. As a first step towards understanding the reasons for the declining gap, we start by asking whether cross-sectional differences in exposure in any given year can be explained by differences in individual characteristics from the Census and ACS. One natural way to do this is to compare the conditional mean differences in pollution exposure between racial groups to the unconditional mean differences using the following linear regression model:

$$P_i = \gamma 1[AfricanAmerican] + X'\beta + \epsilon_i, \quad (1)$$

where pollution for individual *i* is regressed on an indicator for whether or not the individual is African American, controlling for individual demographic characteristics in the vector (*X*). Specifically, we control for individual income, age, education, number of children, gender, and an indicator for homeownership. Regressions are weighted using survey weights, and inference is conducted with cluster-robust standard errors, clustering by commuting zone. Figure 3a plots the coefficient $\hat{\gamma}$ and associated confidence intervals from 15 separate regressions, one per year, representing the conditional mean differences in black-white pollution exposure over time. Figure 3b compares these estimates to the unconditional mean differences in pollution exposure over time. The lines look almost identical. This similarity implies that differences in individual characteristics, such as income, explain almost none of the cross-sectional differences in pollution exposure between African Americans and non-Hispanic whites. This conclusion is striking considering that African American households have mean household income more than $15,000 lower than non-Hispanic whites (see Appendix Table A1). These differences in income and other characteristics explain almost none of the observed differences in pollution exposure. Section A of the Appendix formally decomposes these cross-sectional differences in mean pollution exposure using methods pioneered by Oaxaca (1973) and Blinder (1973). Observable differences in individual and household characteristics are able to explain at most 8 percent of the gap in mean differences in any given year.

Gaps in pollution exposure are also found at different points within the income distribution. Table 1 presents the results from 10 separate estimates of $\hat{\gamma}$ from equation (1), where we stratify the data by income quintiles in both 2000 and 2015. At each quintile, the conditional gap remains similar

to the unconditional gap in a given year. Differences in individual characteristics do little to explain the mean differences in pollution exposure within these income categories.

Equation (1) and the Oaxaca-Blinder exercise in Appendix A are well-suited for decomposing differences in mean black-white pollution exposure. However, Figure 2 suggests there exist additional racial differences at different quantiles of the pollution distribution. A natural question to ask is whether differences in individual or household characteristics are able to explain differences in pollution exposure at other parts of the pollution distribution. Dinardo, Fortin, and Lemieux (1996) provide a straightforward semi-parametric approach to answering this question. The basic idea is to estimate what the entire *distribution* of African American pollution exposure would look like if African Americans had the same observable characteristics as non-Hispanic whites. In practice, this amounts to constructing a weighting function that is then used to re-weight the African American kernel density to provide the relevant counterfactual density. This weighting function boils down to estimating a conditional probability of being a non-Hispanic white individual based on observable, demographic characteristics via a probit regression. This conditional probability is then used, along with the unconditional probabilities of being white/black in our sample to reweight the African American kernel density in a given year.

The results from this exercise are presented in Figure 4. Panel A of Figure 4 presents estimates of the counterfactual pollution distribution in 2000, and Panel B presents estimates of the counterfactual pollution distribution in 2015. In both cases, the counterfactual density looks very similar to the actual density, again suggesting that individual characteristics are able to explain little of the observed pollution gap throughout the distribution. Appendix A and Appendix Table A3 use methods originally proposed by Firpo, Fortin, and Lemieux (2009) to formally decompose the 10^{th}, 50^{th}, and 90^{th} percentiles of the pollution distribution into the parts that can be explained by observable covariates versus the part that remains unexplained. Once again, we find that little can be explained by observable individual-level characteristics from Census data.

Controlling for neighborhood characteristics

While individual and/or household characteristics explain little of the black-white pollution gap, there is substantial racial segregation into different areas within and across communities. Are the differences in characteristics of black and white neighborhoods able to explain these gaps in exposure? Panel 2 of Appendix Table A1 presents mean census tract characteristics, separately for non-Hispanic whites and African Americans. African Americans not only have different individual and household-level characteristics on average, but they also tend to be concentrated in census tracts with a high percentage of other African Americans and with many relatively disadvantaged neighbors. Note that there is nothing mechanical about this—if neighborhoods were perfectly integrated in terms of race and socioeconomic status, then everyone would live in a similar census tract regardless of their own characteristics. Hence, we can control for characteristics of

neighborhoods or census tracts as well as characteristics of individuals in equation (1) in order to explore whether black-white differences in neighborhood characteristics (conditional on individual characteristics) are able to explain some of the observed differences in PM2.5 exposure.

Figure 5 and Appendix Table A4 show the results of adding neighborhood characteristics to equation (1) and the associated Oaxaca-Blinder decomposition, respectively. Specifically, we add census-tract level measures of percent African American, mean public assistance income, the teen pregnancy rate, years of schooling, the share living in single family residences, and the home ownership rate to equation (1). Figure 4, which plots year by year estimates of $\hat{\gamma}$, looks substantially different than Figure 3a. Namely, the conditional gap in pollution exposure between African Americans and non-Hispanic whites is entirely explained by differences in neighborhood characteristics. Appendix A4 explores this result in greater detail. The table shows the fraction of the gap that is "explained" and "unexplained" by each variable separately, as well as the combined effect of these neighborhood characteristics. As one can see from the "Explained" panel of Table 3, black-white differences in neighborhood characteristics explain the entire gap in pollution exposure and more. More specifically, the mean difference in the fraction of a black vs. a white person's Census tract that is African American explains 140% of the overall PM2.5 pollution gap in 2000. In 2015, the mean difference in this single neighborhood covariate explained 135% of the observed pollution gap. There are also substantial differences in the returns or responsiveness of individuals to differences in neighborhood characteristics (i.e. Panel B); a one-year increase in Census tract average education levels translates into substantially less pollution exposure for whites than it does for blacks in our sample, which is reflected in the relatively large difference in estimated "slopes" on the "Tract Years of Schooling" variable in Panel B.

The fact that African Americans live in different neighborhoods than non-Hispanic whites explains virtually all of the gap in pollution exposure but also raises further questions. Does the shrinking of the racial gap in pollution exposure come from disproportionate improvements in air quality in neighborhoods that have higher proportions of African Americans, or is the improvement due to the relative movement of African Americans away from dirty neighborhoods towards cleaner ones? We next present an additional decomposition which attempts to shed light on these questions.

People versus Places: Understanding How Relative Mobility Has Affected Pollution Disparities

One straightforward way to understand the role of mobility in contributing to the convergence in pollution exposure between blacks and whites is to consider what the pollution gap in 2015 would have been if we fixed individuals in their 2000 locations but let pollution in their respective Census

block change to its 2015 level.[10] Table 2 presents results from this type of counterfactual exercise. Columns (1) and (2) present the actual pollution levels experienced by blacks and whites in 2000 and 2015, and the associated gap (row 3) and change in gap (row 4). Column (3) of Table 2 presents the results from a counterfactual exercise, where we simulate what the gap (and change in gap) would have been in 2015 if people's locations had been held fixed. The last row of Column (3) tells us that the change in the pollution gap would have been *smaller* if individuals were unable to move between 2000 and 2015. Relative differences in mobility between blacks and whites over this time period contributed to *additional* convergence in the pollution gap. African Americans moved to relatively cleaner places between 200 and 2015, and the opposite is true for non-Hispanic Whites. This pattern is consistent with recent findings in urban economics which show that, after decades of suburbanization, the urban population became whiter and more college-educated in most large U.S. cities after 2000 (Baum-Snow and Hartley 2019, Couture and Handbury 2019), and similarly, suburbs have become more diverse. The shift in white population shares towards urban centers contributes to higher levels of average pollution exposure among whites than they would have experienced had they remained in predominantly suburban locations.

One can decompose the overall shrinkage in the black white exposure gap into the part which can be explained by mobility versus the part which remains unexplained. If populations were fixed in their 2000 locations, the gain would have been 0.89 μg/m^3 versus 1.02 μg/m^3. Thus, only 12.7% of the improvement in the exposure gap stems from relative mobility differences over this time period, leaving the rest to be explained by a greater improvement in pollution in disproportionately African American areas.

Appendix A presents additional results that explore the year-by-year changes in black-white population shares and pollution exposure in ways that shed further light on the role of mobility in explaining pollution gaps. For example, the once strong negative assortative relationship between white population shares and pollution levels has weakened over time, while the opposite has happened for blacks. Even so, the role of relative mobility differences remains small (certainly less than 15%).

There is a separate question, which we explore in more detail in subsequent sections, as to whether relative mobility is related to *changes* in air quality over this time period. Are non-Hispanic whites moving into the set of cities that have seen the largest improvements in air quality over this time period? It's possible that non-Hispanic whites are moving into areas where air quality has dramatically improved, while also increasing their relative pollution exposure. Consider the case where an individual moves from the relatively less polluted suburbs towards the city center where

[10] For this counterfactual exercise, we use the public-use 100% count population data from the 2000 Decennial Census at the Census block level. We merge these data to the 2000 and 2015 Di et al. (2016) data to calculate mean exposure gaps using block level population counts for non-Hispanic whites and African Americans.

air quality has improved, but air quality still remains above what that individual would have experienced in the suburbs.

IV. The Clean Air Act and Relative Changes in Pollution Exposure

Thus far, we have documented that a) the black-white gap in exposure to particulate matter has declined substantially since 2000, and b) neither individual characteristics nor residential mobility of individuals appears to fully explain why this gap has declined. Rather, it appears that the black-white gap in exposure has declined primarily because African American neighborhoods had greater improvements in air quality. But why did these particulate neighborhoods experience improvements in air quality? We hypothesize that a major reason for the narrowing of the black-white gap in air quality is that changes in the enforcement of the Clean Air Act were more binding in predominantly black neighborhoods. We formally investigate this hypothesis below.

The CAA was first implemented in 1963, but the original legislation provided limited federal oversight of state efforts and led to disappointing results. In response, Congress enacted the Clean Air Act Amendments of 1970 and established the EPA, which vastly increased federal power to address air pollution.[11] The CAA initially focused on common, dangerous, air pollutants known as "criteria air pollutants." Compliance was to be achieved through regulations governing both stationary sources (e.g., factories) and mobile sources (e.g., cars).

For stationary sources, the CAA created pollutant-specific national ambient air quality standards (NAAQS) that specify maximum allowable concentrations of criterion air pollutants. NAAQS were initially established for sulfur dioxide (SO_2), carbon monoxide (CO), nitrogen dioxide (NO_2), lead, particulates (TSP), and eventually ozone (O_3). Stationary source regulations focus primarily on areas that are out of compliance with the NAAQS. Each year in July, the EPA determines the set of counties that are in violation or "nonattainment" of a particular NAAQS standard based on air pollution monitor measurements in those or nearby counties. The consequences of being a "nonattainment" area can be severe. State governments must develop a pollutant-specific plan, known as a State Implementation Plan, describing how areas will improve air quality and come into compliance. The EPA reviews these state plans. If a state fails to act or develops an inadequate plan, the EPA can withhold federal funding for the state air pollution control program, highway construction, and the construction of sewage treatment plants. The EPA can also ban permits for construction of major new and/or modified sources of a pollutant in communities that are out of compliance with NAAQS. In addition, the EPA can impose its own federal plan if it deems a state's plan inadequate. Thus, the CAA gives the EPA sweeping powers to take action to improve air quality, with or without state cooperation.

[11] Currie and Walker (2019) provide a more complete overview of the Clean Air Act and associated research by economists.

