APPENDIX FOR ONLINE PUBLICATION

Clean Identification? The Effects of the Clean Air Act on Air Pollution, Exposure Disparities and House Prices

by Lutz Sager[1,*], Gregor Singer[2,*]

[1]Georgetown University

[2]London School of Economics

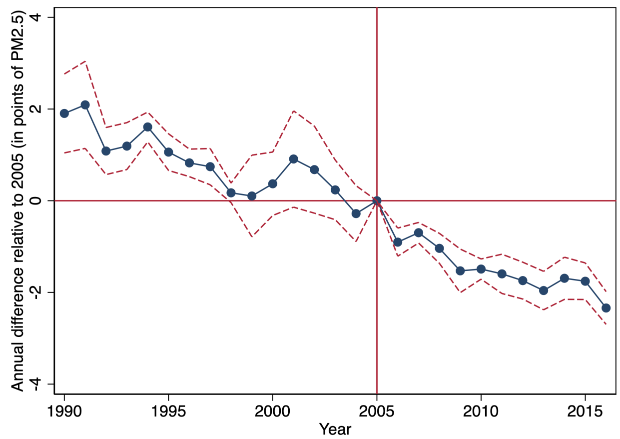
[*]Correspondence to: Lutz Sager: Lutz.Sager@georgetown.edu, Gregor Singer: g.a.singer@lse.ac.uk

A.1 Descriptive Statistics & Graphs

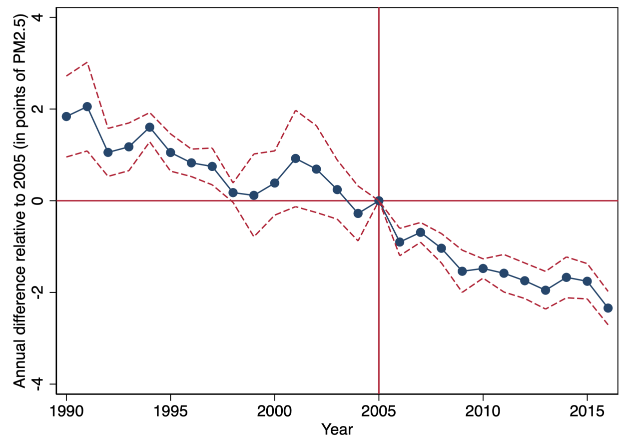
Table A.1: Descriptive Statistics

Panel (a): Baseline Period (2001-2003)				
	mean	sd	min	max
PM25 (Meng et al.)	11.98	3.33	1.35	24.79
PM25 (Di et al.)	12.58	3.31	2.51	29.58
PM25 (Van Donkelaar et al.)	12.19	3.23	3.68	27.30
Observations	72043			
Panel (b): Five Year Post Period (2006-2008)				
	mean	sd	min	max
PM25 (Meng et al.)	10.91	2.57	1.75	18.57
PM25 (Di et al.)	11.12	2.36	2.19	21.18
PM25 (Van Donkelaar et al.)	10.99	2.50	3.67	22.45
Observations	72043			
Panel (c): Ten Year Post Period (2011-2013)				
	mean	sd	min	max
PM25 (Meng et al.)	8.99	1.96	1.22	16.96
PM25 (Di et al.)	9.20	1.80	1.91	18.63
PM25 (Van Donkelaar et al.)	8.99	1.74	3.52	18.19
Observations	72043			

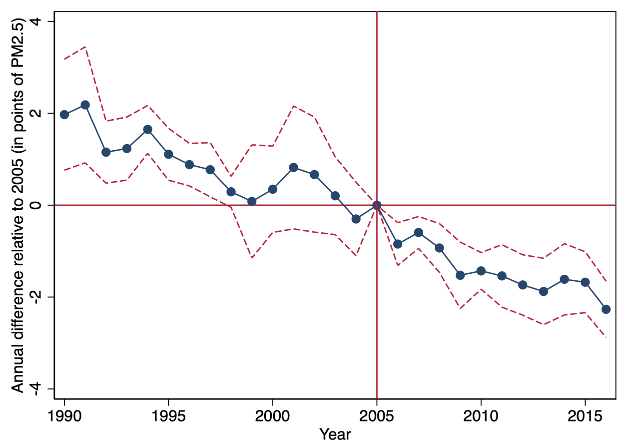
Notes: Tract level summary statistics, averaged over the respective 3-year periods, weighted by population weights accounting for population differences within tracts as well as across tracts. Pollution data is from Meng et al. (2019*b*), Di et al. (2021) and van Donkelaar et al. (2021*b*).