Since 1970, there have been two major amendments to the CAA, in 1977 and 1990, and hundreds of additional policies designed to respond to changing scientific consensus about both the harms from pollution and feasible technologies. In 1997 the EPA tightened the NAAQS pertaining to ozone and particles further, regulating fine particles less than 2.5 micrometers in diameter ($PM_{2.5}$) for the first time. The new standards were extremely controversial and were challenged in the courts for years, but ultimately the EPA prevailed, and the new standards were implemented in April 2005. The EPA revised the PM2.5 (24 hour) standard again in 2006, and the revision went into effect in 2009. However, since all counties which were in nonattainment of the annual PM2.5 standard in 2009 were also in nonattainment of the 24 hour standard, the 2009 designations do not appear to have caused additional new areas to be subject to NAAQS nonattainment regulations. We focus on the effects of mandated reductions in annual $PM_{2.5}$ under the 1997 standards which finally began to be implemented in 2005.

To what extent is the cleanup of predominantly African American areas a result of the implementation of these new standards? We have seen that African Americans lived in more polluted places on average in 2000, and the initial impact of the CAA was therefore likely to be greatest in these places. Figure 6a shows the distribution of the African American and non-Hispanic white populations by vigintile of PM2.5 levels in 2000; the former were much more likely to live in the highest pollution deciles, while for whites the situation is reversed. Figure 6b shows that it was predominantly counties in the highest PM2.5 deciles that were impacted by the enforcement of the standards in 2005. Hence, unless the initial impact of stronger regulation was undone by re-sorting of population shares (which we have shown above was not the case), we would expect the regulations to close gaps between blacks and whites.

We begin by examining the effect of these new air quality standards on pollution exposure in a standard difference-in-difference, event-study design. To better mimic the way the regulations work in practice we adopt a slightly unconventional regression model when compared to the existing literature. Whenever a county exceeds the air quality standard based on a local monitoring station, the regulator decides whether nearby or adjacent counties could also have contributed to this violation. Thus, these nonattainment designations apply to "air regions" or groups of counties in the same local market, typically not a single county. We approximate these air region definitions using county-aggregates in the form of commuting zones (CZs) or local labor markets.[12] Figure 7 shows that, once the standards went into effect in 2005, 62 commuting zones, consisting of 250 counties in 20 states, were designated as nonattainment areas.

We estimate an event study model of the following form:

[12] There is no formal EPA definition of "air regions", as they are decided on a case-by-case basis. Thus, air regions are only defined for nonattainment areas. Our use of commuting zone boundaries is meant to approximate air region boundaries in the attainment areas. In practice, our results are not sensitive to using either county or commuting zone boundaries to assign treatment/control.

$$(2) \qquad P_{ict} = \sum_{i=2000}^{2015} \beta_t (1[\text{Nonattain}_c] \times 1[\text{year} = t]) + \gamma_c + \rho_t + \epsilon_{ict},$$

where pollution P_{ict} for person i residing in CZ c in year t is regressed on a series of interaction terms for whether a CZ is newly designated as nonattainment for the PM2.5 standard ($1[\text{Nonattain}_c] = 1$) interacted with a dummy for each year before and after the regulations went into place. Equation (2) also controls for county fixed effects γ_c and (state-)year fixed effects ρ_t. All regressions are weighted using Census survey weights, and standard errors are clustered at the CZ level.

The coefficients of interest, β_t, compare the CZs that became newly regulated under the PM2.5 standard to areas that were in compliance with the standard, before and after the regulations went into place. The identifying assumption is that the newly regulated CZs would have trended similarly, in terms of pollution levels, to the unregulated CZs in the absence of the treatment. Alternatively, the identifying assumption relies on there being no common shock to the treated CZs in the years after the regulations went into place. While these assumptions are inherently untestable, the event study design affords us a useful indirect test; we can examine the event study coefficients and trends leading up to the regulation in pollution levels between the regulated and unregulated CZs and test whether the two groups were trending similarly in the years prior to the regulation going into place.

We also estimate a version of (2) with an additional, three-way interaction term between indicators for ever being in nonattainment status, year, and whether the individual is African American.[13] This "triple-difference" regression equation allows the impact of the CAA to vary differentially by race (i.e. do black neighborhoods clean up more or less in newly regulated CZs than white neighborhoods)?

Lastly, we explore how the CAA regulatory variation affects different quantiles of the pollution distribution. The basic idea is to combine our difference-in-difference estimator with quantile regression methods to estimate the counterfactual outcome distribution in the absence of the policy intervention. We can then compare this distribution to the actual outcome distribution when subject to the policy intervention. We can go further to calculate mean black-white gap in pollution exposure in this counterfactual world, to better understand the role of the CAA in contributing to the convergence of black-white pollution levels.

[13] Formally, we estimate the following equation

$$P_{ict} = \sum_{i=2000}^{2015} \beta_t (1[\text{Nonattain}_c] \times 1[\text{year} = t] \times 1[African\ American_i]) + \Gamma_{ict} + \gamma_c + \rho_t + X'\eta + \epsilon_{ict}$$

where the vector Γ_{ict} includes the full set of two-way interaction terms to facilitate interpretation.

Recent advances in quantile regression allow us to estimate the causal effect of the Clean Air Act's PM2.5 National Ambient Air Quality Standard on the unconditional pollution distribution (Firpo et al. 2009). The basic idea is to transform the problem by considering a covariate's influence on population shares rather than quantiles. By estimating how a covariate (e.g., nonattainment status) affects the share of the population below various pollution thresholds, the semi-elasticities show the effect of an increase in CAA regulatory stringency on the cumulative distribution function (CDF) of pollution (see e.g., Chernozhukov et al. 2013). We can then invert the impact of nonattainment on the CDF of pollution to estimate the impact on a pollution quantile. The re-centered influence function (RIF) regression approach proposed by Firpo et al. (2009) performs this inversion using a local linear approximation to the counterfactual CDF, rescaling the marginal effect of the CAA on the population share above a pollution cutoff by the probability density of pollution at that cutoff.

In practice, RIF regression requires first transforming the outcome variable, PM2.5 pollution, using a re-centered influence function before projecting this transformation onto the explanatory variables of interest. Firpo et al. (2009) consider the following model of pollution *P*:

$$P = h(X, \epsilon)$$

where X represents the set of independent, explanatory variables and ϵ is the scalar unobserved error term. The unconditional partial effect is defined as the shift in the distribution of a continuous variable X on the distributional statistic $v(F_p)$, which can be expressed as

$$\alpha(v) = \int \frac{dE[\mathrm{RIF}(P, v)|X = x]}{dx} dF(x)$$

where $\mathrm{RIF}(P, v)$ is the recentered influence function. When the distributional statistic v is the τth quantile function $q_t = \inf_q\{q : F_P(q) \geq \tau\}$ the $RIF(P, q_\tau)$ can be represented as

$$\mathrm{RIF}(P, q_\tau) = q_\tau + \frac{\tau - 1\{p \leq q_\tau\}}{f_P(q_\tau)}$$

where $f_P(q_\tau)$ is the density function of pollution P evaluated at quantile q_τ.

The relevant property of a re-centered influence function is that its expectation equals the distributional statistic of interest. For quantile τ denoted Q_τ, the quantile RIF is given by $\mathrm{RIF}(p, Q_\tau) = Q_\tau + \frac{\tau - \mathbf{1}(p < Q_\tau)}{f_p(Q_\tau)}$ and taking expectations verifies $E[RIF(p, Q_\tau)] = Q_\tau$. Since the mean of the RIF is equal to the quantile, we can use the law of iterated expectations to go between conditional and unconditional partial effects. Firpo et al. (2009) show that a regression of the RIF

on covariates yields the approximate effect of the covariates on the distributional statistic of interest (applied to the unconditional distribution). This feature of RIF regressions allows researchers to estimate how treatment effects (e.g., the effects of the CAA PM2.5 regulations on county-year pollution levels) map into the unconditional distribution of pollution.

In practice, this entails first defining a series of pollution cutoffs corresponding to specified quantiles of the empirical pollution distribution, and then for each cutoff, estimating the effect of the Clean Air Act's PM2.5 NAAQS policy on the probability of being above that cutoff. We begin by creating 19 RIF statistics, one for each pollution vigintile from the 5th to the 95th percentiles of the pollution distribution. We then replace the dependent variable in equation (2) with one of the RIF-quantile statistics. The event study coefficients for each RIF-quantile can be interpreted as the effect of the PM2.5 nonattainment designations on the qth quantile of the unconditional PM2.5 exposure distribution.

Figure 8 presents the estimated event-study coefficients $\hat{\beta}_t$ from a version of equation (2) where the dependent variable is the pollution level for an individual survey respondent in a given year. There are two main findings. First, in the years leading up to the implementation of the regulation, the trends in air quality between the newly regulated counties and the unregulated counties are not statistically different from zero. After 2005, when the policy was first enacted, pollution levels in newly regulated counties fall by about 8% and remain there through the end of our sample.

Table 3 presents the difference-in-difference analog to Figure 8 and equation (2). In particular, we replace the $1[\text{year} = t]$ variable in equation (2) with a single indicator equal to one for year>=2005:

$$(3) \qquad P_{ict} = \beta(1[\text{Nonattain}_c] \times 1[\text{year}_t \geq 2005]) + \gamma_c + \rho_t + X'\eta + \epsilon_{ict},$$

The regression coefficient of interest β from this more parsimonious model tells us the average difference in pollution levels in the years after, relative to before, comparing regulated counties to counties not subject to the regulation. Column (1) suggests that PM2.5 levels fell by 1.23 μ/m^3 in nonattainment counties in the years after the regulation went into place. Column (3) presents the same model except the dependent variable has been transformed via the natural log. As suggested by the figures, these estimates suggest that pollution levels in newly regulated counties improved by about 8 percent in the years after the policy went into place.

The even columns of Table 3 add a triple interaction between the county ever being in non-attainment status, a post-2005 time period, and an indicator for whether the individual is African American. Recall that African Americans and non-Hispanic whites tend to live in different parts of the same counties, so there is scope for regulation at the county-level to have a differential impact on different communities within a county. These estimates suggest that the within-county

improvements in air quality were slightly less for African Americans than for non-Hispanic whites, though the difference is not statistically significant in the log specification.

Columns (5) through (8) of Table 3 add state-by-year fixed effects to equation (3) to control for any unobserved, statewide changes to air quality that are common to all CZs in a given year. These controls substantially attenuate the coefficients from columns (1) through (4) by almost 50 percent. There are potentially many reasons for this attenuation, but a leading explanation stems from the late-2000's Great Recession which would have led to improvements in air quality as the scale of output (and PM2.5 precursors) declined substantially at the peak of the Great Recession. From this point forward, we treat the estimates with state-by-year fixed effects as our preferred specification in order to flexibly control for this potential form of omitted variable bias.

While Table 3 and equation (3) tells us about the average effects in the treated counties, relative to the controls, it tells us little about other parts of the pollution distribution that might otherwise be affected by this increased regulatory stringency. To explore the distributional impacts of the PM2.5 NAAQS we next turn to our unconditional quantile regression estimates. Figure 9 presents nineteen separate regression estimates of β from equation (3), where we replace the dependent variable in this model with the RIF counterpart for each respective pollution quantile. The figure visually summarizes the effect of the PM2.5 nonattainment designation on different percentiles of the unconditional PM2.5 pollution distribution. The estimates suggest that the most significant effects of the new standards were to improve air quality in areas between the 50^{th} and 90^{th} percentiles of PM2.5 distribution. This result must be true almost by construction, as the PM2.5 standards only bind at the upper quantiles of the pollution distribution. It is somewhat surprising to see that the estimated effect of the regulations are somewhat smaller at the 95^{th} percentile of the pollution distribution than at the 50^{th}-80^{th} percentiles. This smaller effect may be partly due to the fact that the EPA and local regulators have faced severe difficulties in addressing air quality problems in the most severely polluted parts of the country (e.g., the San Joaquin Valley or parts of Southern California).

Figure 11 presents RIF-Quantile regression results that estimate each quantile specific RIF separately by race. This results in 38 versions of equation (3), and the corresponding estimates are plotted in the figure. While the estimates from Table 3 suggested there was little difference in the treatment effects of nonattainment between African Americans and non-Hispanic whites on average (see e.g., column (6)), the results in Figure 11 suggest that at the upper quantiles of the pollution distribution, African Americans are seeing larger improvements in air quality relative to their non-Hispanic white counterparts. For visual clarity, we have omitted the associated confidence intervals, but they are shown in Appendix Figure A2. For most quantiles, the race-specific confidence intervals overlap with one another, and thus it is difficult to reject the null that the treatment effects at each quantile are equal. That being said, it is unlikely that sampling

variability alone could explain the fact that African Americans have larger treatment effects at every percentile above the 60th percentile of the pollution distribution.

What Fraction of Black-White Convergence is Attributable to the CAA Regulation of PM2.5?

We can use the RIF estimates from Figure 10, combined with population shares of African Americans and non-Hispanic whites in each vigintile of the initial distribution of PM2.5, to compute counterfactual pollution levels in the absence of the implementation of the nonattainment designations.[14] Table 4 walks through these calculations. Columns (1) and (2) of the top panel describe actual pollution levels within each quantile bin in 2000 and 2015, respectively. Columns (3) and (4) use estimates from Figure 10 to calculate what pollution in each quantile would have been in 2015 in the absence of the CAA PM2.5 NAAQS implementation, separately for African Americans and non-Hispanic whites.

The second panel of Table 4 considers the counterfactual gap that would have existed in 2015 in the absence of the CAA PM2.5 NAAQS. The actual gap in 2015 was 0.61 μg/m^3, whereas the counterfactual gap is 0.97 μg/m^3. The actual change in the black-white gap between 2005 and 2015 was 0.59 μg/m^3, and the implied counterfactual change is a narrowing of 0.23 μg/m^3. Since, we would have observed a 0.23 μg/m^3 improvement in the black-white gap in the absence of the policy, we conclude that the CAA can account for over 60% of the relative improvement in black-white outcomes.[15]

The last panel of Table 4 considers an alternative counterfactual, recognizing that there may have been some endogenous mobility response in response to these CAA-induced improvements in air quality that may have attenuated or exacerbated some of the relative improvements in outcomes between African Americans and non-Hispanic whites. To explore this counterfactual, we weight columns (3) and (4) by the 2005 population shares of whites and blacks in the respective quantile bins. Holding fixed the relative population shares, the counterfactual change in the gap would have been only 0.04 μg/m^3 instead of 0.59 μg/m^3. Thus, between 2005 and 2015 there were large relative shifts in population that unraveled some of the relative black-white convergence in mean air quality exposure.

Appendix Figure A3 explores the mobility responses in more detail. Figure A3a plots the within-quantile change in the black population share between 2005 and 2015; negative numbers imply a reduction in the black share in the particular quantile bin over this time period. These changes in population are plotted against the RIF-quantile treatment effects for blacks in the respective quantile (i.e. taken from Figure 10). We include a linear fit that suggests the areas that saw the

[14] For the calculations in this exercise, we use tract-level population shares by race from the public-use American Community Survey 5-year files.
[15] This is calculated as (0.59-0.23)/0.59.

largest improvements for African Americans over this time period are also the areas where the share of African Americans declined over this time period. Figure A3b shows the opposite is true for non-Hispanic whites. The quantiles which saw the largest improvements in air quality for non-Hispanic whites saw the largest increases in white population shares over this time period. These relative shifts in mobility served to offset some of the CAA-induced convergence over this time period.

How does one reconcile the results in Figures A3 with the earlier mobility results from Table 2 that suggested whites have moved to relatively more polluted areas in 2015 versus where they were in 2000? These findings are actually consistent with one another. While whites *are* disproportionately moving to urban areas (i.e. Table 2), they are moving to the urban areas that have experienced large improvements in air quality between 2000-2015 (i.e. Figure A3). By the same token, African-Americans who were already living in polluted urban areas became less dominant in these areas by 2015 as non-Hispanic whites moved in and so saw less of the gains than they otherwise would have.

V. Conclusions

This paper shows that racial differences in ambient particulate exposure declined significantly between 2000 and 2015. We add to the small but growing literature using high-resolution, nationwide data on pollution to examine racial differences in potential pollution exposure. We focus on PM2.5 and show that the gap between African Americans and non-Hispanic whites narrowed from -1.6 $\mu g/m^3$ in 2000 to -0.5 $\mu g/m^3$ by 2015. To our knowledge, this is the first paper to explore the underlying causal drivers that contributed to the narrowing of this gap. We find that very little of the decline in the gap in mean exposure levels can be accounted for by changes in mobility, individual, or household-level characteristics. Similarly, we find that racial gaps in exposure have narrowed at each quantile of the PM2.5 distribution, and that little of this narrowing can be explained by the demographic characteristics available in Census data. Instead, we find that virtually all of the closure of the gap is due to falling pollution levels in the areas where African Americans are more likely to live. There is little evidence that movement of African Americans to relatively cleaner neighborhoods or non-Hispanic whites to relatively dirtier neighborhoods has played a significant role in the observed convergence.

Why then has pollution fallen faster in areas that are home to disproportionate numbers of African Americans? Since African Americans have historically been concentrated in areas with more polluted air, and because the CAA targets the dirtiest areas for cleanup, it is reasonable to expect that the CAA could have a disproportionate impact on African Americans. However, set against this expectation are market forces that might cause African Americans in search of lower rents to move away from newly cleaner places. Our analysis shows that implementation of new PM2.5 standards in 2005 sharply reduced pollution in a way that was sustained over time and is

responsible for much of the closure of the racial gap in PM2.5 exposure between 2000 and 2015. We find some evidence however to suggest that changes in PM2.5 levels were associated with re-sorting of African Americans across areas in ways that undid some of these gains.

These findings suggest that the CAA has likely played a significant role in reducing racial gaps in exposure to air pollution, because the legislation systematically targeted the dirtiest areas for cleanup, and African Americans were disproportionately likely to live in areas with dirty air. Hence, although it was not their primary intent, the CAA has contributed to reductions in environmental inequality between racial groups in the United States.

References

Abel, Troy and Jonah White. 2011. “Skewed Riskscapes and Gentrified Inequities: Environmental Exposure Disparities in Seattle, Washington,” Am J Public Health. 2011 December; 101(Suppl 1): S246–S254.

Almond, Douglas, Janet Currie and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II," Journal of Economic Literature, vol 56(4), pages 1360-1446.

Apte, J. S., Messier, K. P., Gani, S., Brauer, M., Kirchstetter, T. W., Lunden, M. M., ... & Hamburg, S. P. (2017). High-resolution air pollution mapping with Google street view cars: exploiting big data. Environmental science & technology, 51(12), 6999-7008.

Ard, K. 2015. “Trends in exposure to industrial air toxins for different racial and socioeconomic groups: A spatial and temporal examination of environmental inequality in the U.S. from 1995 to 2004,” Social Science Research, Sept. 53:375-390.

Arias, Elizabeth and Jiaquan Xu. 2019. “United States Life Tables, 2017.” NCHS National Vital Statistics Reports, Vol. 68, No. 7.

Aufhammer, Maximilien, A. Bento and S. Lowe. 2009. “Measuring the Effects of the Clean Air Act Amendments on Ambient PM Concentrations: The Critical Importance of a Spatially Disaggregated Analysis,” Journal of Environmental Economics and Management 58(1): 15-26.

Banzhaf, Spencer. 2012. The Political Economy of Environmental Justice, Stanford University Press, 2012.

Banzhaf, H. Spencer, and Randall P. Walsh. 2008. "Do People Vote with Their Feet? An Empirical Test of Tiebout." American Economic Review, 98 (3): 843-63.

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.

Baum-Snow, N., & Hartley, D. A. (2019). Accounting for central neighborhood change, 1980-2010. Working Paper.

Bayer, Patrick, Nathaniel Keohane and Christopher Timmins, 2009. "Migration and hedonic valuation: The case of air quality," Journal of Environmental Economics and Management, Elsevier, vol. 58(1), pages 1-14, July.

Blinder, Alan S. “Wage Discrimination: Reduced Form and Structural Estimates.” *The Journal of Human Resources*, vol. 8, no. 4, 1973, pp. 436–455.

Brajer, Victor and Jane Hall., 2005. “Changes in the Distribution of Air Pollution Exposure in the Los Angeles Basin from 1990 to 1999,” Contemporary Economic Policy, 23(1): 50-58.

Cameron, Trudy Ann and Ian T. McConnaha. 2006. “Evidence of Environmental Migration.” Land Economics, 82(2): 273-290.

Chavis, Benjamin and Charles Lee. 1987. Toxic Wastes and Race in the United States, (United Church of Christ Committee on Racial Justice: New York).

Chay, Kenneth Y. and Michael Greenstone, 2005. "Does Air Quality Matter? Evidence from the Housing Market," Journal of Political Economy, University of Chicago Press, vol. 113(2), pages 376-424, April.

Chay, Kenneth Y. and Michael Greenstone, “Air Quality, Infant Mortality, and the Clean Air Act of 1970,” NBER Working Paper 10053 http://www.nber.org/papers/w10053, Oct. 2003.

Chernozhukov, V., Fernández-Val, I., & Melly, B. (2013). Inference on counterfactual distributions. Econometrica, 81(6), 2205-2268.

Christensen, Peter and Christopher Timmins. “Sorting or Steering: Experimental Evidence on the Economic Effects of Housing Discrimination,” NBER Working Paper #24826, July 2018.

Clark, Lara, Dylan B. Millet, and Julian D. Marshall. 2014. “National Patterns in Environmental Justice and Inequality: Outdoor NO_2 Air Pollution in the United States,” PLoS ONE, 9(4), e94431.

Clark, Lara, Dylan B. Millet, and Julian D. Marshall. 2017. “Changes in Transportation-Related Air Pollution Exposures by Race-Ethnicity and Socioeconomic Status: Outdoor Nitrogen Dioxide in the United States in 2000 and 2010,” Environmental Health Perspectives 125:9.

Couture, V., & Handbury, J. (2019). Urban revival in America, 2000 to 2010 (No. w24084). National Bureau of Economic Research.

Currie, Janet. "Inequality at Birth: Some Causes and Consequences," American Economic Review, American Economic Association, vol. 101(3), May 2011, p. 1-22.

Currie, Janet, Lucas Davis, Michael Greenstone, and Reed Walker. 2015. "Environmental Health Risks and Housing Values: Evidence from 1,600 Toxic Plant Openings and Closings." American Economic Review, 105(2): 678-709.

Currie, Janet and Matthew Neidell. "Air Pollution And Infant Health: What Can We Learn From California's Recent Experience?," Quarterly Journal of Economics, 2005, v120(3,Aug), 1003-1030.

Currie, Janet and Reed Walker, 2011. “Traffic Congestion and Infant Health: Evidence from E-ZPass,” American Economic Journal: Applied Economics, January 2011, 3(1) 65-90.

Currie, J., & Walker, R. (2019). What Do Economists Have to Say about the Clean Air Act 50 Years after the Establishment of the Environmental Protection Agency?. The Journal of Economic Perspectives, 33(4), 3-26.

Depro, Brooks, Christopher Timmins, and Maggie O'Neil. "White flight and coming to the nuisance: can residential mobility explain environmental injustice?" Journal of the Association of Environmental and resource Economists 2.3 (2015): 439-468.

Deryugina, Tatyana, Garth Heutel, Nolan Miller, David Molitor, and Julian Reif. "The Mortality and Medical Costs of Air Pollution: Evidence from Changes in Wind Direction,"
NBER Working Paper No. 22796, February 2018.

Dinardo, John, Nicole Fortin, and Thomas Lemieux. "Labor Market Institutions and the Distribution of Wages, 1973-1992: A Semiparametric Approach," Econometrica, 64 (5):1001-1044, September 1996.

Di, Qian, I Kloog, P Koutrakis, A Lyapustin, Y Wang, J Schwartz, 2016. "Assessing $PM_{2.5}$ Exposures with High Spatiotemporal Resolution across the Continental United States," Environmental science & technology 50 (9), 4712-4721.