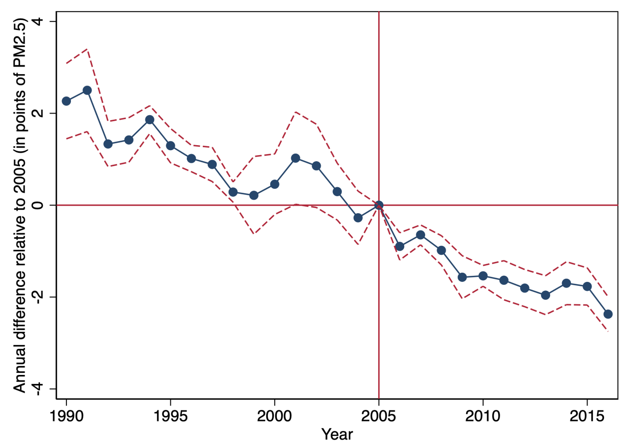
(a) 2000 block borders and block level analysis



(b) Interpolated population weights 2000, 2010, 2020



(c) Nonattainment based on commuting zones



(d) Dropping neighboring controls

Figure A.1: Robustness of differences in pre-trends in event study

Notes: The figure replicates the event study graph from Panel (b) of Figure 3. Panel (a) uses borders and population counts from the 2000 Census instead of the 2010 Census. In addition, the analysis is at the Census block level, rather than pre-aggregating to the Census tract level using Census block weights as in our main analysis (the results are equivalent using either). Panel (b) uses population weights that are interpolated between the 2000, 2010 and 2020 Census, using the IPUMS NHGIS crosswalk, instead of constant population weights at the 2010 level. Panel (c) assigns all counties in a commuting zone into nonattainment, as long as a single county in that commuting zone is in nonattainment, resulting in 428 nonattainment counties compared to the 208 actual nonattainment counties based on EPA air regions. Panel (d) drops all attainment counties that border a nonattainment county allowing for possible spatial spillovers (dropping 300 counties). All results based on Meng et al. (2019*b*), but the same patterns hold for data from Di et al. (2021) or van Donkelaar et al. (2021*b*). Standard errors are clustered at the county level except for Panel (c) where we cluster at the commuting zone level.

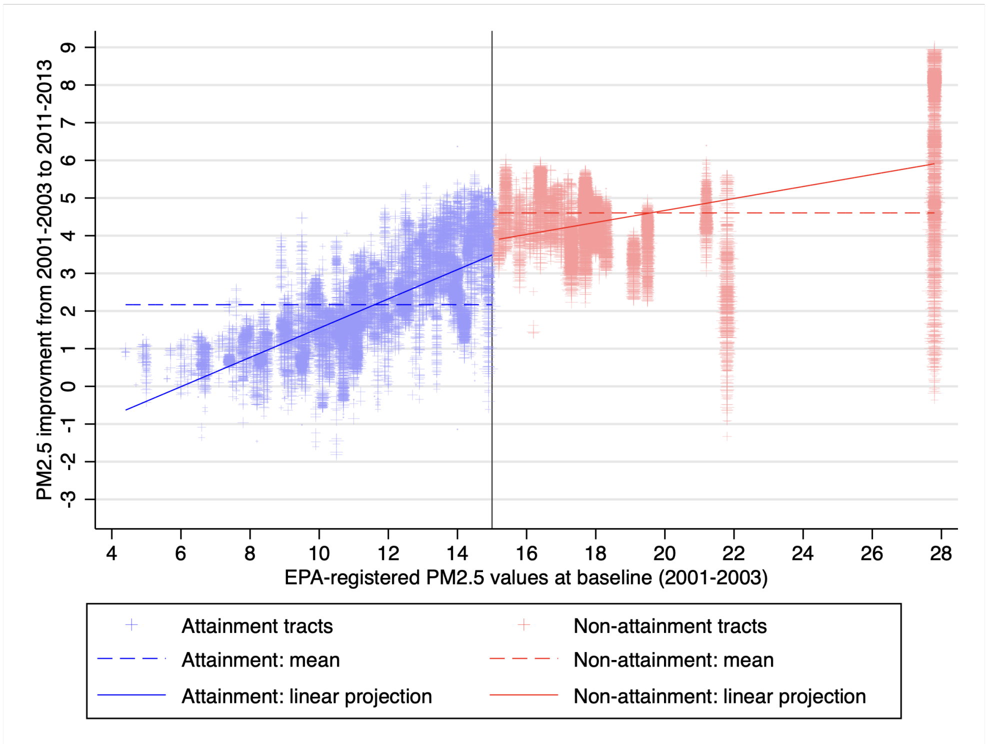


Figure A.2: Ten year improvement in tract $PM_{2.5}$ averages and EPA-registered $PM_{2.5}$ values

Notes: The figure shows the improvement in $PM_{2.5}$ averages at the tract level between two periods, 2001-2003 and 2011-2013. The size of the markers reflect tract level populations. The $PM_{2.5}$ improvements are plotted against the EPA-registered $PM_{2.5}$ values of each attainment/nonattainment area, each of which usually comprises multiple counties and tracts. The dashed line plots the average $PM_{2.5}$ improvement for tracts in nonattainment and attainment areas separately, weighted by tract population. The solid lines plot the linear projection of tract level $PM_{2.5}$ improvements on the EPA-registered $PM_{2.5}$ values of the nonattainment and attainment areas separately, weighted by tract population. Based on data from Meng et al. (2019*b*).

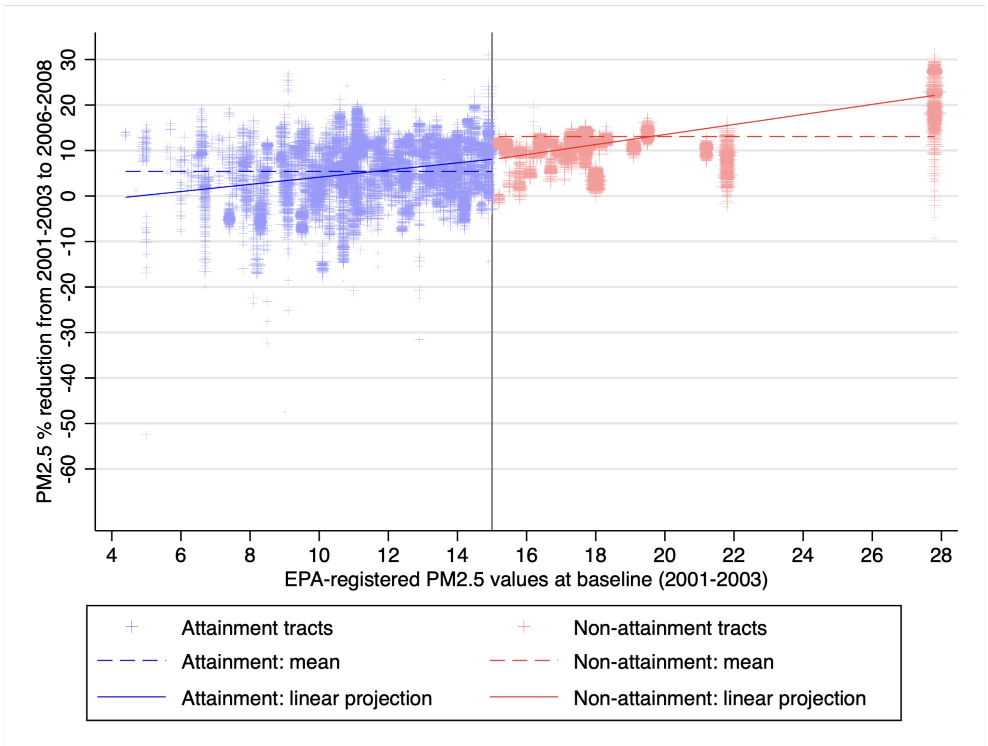


Figure A.3: Relative improvement in tract $PM_{2.5}$ averages and EPA-registered $PM_{2.5}$ values

Notes: The figure replicates Figure 4 but converts the vertical axis to percentage changes. It shows the improvement in $PM_{2.5}$ averages at the tract level between two periods, 2001-2003 and 2006-2008, expressed in percent of the 2001-2003 values. The size of the markers reflect tract level populations. The $PM_{2.5}$ improvements are plotted against the EPA-registered $PM_{2.5}$ values of each attainment/nonattainment area, each of which usually comprises multiple counties and tracts. The dashed line plots the average $PM_{2.5}$ improvement for tracts in nonattainment and attainment areas separately, weighted by tract population. The solid lines plot the linear projection of tract level $PM_{2.5}$ improvements on the EPA-registered $PM_{2.5}$ values of the nonattainment and attainment areas separately, weighted by tract population. Based on data from Meng et al. (2019*b*).