Ebenstein et al. (2017) "New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River policy," Proceedings of the National Academy of Sciences, 114(39): 10384-10389.

Firpo, Sergio, Nicole M. Fortin, and Thomas Lemieux (2009) "Unconditional Quantile Regressions," Econometrica 77(3): 953-973.

Fowlie, M., Rubin, E., & Walker, R. (2019, May). Bringing Satellite-Based Air Quality Estimates Down to Earth. In AEA Papers and Proceedings (Vol. 109, pp. 283-88).

Graff Zivin, Joshua, and Matthew Neidell. 2013. "Environment, Health, and Human Capital." Journal of Economic Literature, 51 (3): 689-730.

Hamilton James T. Testing for Environmental Racism: Prejudice, Profits, Political Power? Journal of Policy Analysis and Management. 1995;14:107–32.

Hipp, John R and Cynthia M Lakon. 2010, "Social disparities in health: disproportionate toxicity proximity in minority communities over a decade," Health and Place, 16(4):674-83.

Hsiang, S., Oliva, P., & Walker, R. (2019). The distribution of environmental damages. Review of Environmental Economics and Policy, 13(1), 83-103.

Isen, A., Rossin-Slater, M., & Walker, W. R. (2014, January). Every breath you take – every dollar you'll make: the long term consequences of the Clean Air Act of 1970. *Journal of Political Economy*, University of Chicago Press, *vol 125(3)*, pp. 848 – 902.

Kravitz-Wirtz, Nicole, K Crowder, A Hajat, V Sass. 2016. "The Long-Term Dynamics of Racial/Ethnic Inequality in Neighborhood Air Pollution Exposure, 1990-2009," Du Bois Review: Social Science Research on Race, 13(2): 237-259.

Liebler, Carolyn, Sonya Porter, Leticia Fernandez, James Noon and Sharon Ennis. 2017. "America's Chruning Races: Race and Ethnicity Response Cahnges Between Census 2000 and the 2010 Census," Demography, 54 (1): 259-284.

Mohai, P., Pellow, D., & Roberts, J. T. (2009). Environmental justice. Annual Review of Environment and Resources, 34, 405-430.

Mohai, P., & Saha, R. (2006). Reassessing racial and socioeconomic disparities in environmental justice research. Demography, 43(2), 383-399.

Oaxaca, Ronald. 1973. "Male-Female Wage Differentials in Urban Labor Markets," International Economic Review, 14(3): 693-709.

Olley, G. S., & Pakes, A. (1996). The Dynamics of Productivity in the Telecommunications Equipment. Econometrica, 64(6), 1263-1297.

Pais, Jeremy, Kyle Crowder, Liam Downey, 2014. "Unequal Trajectories: Racial and Class Differences in Residential Exposure to Industrial Hazard," Social Forces, Mar 1; 92(3): 1189–1215.

Parry, Ian, Hilary Sigman, Margaret Walls, Roberton C. Williams III. The Incidence of Pollution Control Policies in Tom Tietenberg and Henk Folmer (eds.) International Yearbook of Environmental and Resource Economics 2006/2007 (Cheltenham, UK: Edward Elgar) 2006.

Schlenker, W., and Walker, W. R. (2016). Airports, air pollution, and contemporaneous health. *The Review of Economic Students, vol 83(2)*, pp 768-809.

Smith, V. Kerry and Ju-Chin Huang. 1995. "Can Markets Value Air Quality? A Meta-Analysis of Hedonic Property Value Models." The Journal of Political Economy, 103(1): 209-227.

U.S. General Accounting Office (GAO), 1983. Siting of Hazardous Waste Landfills and their Correlation with Racial and Economic Status of Surrounding Communities, (Washington, DC: General Accounting Office).

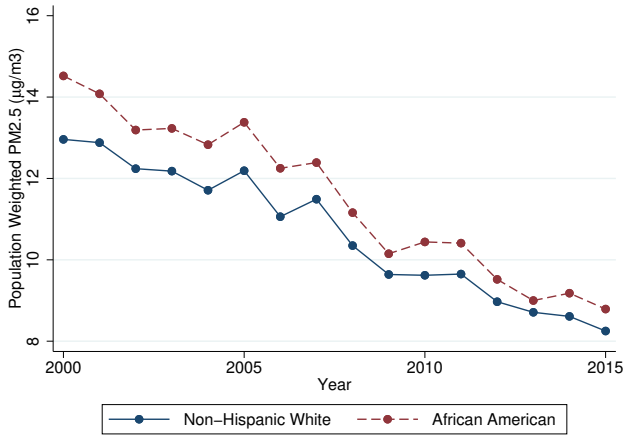
Voorheis, John. 2016. "Trends in Environmental Inequality in the United States: Evidence from Satellite Data," working paper.

Voorheis, John. 2018. "Longitudinal Environmental Inequality and Environmental Gentrification: Who Gains from Cleaner Air?" working paper.

Wagner, Deborah and Mary Layne. 2014. "The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software," CARRA Working Paper 2014-01.

1 Tables and Figures

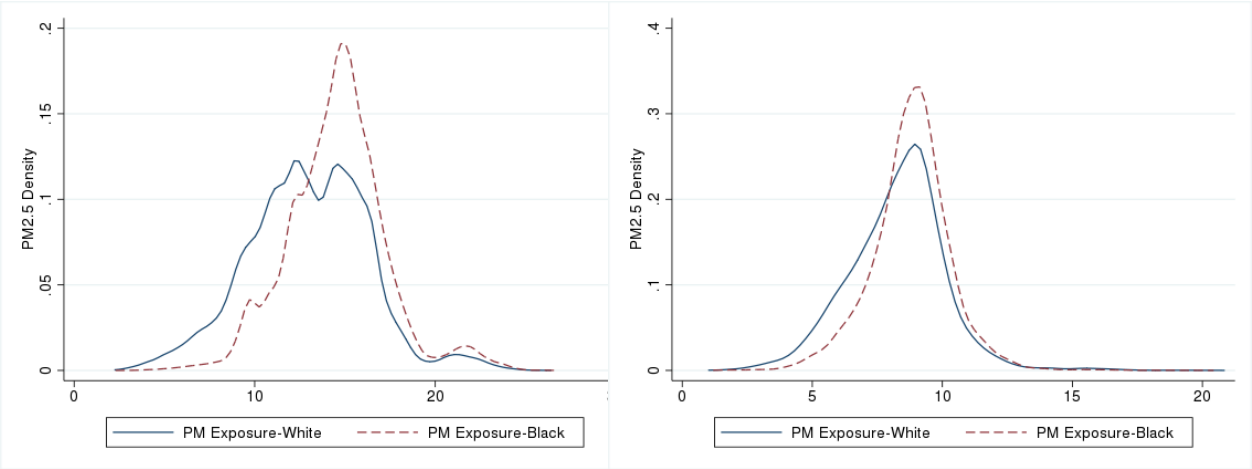
Figure 1: Trends in Pollution Exposure by Race



NOTES: This figure plots mean PM2.5 exposure by year, separately for African-Americans and the non-Hispanic white population. Source: Decennial Census, American Community Survey, and Di et al. (2016).

Figure 2: Distributions of Pollution Exposure, 2000 and 2015

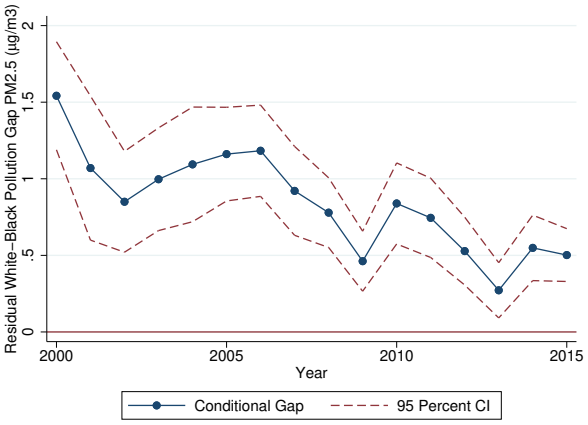
(a) 2000 (b) 2015



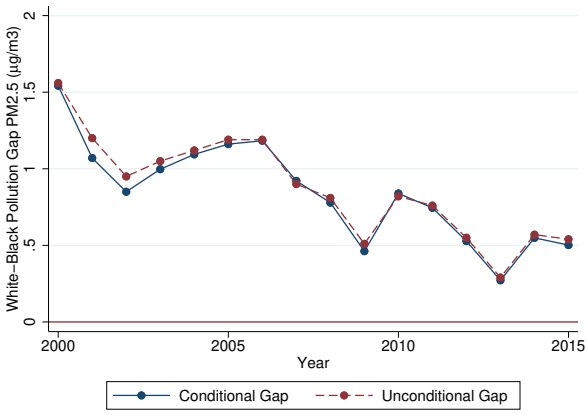
NOTES: This figure plots the PM2.5 density, separately for African-Americans and the non-Hispanic white population in both 2000 and 2015. Due to Census disclosure avoidance review, we were forced to trim the upper 97th and lower 3rd percentiles of each density. Source: Decennial Census, American Community Survey, and Di et al. (2016).

Figure 3: Residual Black-White Pollution Gap: PM2.5

(a) Conditional Gap with 95% Confidence Intervals



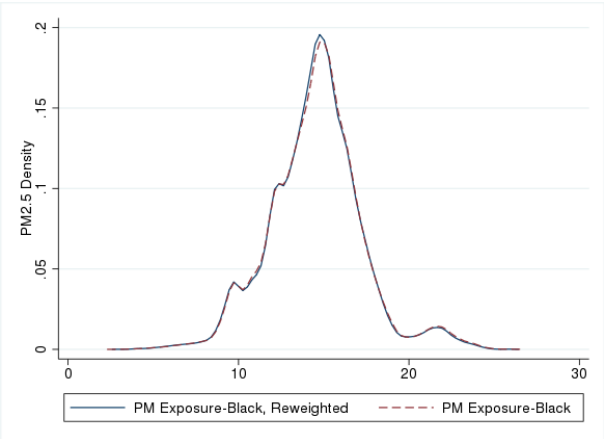
(b) Conditional vs. Unconditional Gap



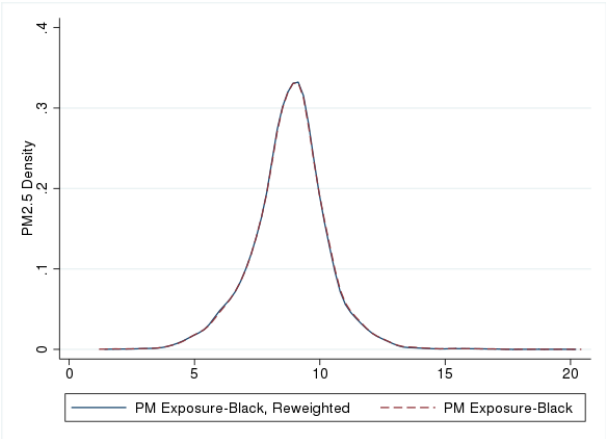
NOTES: Figure (a) plots the regression coefficients from 15 separate estimates of equation (1), regressing pollution on an indicator for whether an individual is an African-American, controlling for individual and household characteristics. The coefficient estimates correspond to the conditional mean black-white difference in air pollution, after adjusting for differences in observable individual and household characteristics. Dashed lines correspond to 95 percent confidence intervals, where standard errors have been clustered by commuting zone. Regressions are weighted using Census survey weights. Figure (b) compares the conditional mean coefficients to the unconditional mean differences by year. See text for details. Source: Decennial Census, American Community Survey, and Di et al. (2016).