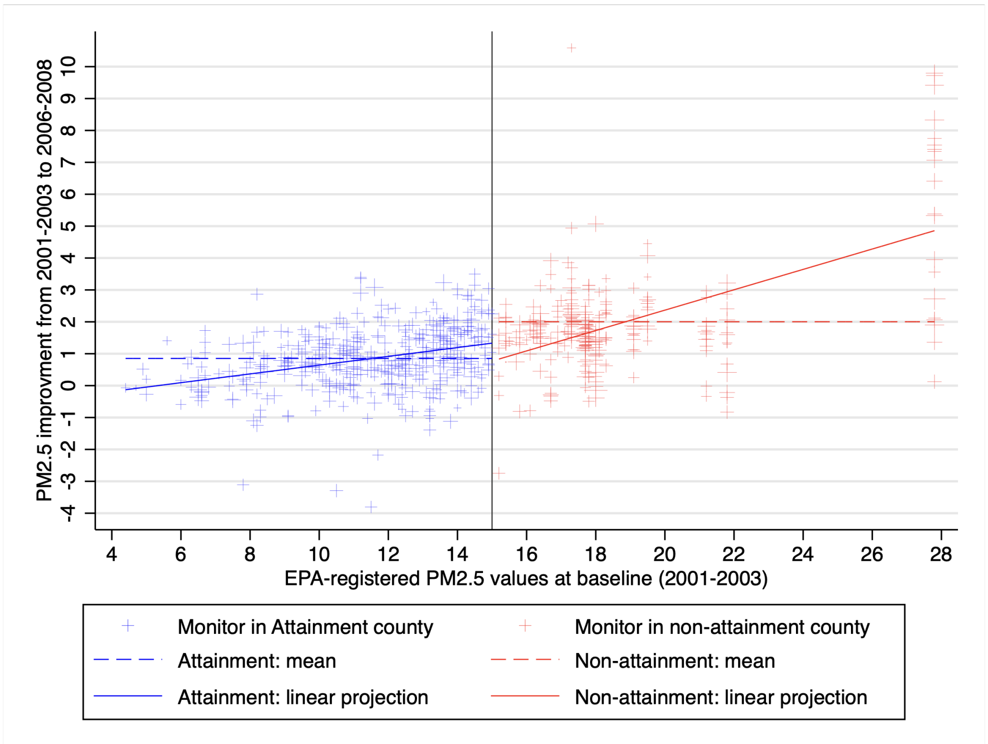


Figure A.4: Improvement in EPA monitor $PM_{2.5}$ averages and EPA-registered $PM_{2.5}$ values

Notes: The figure shows the improvement in $PM_{2.5}$ averages at the EPA monitor level between two periods, 2001-2003 and 2006-2008, taking the average $PM_{2.5}$ for each monitor. The size of the markers reflect tract level populations in which the monitor is situated. The $PM_{2.5}$ improvements are plotted against the EPA-registered $PM_{2.5}$ values of each attainment/nonattainment area, each of which usually comprises multiple counties and tracts. The dashed line plots the average $PM_{2.5}$ improvement for monitors in nonattainment and attainment areas separately, weighted by tract population. The solid lines plot the linear projection of monitor level $PM_{2.5}$ improvements on the EPA-registered $PM_{2.5}$ values of the nonattainment and attainment areas separately, weighted by tract population. Based on data from EPA (2022*a*).

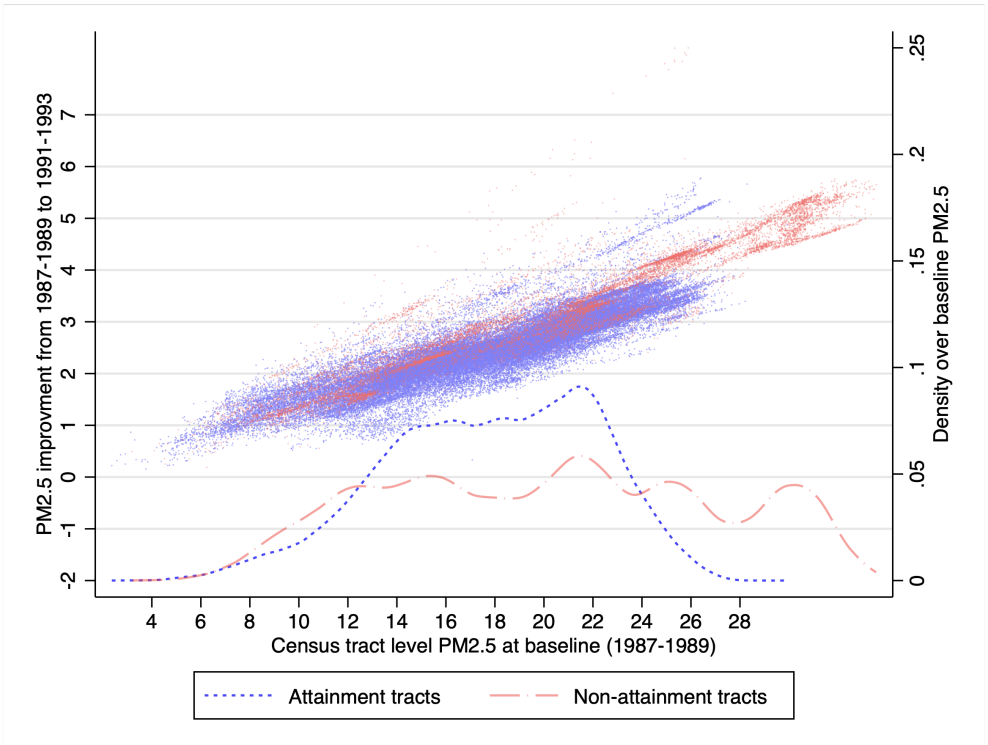


Figure A.5: Improvement in tract $PM_{2.5}$ averages and baseline $PM_{2.5}$ levels for the 1987 PM_{10} rules coming into effect in 1990

Notes: The markers in the figure show the improvement in $PM_{2.5}$ averages at the tract level between two periods, 1987-1989 and 1991-1993. The $PM_{2.5}$ improvements are plotted against the baseline $PM_{2.5}$ levels of each tract, using two different colors for tracts in nonattainment and attainment areas, based on the 1987 PM_{10} NAAQS EPA designations. The kernel density (right axis) shows the overlap between the baseline $PM_{2.5}$ distributions of nonattainment and attainment tracts, weighted by tract population. The figure is based on data from Meng et al. (2019*b*).

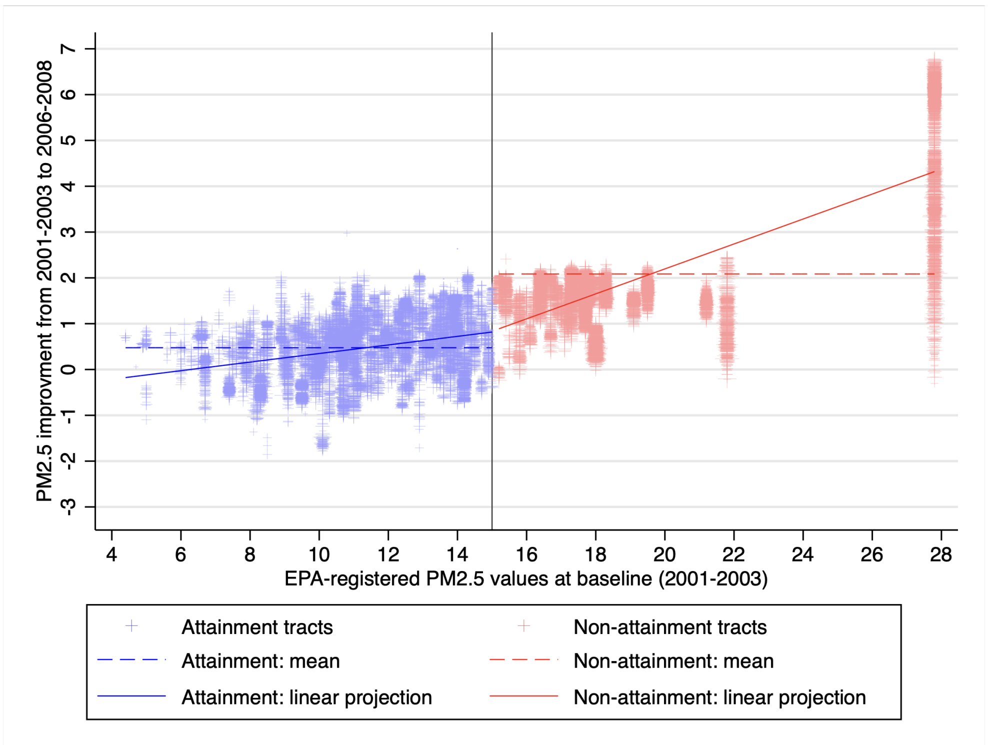


Figure A.6: Improvement in tract $PM_{2.5}$ averages and EPA-registered $PM_{2.5}$ values excluding attainment counties that share a border with a nonattainment county

Notes: The figure shows the improvement in $PM_{2.5}$ averages at the tract level between two periods, 2001-2003 and 2006-2008. Tracts in attainment counties that share a border with a nonattainment county are dropped. The size of the markers reflect tract level populations. The $PM_{2.5}$ improvements are plotted against the EPA-registered $PM_{2.5}$ values of each attainment/nonattainment area, each of which usually comprises multiple counties and tracts. The dashed line plots the average $PM_{2.5}$ improvement for tracts in nonattainment and attainment areas separately, weighted by tract population. The solid lines plot the linear projection of tract level $PM_{2.5}$ improvements on the EPA-registered $PM_{2.5}$ values of the nonattainment and attainment areas separately, weighted by tract population. Based on data from Meng et al. (2019*b*).