Figure 4: Actual versus Counterfactual African American Pollution Distribution: PM2.5

(a) Reweighted vs. Actual PM2.5 Density
African Americans, 2000

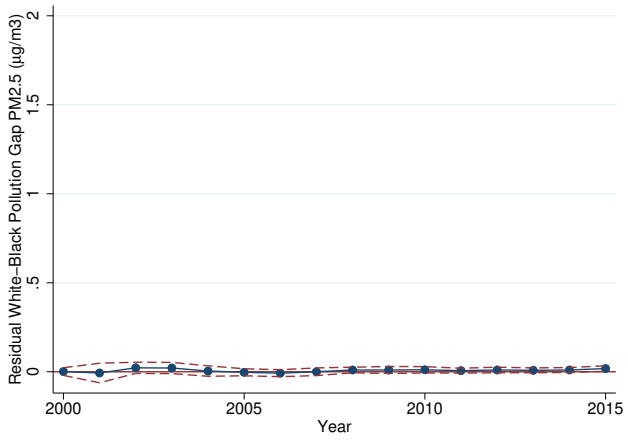


(b) Reweighted vs. Actual PM2.5 Density
African Americans, 2015



NOTES: These figures plot the actual versus counterfactual densities of pollution exposure for African Americans in 2000 and 2015. The counterfactual densities stem from an application of Dinardo, Fortin, Lemieux (1996), whereby we reweight the African American pollution distribution to reflect what the distribution would have looked like if they had the same individual characteristics as non-Hispanic whites in our sample. See text for details. Source: Decennial Census, American Community Survey, and Di et al. (2016).

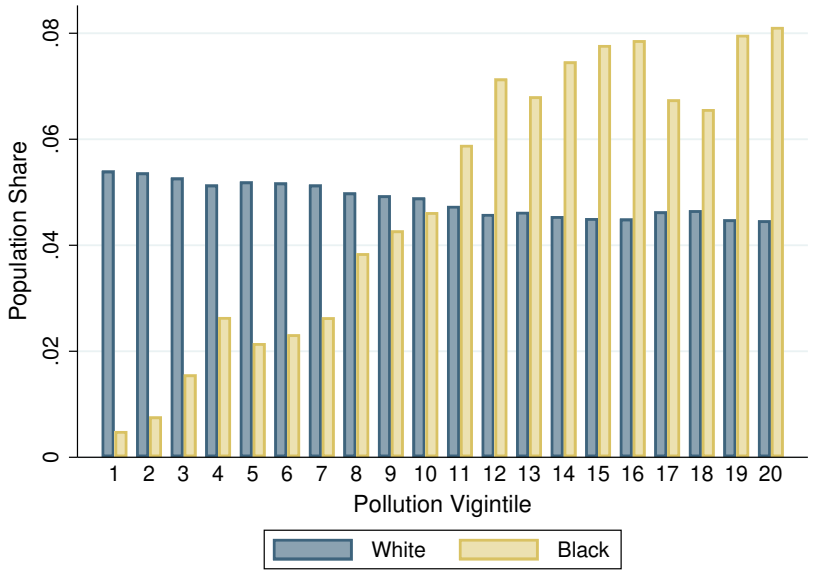
Figure 5: Residual Black-White Pollution Gap After Controlling for Neighborhood Characteristics: PM2.5



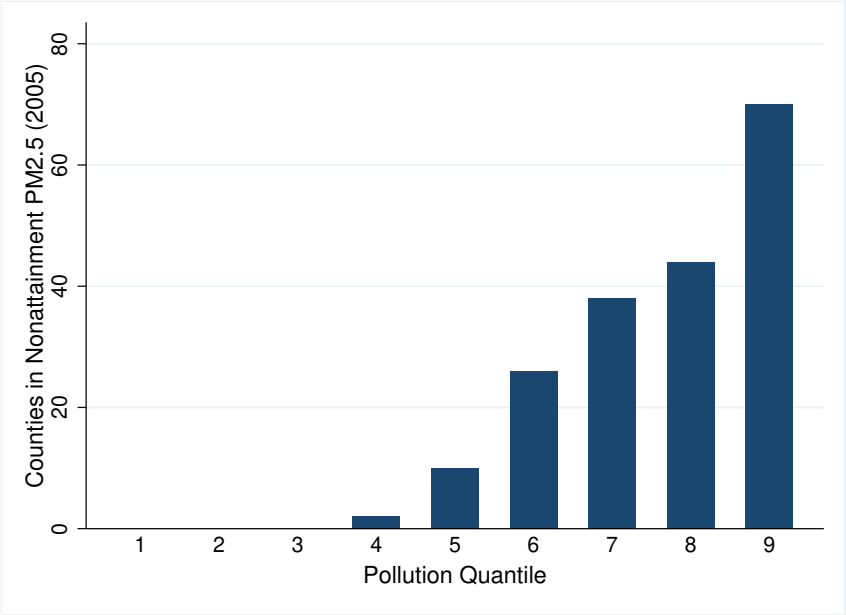
NOTES: This figure plots the regression coefficients from 15 separate estimates of equation (1), regressing pollution on an indicator for whether an individual is an African-American, controlling for individual, household, and census tract characteristics. The coefficient estimates correspond to the conditional mean black-white difference in air pollution, after adjusting for differences in observable individual, household, and census tract characteristics. Dashed lines correspond to 95 percent confidence intervals, where standard errors have been clustered by commuting zone. Regressions are weighted using Census survey weights. See text for details. Source: Decennial Census, American Community Survey, and Di et al. (2016).

Figure 6: Racial Distribution of Population and Impact of CAA by Pollution Decile

(a) Distribution of African-American and non-Hispanic White population by decile of PM2.5 in 2000

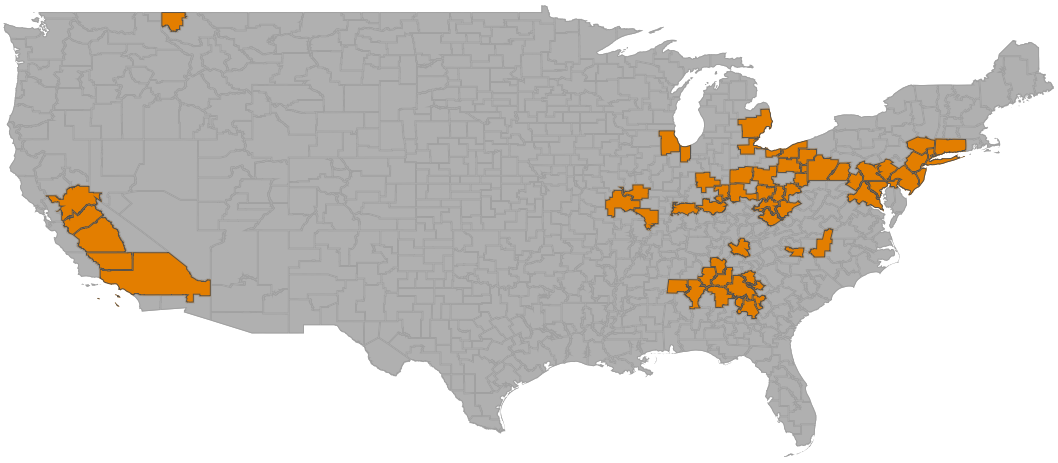


(b) Counties impacted by the CAA, by decile of PM2.5 in 2005



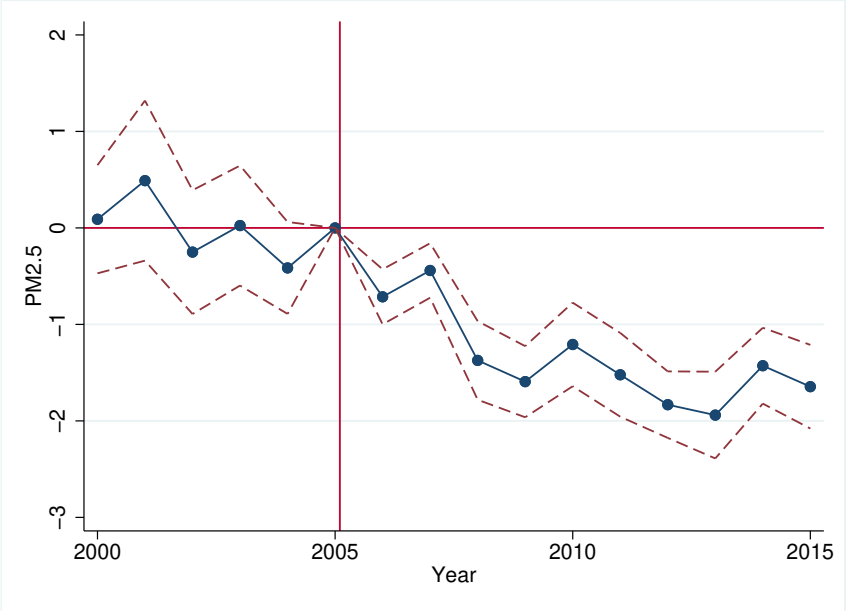
NOTES: Figure 5a plots population shares by pollution decile, separately for African American and non-Hispanic whites. Figure 5b shows the total number of counties subject to the Clean Air Act's 1997 NAAQS PM2.5 standard, by pollution decile. Source: Decennial Census, American Community Survey, EPA NAAQS Greenbook, and Di et al. (2016).

Figure 7: Commuting Zones in Nonattainment After the Adoption of the PM2.5 National Ambient Air Quality Standards



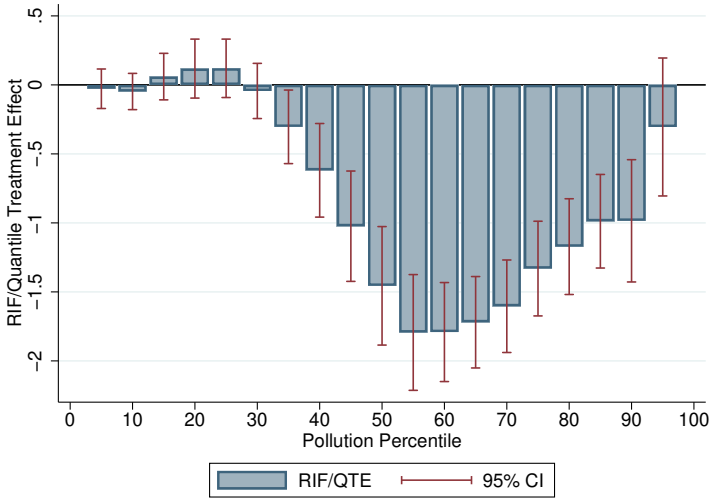
NOTES: This figure plots all the commuting zones containing at least one nonattainment county in 2005 for the Clean Air Act's 1997 NAAQS PM2.5 standard. While the PM2.5 NAAQS was initially proposed in 1997, the first year of regulatory enforcement began in 2005. Source: EPA NAAQS Greenbook.

Figure 8: The Effect of the PM2.5 NAAQS on Newly Regulated Commuting Zones



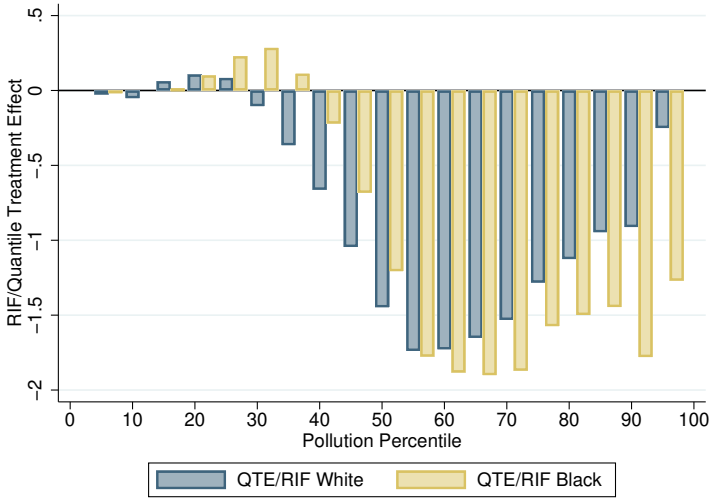
NOTES: This figure plots the event-time coefficient estimates from a version of equation (2), where the dependent variable consists of PM2.5 exposure ($\mu g/m^3$) for a given individual-year. The regression model controls for county and year fixed effects. The dashed lines represent 95% confidence intervals. Regressions are weighted by Census survey weights and errors are clustered by commuting zone. Source: Decennial Census, American Community Survey, EPA NAAQS Greenbook, Di et al. (2016).