A.2 Difference in differences with different bandwidths or subsamples

Table A.2: Difference-in-differences estimates using different bandwidths or sub-samples using Chay & Greenstone (2005) approach.

	All tracts	with RV	Tracts with RV: 10-20	13-17	OBW	Binding w/ RV: 13-17	Daily RV: 28-63
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel (a): Change from 2001-03 to 2006-08							
Nonattainment	-1.47	-1.48	-0.83	-0.64	-0.36	-0.62	-0.72
	(0.34)	(0.35)	(0.093)	(0.18)	(0.28)	(0.22)	(0.094)
Observations	72043	47962	37366	12738	7026	10388	35820
Panel (b): Change from 2001-03 to 2011-13							
Nonattainment	-2.35	-2.44	-1.85	-1.48	-1.26	-1.45	-1.78
	(0.27)	(0.28)	(0.12)	(0.22)	(0.35)	(0.29)	(0.11)
Observations	72043	47962	37366	12738	6137	10388	35820

Notes: The table shows coefficient estimates from a simple difference-in-differences estimation following equation 2. Panel (a) uses average $PM_{2.5}$ across years 2006-2008 as post-treatment outcome. Panel (b) uses average $PM_{2.5}$ across years 2011-2013 as post-treatment outcome. Both use 2001-2003 as pre-treatment period. Column 1 uses full sample of tracts, Column 2 only those tracts for which EPA-registered $PM_{2.5}$ values are available, Column 3 only those tracts in a narrow window of these values around treatment cutoff ($10 < RV < 20$), Column 4 an even narrower window ($13 < RV < 17$), and Column 5 an optimal bandwidth as discussed in the section on regression discontinuity. Column 6 is the same as Column 4 but additionally restricts the treated counties to only contain those counties that have the highest EPA pollution readings within each nonattainment area and are therefore the binding counties that assign an area into nonattainment. Column 7 follows a strategy similar to Chay & Greenstone (2005), restricting the sample to areas in attainment of the daily standard and in the overlapping range of daily RV (28-63) shown in Figure 1b. Data from Meng et al. (2019*b*). Standard errors clustered at the county level in parentheses.

A.3 Difference in differences with baseline controls

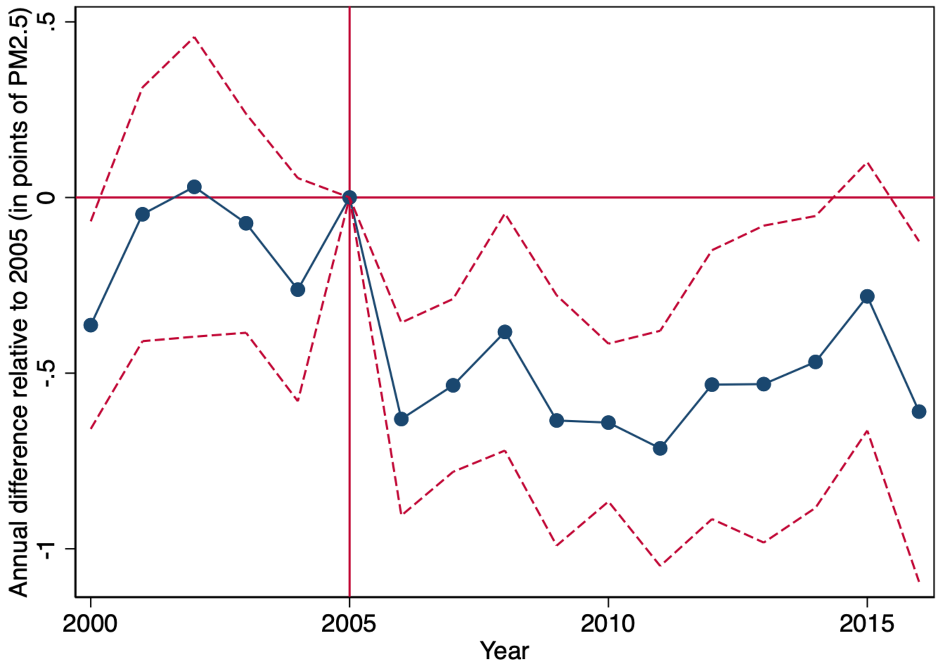


Figure A.7: Event study analysis with controls for baseline pollution

Notes: The Figure replicates the event study graph from Panel (b) of Figure 3 but controls for an interaction between time dummies and baseline pollution equivalent to Column 2 Table 1. All results based on Meng et al. (2019*b*).

Table A.3: Difference-in-differences estimates using different polynomials of baseline $PM_{2.5}$.

	Linear (1)	Quadratic (2)	Cubic (3)	Quartic (4)
Change from 2001-03 to 2006-08				
Nonattainment	-0.49	-0.41	-0.52	-0.51
	(0.098)	(0.070)	(0.072)	(0.071)
Observations	72043	72043	72043	72043
Panel (b): Change from 2001-03 to 2011-13				
Nonattainment	-0.56	-0.55	-0.52	-0.53
	(0.096)	(0.094)	(0.096)	(0.094)
Observations	72043	72043	72043	72043

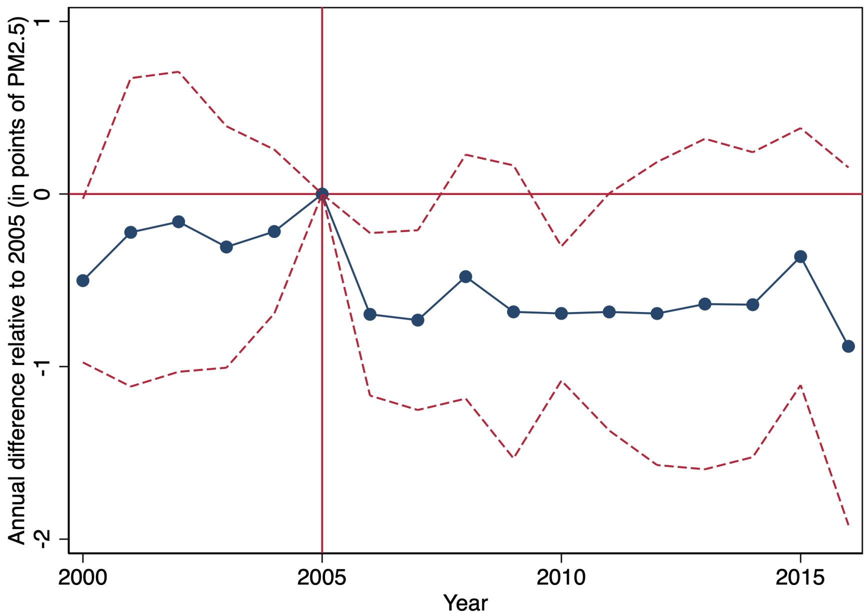
Notes: The table shows coefficient estimates from specifications with control for baseline $PM_{2.5}$ (DiDwb). Column 1 uses linear control and is identical to Column 2 of Table 1. Columns 2, 3 and 4 successively add quadratic, cubic and quartic terms. Data from Meng et al. (2019*b*). Standard errors clustered at the county level in parentheses.

A.4 Matching

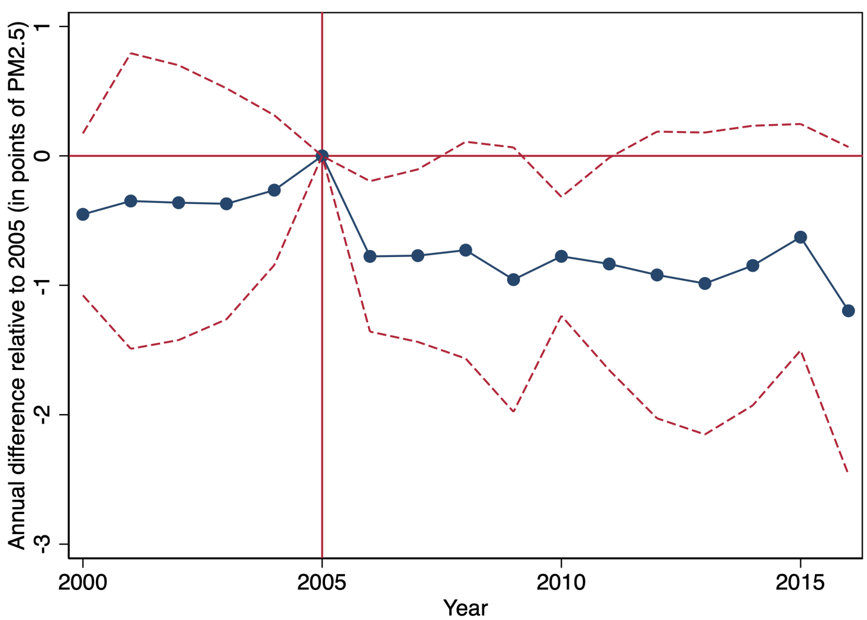
Table A.4: Matched samples - balance tests