Figure 9: RIF-Quantile Treatment Effects of the 2005 CAA PM2.5 NAAQS Implementation



NOTES: This figure plots the regression coefficient $\hat{\beta}$ from 19 separate versions of equation (3), where the dependent variable consists of the RIF-Quantile transformation of the respective PM2.5 vigintile (indicated by the x-axis). The regression model controls for county fixed effects and state-by-year fixed effects. The solid red lines represent 95% confidence intervals. Regressions are weighted by Census survey weights and errors are clustered by commuting zone. Source: Decennial Census, American Community Survey, EPA NAAQS Greenbook, Di et al. (2016).

Figure 10: Race-Specific RIF-Quantile Treatment Effects of the 2005 CAA PM2.5 NAAQS Implementation



NOTES: This figure plots the regression coefficient $\hat{\beta}$ from 38 separate versions of equation (3), 19 regressions for each race, where the dependent variable consists of the RIF-Quantile transformation of the respective PM2.5 vigintile (indicated by the x-axis). The regression model controls for county fixed effects and state-by-year fixed effects. Regressions are weighted by Census survey weights and errors are clustered by commuting zone. Source: Decennial Census, American Community Survey, EPA NAAQS Greenbook, Di et al. (2016).

Table 1: Residual Black-White Pollution Gap by Income Quintile: PM2.5

	Income Quintile				
	1	2	3	4	5
	Panel A: Year 2000				
1[African-American]	1.535	1.603	1.616	1.590	1.396
	(0.180)	(0.190)	(0.193)	(0.195)	(0.183)
Observations	1791000	1967000	2397000	2277000	2204000
	Panel B: Year 2015				
1[African-American]	0.499	0.463	0.512	0.558	0.570
	(0.097)	(0.093)	(0.096)	(0.089)	(0.099)
Observations	235000	223000	224000	226000	244000

Notes: This table presents the regression coefficients from 10 separate estimates of equation (1), 5 per panel. We regress pollution on an indicator for whether an individual is an African-American, controlling for individual and household characteristics, and we stratify the data by income quintile. The coefficient estimates correspond to the conditional mean black-white difference in air pollution, after adjusting for differences in observable individual and household characteristics. Panel A does this for the year 2000, and Panel B repeats this exercise in 2015. Regressions are weighted using Census survey weights, and standard errors are clustered by commuting zone. See text for details. Source: Decennial Census, American Community Survey, and Di et al. (2016).

Table 2: Counterfactual Pollution Levels and Gaps Holding Location Fixed

	(1) Actual 2000 Exposure	(2) Actual 2015 Exposure	(3) Counterfactual 2015 using 2000 locations
White PM2.5 $\mu g/m^3$	12.96	8.25	8.22
Black PM2.5 $\mu g/m^3$	14.52	8.79	8.89
Black-White Difference	1.56	0.54	0.67
Change in B-W Diff		1.02	0.89

Notes: Rows (1) and (2) of columns (1) and (2) present mean pollution exposure separately for African American and non-Hispanic Whites in years 2000 and 2015. Row (3) presents the mean gap in pollution exposure in either each year. Row (4) presents the change in black-white gap between 2000 and 2015. Column (3) presents a counterfactual exercise, whereby we ask what pollution levels would be and by how much the gap would have converged between 2000-2015 if we fixed the population in their 2000 location and assigned the 2015 pollution levels for their respective Census block. Source: Decennial Census, American Community Survey, and Di et al. (2016).

Table 3: The Impact of the 2005 Implementation of PM2.5 Standards on PM2.5 levels

	(1) PM2.5	(2) PM2.5	(3) ln(PM2.5)	(4) ln(PM2.5)	(5) PM2.5	(6) PM2.5	(7) ln(PM2.5)	(8) ln(PM2.5)
PM2.5 Nonattain×Post	-1.230	-1.237	-0.075	-0.076	-0.727	-0.726	-0.036	-0.036
	(0.335)	(0.334)	(0.020)	(0.020)	(0.080)	(0.082)	(0.006)	(0.006)
PM2.5 Non×Black×Post		0.149		0.008		0.048		0.004
		(0.088)		(0.007)		(0.091)		(0.005)
Year FE	X	X	X	X				
State-Year FE					X	X	X	X
County FE	X	X	X	X	X	X	X	X
Observations	32360000	32360000	32360000	32360000	32360000	32360000	32360000	32360000

Notes: This table presents regression coefficients from 8 separate versions of equation (3), one per column, where the dependent variable consists of PM2.5 or ln(PM2.5) for an individual in a given year. Columns (2), (4), (6), and (8) add an additional interaction for African Americans to test for heterogeneity in regulatory impacts for African Americans. Regressions are weighted by Census survey weights and errors are clustered by commuting zone. Source: Decennial Census, American Community Survey, EPA NAAQS Greenbook, Di et al. (2016).

Table 4: Calculating the Effect of CAA Regulations on the Black-White PM2.5 Gap

(1) PM2.5 Quantile Bin	(2) Actual PM2.5 in 2005	(3) Actual PM2.5 in 2015	(4) White Counterfactual PM2.5 in 2015 Without CAA	(5) Black Counterfactual PM2.5 in 2015 Without CAA
5	5.32	4.34	4.37	4.36
10	7.87	5.63	5.69	5.63
15	8.91	6.25	6.18	6.24
20	9.65	6.72	6.62	6.62
25	10.33	7.11	7.03	6.88
30	10.90	7.45	7.56	7.17
35	11.42	7.75	8.12	7.64
40	11.90	8.01	8.67	8.23
45	12.34	8.24	9.28	8.92
50	12.73	8.44	9.89	9.65
55	13.09	8.65	10.39	10.43
60	13.44	8.84	10.57	10.73
65	13.80	9.03	10.68	10.93
70	14.15	9.22	10.75	11.09
75	14.51	9.42	10.71	11.00
80	14.91	9.67	10.80	11.17
85	15.27	9.98	10.93	11.43
90	15.72	10.49	11.41	12.27
95	17.01	12.21	12.46	13.48
Main Counterfactual: Including 2005-2015 Mobility Responses				
2005 Actual Black-White Gap:				1.20
2015 Counterfactual Black-White Gap:				0.97
Counterfactual Change in Black-White Gap:				-0.23
Actual Change in Black-White Gap:				-0.59
% of Actual Gap Attributable to CAA:				**61.2%**
Alternative Counterfactual: Holding Population Fixed in 2005 Locations				
2005 Actual Black-White Gap:				1.20
2015 Counterfactual Black-White Gap:				1.16
Counterfactual Change in Black-White Gap:				-0.04
Actual Change in Black-White Gap:				-0.59

Notes: This table presents calculations used to explore what fraction of the observed racial convergence in mean PM2.5 levels can be attributed to the regulatory variation embedded into the Clean Air Act's 2005 PM2.5 NAAQS. The top panel describes actual pollution levels within each quantile bin in 2000 and 2015. Columns (3) and (4) use estimates from Figure 10 to calculate what pollution would be in 2015 in the absence of the CAA PM2.5 NAAQS implementation, separately for African Americans and non-Hispanic whites. The second panel computes the counterfactual gap in 2015 in the absence of the CAA NAAQS and the implied 2005-2015 change in the gap. The last panel asks what the change in the black-white gap would have been if people did not move between 2005-2015, and the CAA NAAQS was not implemented. Source: Decennial Census, American Community Survey, EPA NAAQS Greenbook, Di et al. (2016).

Online Appendix: Not for Publication

Appendix A: Pollution Decompositions

Oaxaca-Blinder Decomposition of Racial Pollution Gaps

Assume that for an individual i in group $g \in \{b, w\}$ pollution exposure P_{gi} can be written as a linear function of observed characteristics (X_{gi}) and an error term μ_{gi}

$$P_{gi} = \beta_g X_{gi} + \mu_{gi}, \quad (1)$$

where β_g are defined so that $E[\mu_{gi}|X_{gi}] = 0$. The difference in expected pollution exposure for blacks and whites can be written as:

$$P_b - P_w = (X_b - X_w)\beta_b + (\beta_b - \beta_w)X_w, \quad (2)$$

where P_g and X_g represent mean pollution and mean characteristics for all individuals in group g. The first term on the right-hand side of equation (2) is the part of the gap explained by differences in the average observable characteristics of the two groups; namely, how large would the observed pollution gap be if we gave blacks the same mean characteristics as whites? The second term is the part of the mean difference in pollution exposure that is not explained by differences in characteristics; instead it reflect differences in the "returns" to the observable characteristics (that is, the differences in slopes between groups, $\beta_b - \beta_w$, for each observed characteristic).

The results of this initial decomposition are shown in Appendix Table 2 for the two end points of our sample, 2000 and 2015. The first row shows the predicted difference in pollution exposure obtained from left hand side of equation (2): The gap is -1.616 μg/m^3 in 2000, falling to -0.544 μg/m^3 by 2015. The first panel of Table 1 ("Explained") shows that very little of the racial gap in predicted pollution exposure can be explained by an individual's observable characteristics. While African American households have mean household income more than $15,000 less than non-Hispanic whites (see Appendix Table 1), these differences in income explain almost none of the observed differences in pollution exposure. In fact, almost none of the observed individual or household characteristics are able to explain *any* portion of the observed difference in pollution exposure. Only 4.8 percent of the gap (-0.078/-1.616) is explained by differences in income, age, schooling, children, gender, and/or homeownership. In 2015, differences in these characteristics between blacks and whites are able to explain only 8 percent (-0.044/-0.544) of the gap in pollution exposure between black and whites.[1] Hence, differential racial exposure to pollution cannot be explained simply by the fact that African Americans are more disadvantaged in terms of measured characteristics on average than non-Hispanic whites.

The second panel of Table 1 ("Unexplained") shows that the bulk of the racial gap in pollution exposure instead reflects racial differences in the "returns" to observed characteristics and/or

[1] We have experimented with a range of more flexible functional forms for all the control variables, and the qualitative results are nearly identical when including higher order polynomials and/or more flexible dummy variable transformations of the observed continuous variables.

differences in unmeasured characteristics. In particular, differences in the coefficients on age and education explain a significant portion of the gap in both 2000 and 2015.

Decompositions Using Recentered Influence Functions (RIF)

While the previous section explored differences in mean outcomes between blacks and whites, it is also possible to decompose other parts of the pollution distribution (e.g., the 10th or 90th percentiles of the white and black pollution distributions). Are observable characteristics able to explain more of the difference in outcomes at the 90th percentiles? What about the 10th percentiles? Recent advances in quantile regression allow us to decompose differences in quantiles of the unconditional pollution distribution using recentered influence functions (RIF) (Firpo et al. 2009). The basic idea is to transform the problem by considering a covariate's influence on population shares rather than quantiles. By estimating how a covariate (e.g., income) affects the share of the population below various pollution thresholds, we can identify the marginal effect of income on the cumulative distribution function (CDF) of pollution. We can then invert the impact of income on the CDF of pollution to estimate the impact on a pollution quantile. The RIF regression approach proposed by Firpo et al. (2009) performs this inversion using a local linear approximation to the counterfactual CDF, rescaling the marginal effect of each covariate on the share above a pollution cutoff by the probability density of pollution at that cutoff.

In practice, RIF regression requires first transforming the outcome variable, PM2.5 pollution, using a recentered influence function before projecting this transformation on the explanatory variables of interest. RIF-regression methods provide a simple way of performing decompositions for any distributional statistic for which an "influence function" can be computed.

Firpo et al. (2009) consider the following model of pollution *P*:

$$P = h(X, \epsilon)$$

where X represents the set of independent, explanatory variables and ϵ is the scalar unobserved error term. The unconditional partial effect is defined as the shift in the distribution of a continuous variable X on the distributional statistic $v(F_p)$, which can be expressed as

$$\alpha(v) = \int \frac{dE[\text{RIF}(P, v)|X = x]}{dx} dF(x)$$

where $\text{RIF}(P, v)$ is the recentered influence function. When the distributional statistic v is the τth quantile function $q_t = \inf_q\{q : F_P(q) \geq \tau\}$ the $RIF(P, q_\tau)$ can be represented as

$$\text{RIF}(P, q_\tau) = q_\tau + \frac{\tau - 1\{p \leq q_\tau\}}{f_P(q_\tau)}$$

where $f_P(q_\tau)$ is the density function of pollution P evaluated at quantile q_τ.

The relevant property of a recentered influence function is that its expectation equals the distributional statistic of interest. For quantile τ denoted Q_τ, the quantile RIF is given by

$\text{RIF}(p, Q_\tau) = Q_\tau + \frac{\tau - \mathbf{1}(p < Q_\tau)}{f_p(Q_\tau)}$ and taking expectations verifies $E[RIF(p, Q_\tau)] = Q_\tau$. Since the mean of the RIF is equal to the quantile, we can use the law of iterated expectations to decompose each unconditional quantile, as Oaxaca (1973) and Blinder (1973) do when they decompose the mean.

Firpo et al. (2009) show that a regression of the RIF on covariates yields the approximate effect of the covariates on the distributional statistic of interest (applied to the unconditional distribution). As we show in the text, this feature of RIF regressions provides a natural bridge to exploring how treatment effects (e.g., the effects of the CAA PM2.5 regulations on county-year pollution levels) map into the unconditional distribution of pollution.

Appendix Table A3 shows the results of decompositions of the 10th, 50th, and 90th percentiles of the pollution distributions for non-Hispanic whites and African Americans using re-centered influence functions. The first three columns show the estimates for 2000, while the last three columns show the estimates for 2015. Both sets of estimates indicate larger gaps at the 10th percentile than at the 90th percentile. These patterns of quantile differences can largely be explained by recognizing that both non-Hispanic whites and African American live in large cities with high levels of air pollution (i.e. the 90th percentiles of their respective pollution distributions are somewhat similar); whereas rural locations, that also tend to be the least polluted parts of the United States, are disproportionately white, leading to larger gaps in the 10th percentiles of the respective race-specific pollution distributions. As in the Oaxaca-Blinder decompositions, these breakdowns indicate that relatively little of the predicted differences in exposure can be explained by differences in individual and household level characteristics. By 2015, the racial gap in the 90th percentile of exposure has narrowed significantly, and, for the first time, most of the gap can be accounted for by differences in observable characteristics. For the most part, however, the quantile decompositions present a pattern similar to the original mean decompositions; black-white differences in individual or household-level characteristics explain very little of the black-white difference in PM2.5 pollution exposure, especially in earlier years.

<u>People versus Places: Decomposing the Role of Population Shifts in Changing the Pollution Gap</u>

We decompose the pollution exposure to explore the role of reallocation of population shares within racial groups over time. We consider pollution exposure (Ω) as the population share (denoted by s_{it}) weighted average of tract level pollution ω_{it}. We rely on the following definition of nationwide pollution exposure $\Omega_t = \sum_i s_{it}\omega_{it}$, which is different from the unweighted average of census tract pollution exposure $\overline{\omega_t} = \frac{1}{N_t}\sum_i \omega_{it}$.

We decompose this nationwide pollution term into unweighted tract-level pollution and the covariance between pollution and tract-level population shares. The same decomposition can be applied by race group ϕ, which can inform us whether the observed pollution changes are driven mostly by improvements in average census tracts versus shifts in race-specific population shares across census tracts over time. Denote the population share of each race group as $s(\phi)_t = \sum_{i \in \phi} s_{it}$. Likewise, denote race-specific pollution is $\Omega_t(\phi)$, while the average tract-level pollution within a race group is $\overline{\omega}_t(\phi)$. Then we can write nationwide pollution exposure as a weighted average of the race-specific components

$$\Omega_t = \sum_{\phi \in b,w} s_t(\phi) \left((\overline{\omega}_t(\phi) + \sum_{i \in \phi} (\omega_{it} - \overline{\omega}_t(\phi))(s_{it}(\phi) - \bar{s}_t(\phi)) \right)$$

$$= \sum_{\phi \in b,w} s_t(\phi) \left(\overline{\omega}_t(\phi) + \Gamma_t^{\mathrm{OP}}(\phi) \right)$$

where $\Gamma_t^{\mathrm{OP}}(\phi)$ reflects the covariance between population shares and pollution levels for a respective race group. If African Americans are more concentrated in census tracts with high pollution levels, this term will be positive for African Americans.

This equation allows us to explain changes in nationwide pollution, through (i) changes in the average pollution of black and white census tracts ($\overline{\omega}_t(\phi)$), and (ii) changes in the covariance between population shares and pollution, separately for both blacks and whites ($\Gamma_t^{\mathrm{OP}}(\phi)$).

We perform this decomposition year by year to better understand the dynamics of this covariance term; if the covariance is falling over time, this would suggest that African Americans were becoming less concentrated in neighborhoods with the worst air quality. If the covariance term were relatively constant over time, the observed improvement in air quality for African Americans is primarily the result of the average African American census tract getting cleaner.

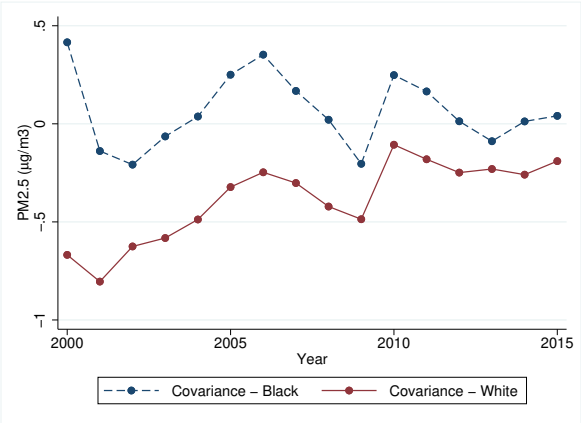
Appendix Figure A1 shows the results of implementing the decomposition in equation (4) year-by-year, separately for non-Hispanic whites and African Americans. Appendix Table A5 similarly decomposes the 2000-2015 difference in PM2.5 exposure. Figure A1a shows that there is little change in the covariance between air pollution levels and share African American in a Census tract. However, over time, the negative relationship between pollution and non-Hispanic white population shares weakens.

Figure A1b shows that for African Americans, there is virtually a one to one relationship between average individual exposure and the unweighted mean census tract pollution level, consistent with the flat trend in the covariances in Figure A1a. Hence, among African Americans, virtually all of the reduction in pollution exposure can be accounted for by the average African American tract cleaning up, rather than by relocation of African Americans to relatively cleaner tracts. Formally, Column (2) of Table A5 suggests 93% of the 2000-2015 improvement in air quality for African Americans can be explained by changes in the average African American tract. The remaining seven percent can be explained by a weakening of the covariance between African American population shares and pollution exposure. Figure A1c shows that the average white pollution exposure is slightly lower than average tract-level exposure, indicating that whites live in cleaner tracts. However, this gap narrows over time, consistent with the trend in covariances shown in Figure A1a. Overall then, this decomposition indicates that a small part of the gap between African Americans and non-Hispanic whites can be accounted for by whites becoming less concentrated in the relatively cleanest neighborhoods, but that most of the relative air quality improvement for African Americans reflects clean ups within tracts rather than relative shifts in black-white population shares to cleaner or dirtier tracts.

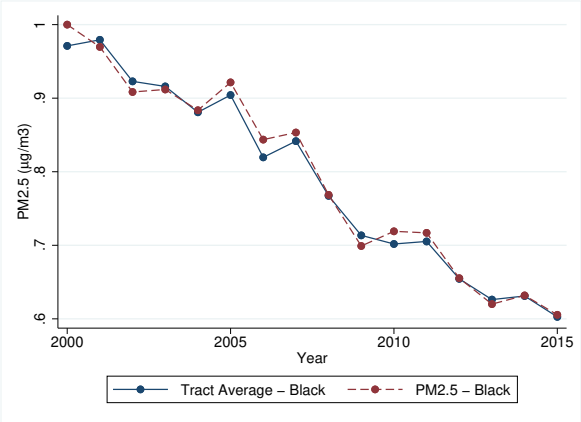
A Appendix Tables and Figures

Figure A1: Individual vs. Tract-Level Exposure and Covariance Between Race Share and Air Quality

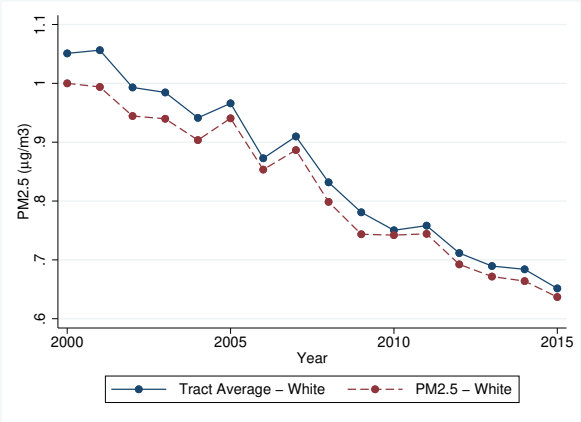
(a) Tract PM2.5 and Population Share Covariances



(b) Individual and Tract Average PM2.5 Exposure: African Americans

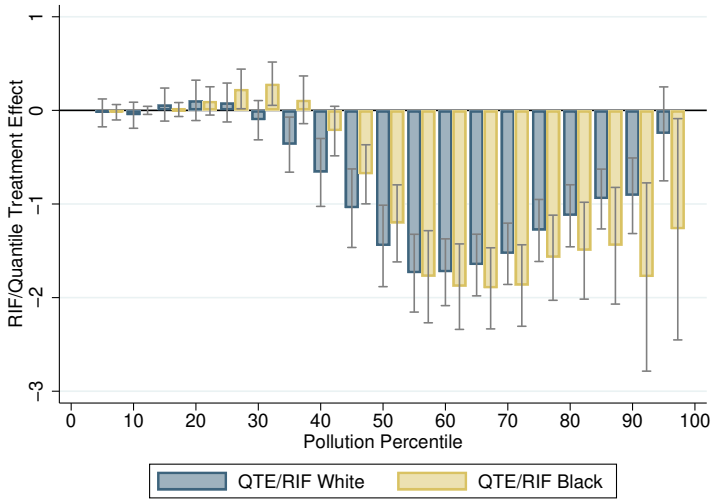


(c) Individual and Tract Average PM2.5 Exposure: non-Hispanic White



NOTES: These figures present results from decomposing average PM2.5 exposure into the unweighted average Census tract exposure for each race group and the assortative relationship between race-specific population shares and pollution levels. Figure (a) presents the year-by-year covariance between race-specific population shares and pollution levels, separately by race. Figure (b) plots the population-weighted trend in PM2.5 exposure for African Americans (dashed line) and the trend in unweighted Census tract exposure for African Americans (solid line). Figure (c) replicates Figure (b) for Non-Hispanic whites. See text for details. Source: Decennial Census, American Community Survey, and Di et al. (2016).

Figure A2: Race-Specific RIF-Quantile Treatment Effects of the 2005 CAA PM2.5 NAAQS Implementation

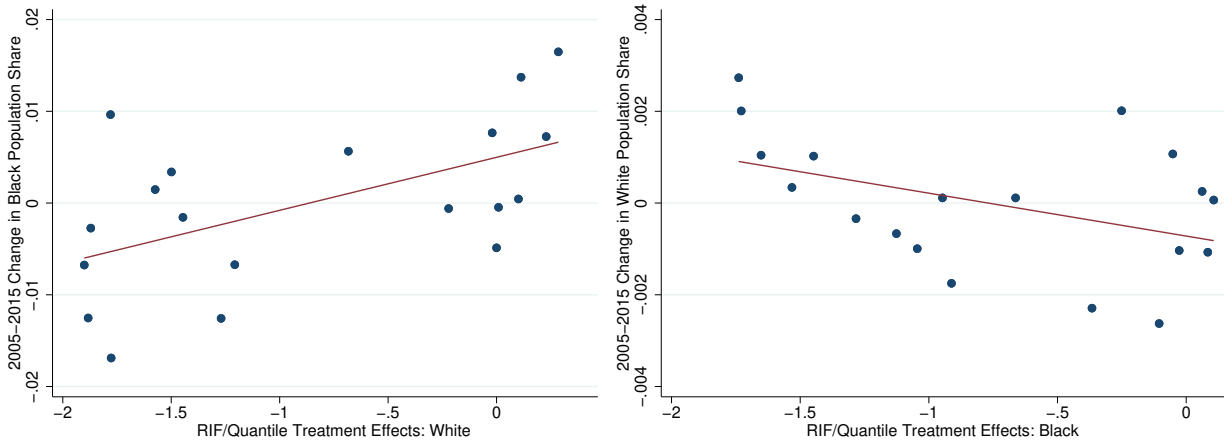


Notes: This figure reproduces Figure 10, including confidence intervals for the regression coefficients. The figure plots the regression coefficient $\hat{\beta}$ from 38 separate versions of equation (3), 19 regressions for each race, where the dependent variable consists of the RIF-Quantile transformation of the respective PM2.5 vigintile (indicated by the x-axis). The regression model controls for county fixed effects and state-by-year fixed effects. The solid gray lines represent 95% confidence intervals. Regressions are weighted by Census survey weights and errors are clustered by commuting zone. Source: Decennial Census, American Community Survey, EPA NAAQS Greenbook, Di et al. (2016).

Figure A3: 2005-2015 Change in Population Plotted Against Quantile Treatment Effects

(a) African Americans

(b) Non-Hispanic White



Notes: Figure A3a plots the within-quantile change in the black population share between 2005 and 2015 against the RIF-quantile treatment effects for blacks in the respective quantile (i.e. taken from Figure 10). Figure A3b repeats this exercise for non-Hispanic whites. An observation corresponds to a particular quantile bin. Source: Decennial Census, American Community Survey, and Di et al. (2016).

Table A1: Summary Statistics by Race, Overall, and Sub-Periods

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Overall		African-American		Non-Hispanic White		Mean Diff.	
	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.	(5)-(3)	p-value
Panel A: Individual Characteristics								
Age	40.010	8.573	39.410	8.536	40.100	8.575	0.687	(0.000)
Years of School	13.650	2.650	13.130	2.405	13.720	2.674	0.587	(0.000)
Sex (1=Female)	0.514	0.500	0.549	0.498	0.509	0.500	-0.041	(0.000)
Homeowner	0.703	0.457	0.486	0.500	0.733	0.443	0.247	(0.000)
Number of Children	1.070	1.213	1.039	1.267	1.074	1.206	0.035	(0.000)
Income	48130	51590	34300	34630	50070	53250	15760	(0.000)
Bottom Income Quintile	0.200	0.400	0.264	0.441	0.191	0.393	-0.073	(0.000)
Top Income Quintile	0.200	0.400	0.106	0.307	0.213	0.410	0.108	(0.000)
PM2.5 (Satellite, Block)	10.770	2.980	11.460	2.748	10.680	2.999	-0.780	(0.000)
PM2.5 (Satellite, County)	10.770	2.812	11.390	2.608	10.680	2.829	-0.705	(0.000)
PM2.5 (EPA Monitors, County)	11.460	2.948	12.040	2.781	11.360	2.964	-0.679	(0.000)
Panel B: Census Tract Characteristics in 2000								
African American	0.123	0.131	0.262	0.175	0.103	0.110	-0.158	(0.000)
Public Assistance Income	34.04	39.98	34.35	40.89	33.99	39.85	-0.352	(0.902)
Income	48130	12920	47660	12780	48200	12930	540	(0.392)
Years of Schooling	13.640	0.708	13.680	0.683	13.640	0.712	-0.035	(0.276)
% Worked Last Year	0.834	0.047	0.828	0.046	0.835	0.047	0.007	(0.012)
Housing Value	292500	183000	292200	178900	292500	183500	299	(0.980)
Housing Rent	1096	317	1116	294	1094	320	-22.220	(0.203)
% Home Owners	0.703	0.111	0.657	0.120	0.709	0.108	0.053	(0.000)
% Single Family Residence	0.831	0.051	0.822	0.049	0.833	0.051	0.011	(0.000)
% in Urban County	0.992	0.089	0.997	0.057	0.991	0.092	-0.005	(0.000)
% Manufacturing Emp.	0.133	0.095	0.115	0.086	0.136	0.096	0.022	(0.000)
Panel C: County-Level Characteristics in 2000								
African American	0.129	0.232	0.556	0.320	0.069	0.134	-0.487	(0.000)
Welfare Income	30.50	135.80	51.71	205.90	27.53	122.60	-24.180	(0.000)
Years of School	13.590	1.409	13.150	1.305	13.650	1.412	0.496	(0.000)
Single Family Residence	0.824	0.163	0.792	0.186	0.829	0.159	0.037	(0.000)
Teen Pregnancy	-0.042	0.061	-0.063	0.074	-0.039	0.058	0.024	(0.000)
Home Ownership	0.720	0.204	0.603	0.236	0.737	0.194	0.134	(0.000)

Notes:

Table A2: Decomposition of Mean Differences in Pollution Exposure into Components Explained by Differences in Individual Characteristics and due to Differences in "Returns" to Characteristics

	(1) Year 2000	(2) Year 2015
Predicted difference	-1.616	-0.544
Panel A: Explained Gap		
Income	-0.001	0.000
Age	-0.009	-0.002
Schooling	-0.011	-0.010
Kids	0.003	0.001
Gender	0.000	0.000
Homeowner	-0.061	-0.033
Total	-0.078	-0.044
Panel B: Unexplained Gap		
Income	0.040	0.013
Age	-0.412	-0.251
Schooling	-0.419	-0.456
Kids	0.018	0.049
Gender	-0.009	-0.002
Homeowner	-0.002	0.000
Constant	-0.755	0.146
Total	-1.537	-0.500
N	10550000	1185000

Notes: This table plots the results from an Oaxaca-Blinder decomposition of mean differences in PM2.5 exposure between African-Americans and non-Hispanic whites. Column (1) performs this decomposition for the year 2000, whereas column (2) decomposes differences originating in 2015. Panel A displays the amount by which black-white differences in the respective covariates explain the gap in mean PM2.5 exposure between groups. Panel B presents the amount by which black-white differences in the respective coefficient estimates explain the gap in mean PM2.5 exposure between groups. Source: Decennial Census, American Community Survey, and Di et al. (2016).

Table A3: Decomposition of Quantile Differences in Pollution Exposure into Components Explained by Differences in Individual Characteristics and due to Differences in "Returns" to Characteristics

	(1)	(2)	(3)	(4)	(5)	(6)
	Year 2000			Year 2015		
	10th Percentile	50th Percentile	90th Percentile	10th Percentile	50th Percentile	90th Percentile
Predicted difference	-2.360	-1.607	-0.861	-1.081	-0.434	-0.130
			Panel A: Explained Gap			
Income	0.004	-0.003	-0.001	0.001	0.000	0.001
Age	-0.015	-0.008	-0.002	-0.003	-0.002	-0.002
Schooling	0.002	-0.006	0.001	0.003	-0.001	-0.021
Kids	0.004	0.004	0.002	-0.001	0.000	0.002
Gender	-0.002	0.001	0.001	0.0001	0.000	0.000
Homeowner	-0.041	-0.010	-0.120	-0.034	-0.018	-0.056
Total	-0.049	-0.023	-0.119	-0.034	-0.019	-0.077
			Panel B: Unexplained Gap			
Income	0.240	-0.057	0.011	0.016	-0.008	0.046
Age	-0.696	-0.352	0.036	-0.399	-0.205	-0.262
Schooling	0.316	-0.147	-0.981	-0.114	-0.263	-0.696
Kids	0.045	-0.004	0.013	0.040	0.034	0.059
Gender	-0.018	-0.007	-0.002	-0.007	-0.002	0.001
Homeowner	-0.006	0.004	-0.000	0.001	-0.003	0.003
Constant	-2.191	-1.020	0.181	-0.585	0.031	0.795
Total	-2.311	-1.584	-0.742	-1.047	-0.415	-0.053
N	10550000	10550000	10550000	1185000	1185000	1185000

Notes: This table plots the results from six RIF decompositions of quantile differences in PM2.5 exposure between African-Americans and non-Hispanic whites, where quantiles are indicate in the column headings. Columns (1)-(3) perform this decomposition for the year 2000, whereas columns (4)-(6) decompose differences originating in 2015. Panel A displays the amount by which black-white differences in the respective covariates explain the gap in quantile PM2.5 exposure between groups. Panel B presents the amount by which black-white differences in the respective coefficient estimates explain the gap in quantile PM2.5 exposure between groups. Source: Decennial Census, American Community Survey, and Di et al. (2016).

Table A4: Oaxaca-Blinder Decomposition, Census Tract Characteristics

	(1) Year 2000	(2) Year 2015
Predicted difference	-1.617	-0.772
Panel A: Explained Gap		
Income	0.000	0.000
Age	-0.008	-0.003
Schooling	-0.002	-0.003
Kids	0.002	0.001
Gender	-0.001	0.000
Homeowner	0.001	0.019
Neighborhood/Tract		
Tract % Black	-2.311	-0.502
Tract Public Assistance Income	-0.037	-0.007
Tract Years of Schooling	-0.034	-0.021
Tract % Single Family Residence	0.280	0.010
Tract Teen Pregnancy Rate	0.032	-0.002
Tract Home Ownership Rate	-0.266	-0.116
Total	-2.344	-1.032
Panel B: Unexplained Gap		
Income	-0.052	0.028
Age	-0.221	-0.218
Schooling	-0.359	-0.292
Kids	0.039	0.047
Gender	-0.003	-0.003
Homeowner	0.002	-0.004
Neighborhood/Tract		
Tract % Black	1.687	0.185
Tract Public Assistance Income	-0.190	0.010
Tract Years of Schooling	-3.274	-1.046
Tract % Single Family Residence	1.692	0.139
Tract Teen Pregnancy Rate	-0.074	-0.015
Tract Home Ownership Rate	-0.875	-0.403
Constant	2.356	0.866
Total	0.727	0.260
N	10550000	1185000

Notes: This table plots the results from two Oaxaca-Blinder decompositions of mean differences in PM2.5 exposure between African-Americans and non-Hispanic whites. Column (1) performs this decomposition for the year 2000, whereas column (2) decomposes differences originating in 2015. Panel A displays the amount by which black-white differences in the respective covariates explain the gap in mean PM2.5 exposure between groups. Panel B presents the amount by which black-white differences in the respective coefficient estimates explain the gap in mean PM2.5 exposure between groups. Source: Decennial Census, American Community Survey, and Di et al. (2016).

Table A5: Decompositions of PM2.5 Exposure Changes 2000-2015 by Race

	(1) Non-Hispanic White	(2) African American
Average PM2.5 Change by Race: 2000-2015 ($\Omega(\phi)$)	-4.716	-5.728
Decomposition:		
Unweighted Average Tract Change: 2000-2015 ($\bar{\omega}(\phi)$)	-5.176 (110%)	-5.349 (93.3%)
Change in Covariance 2000-2015 ($\Gamma^{OP}(\phi)$)	0.478 (-0.10%)	-0.375 (6.6%)

Notes: This table decomposes the average 2000-2015 change in pollution exposure by race into that which can be explained by unweighted Census-tract level changes in PM2.5 versus reallocation of population shares to cleaner/dirtier Census tracts over time. Source: Decennial Census, American Community Survey, and Di et al. (2016).