	All Tracts		All Tracts with RV	
	unmatched	matched	unmatched	matched
	(1)	(2)	(3)	(4)
	M1: Matching on baseline $PM_{2.5}$			
Baseline PM2.5	0.95	-0.064	1.12	-0.063
	(0.20)	(0.27)	(0.25)	(0.31)
Population	182.7	229.0	195.0	267.9
	(88.4)	(111.8)	(102.7)	(137.5)
Pop. Density	7148.4	6582.5	5992.9	5684.0
	(2605.3)	(2645.6)	(2617.7)	(2654.7)
Observations	28291	28291	26647	26647
	M2: Matching on baseline $PM_{2.5}$, population, density			
Baseline PM2.5	1.25	0.054	1.15	-0.17
	(0.24)	(0.38)	(0.26)	(0.35)
Population	142.4	113.5	86.4	127.6
	(81.2)	(117.7)	(119.0)	(132.0)
Pop. Density	5530.8	-1821.7	4791.2	3854.1
	(2242.0)	(5803.9)	(2396.8)	(2463.8)
Observations	28909	28909	26637	26637

Notes: The table shows comparisons of average pre-treatment differences before and after our matching procedure. Shown are population-weighted average differences between nonattainment tracts and attainment tracts for baseline $PM_{2.5}$ (2001-03), baseline population (2000), and baseline population density (2000), without using matching weights (*unmatched*) and with using matching weights (*matched*). Matching approach M1 is the same as used in Columns 3 and 7 of Table 1, M2 is the same as in Columns 4 and 8. Standard errors in parentheses are cluster-robust at the level of counties. All results based on Meng et al. (2019*b*).



(a) M1: Matching on baseline $PM_{2.5}$



(b) M2: Matching on baseline $PM_{2.5}$, population, density

Figure A.8: Matched samples - event study analysis

Notes: The figure replicates the event study graph from Panel (b) of Figure 3 with the matched sample and weights underlying the matched difference-in-differences estimation shown in Columns 3 and 4 of Table 1. All results based on Meng et al. (2019*b*).

A.5 Regression discontinuity

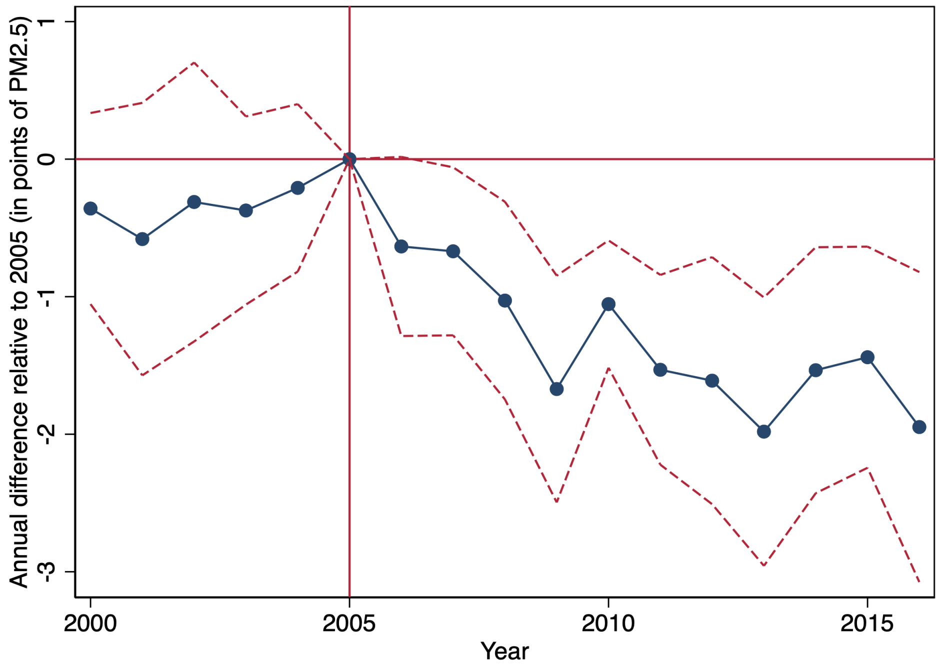
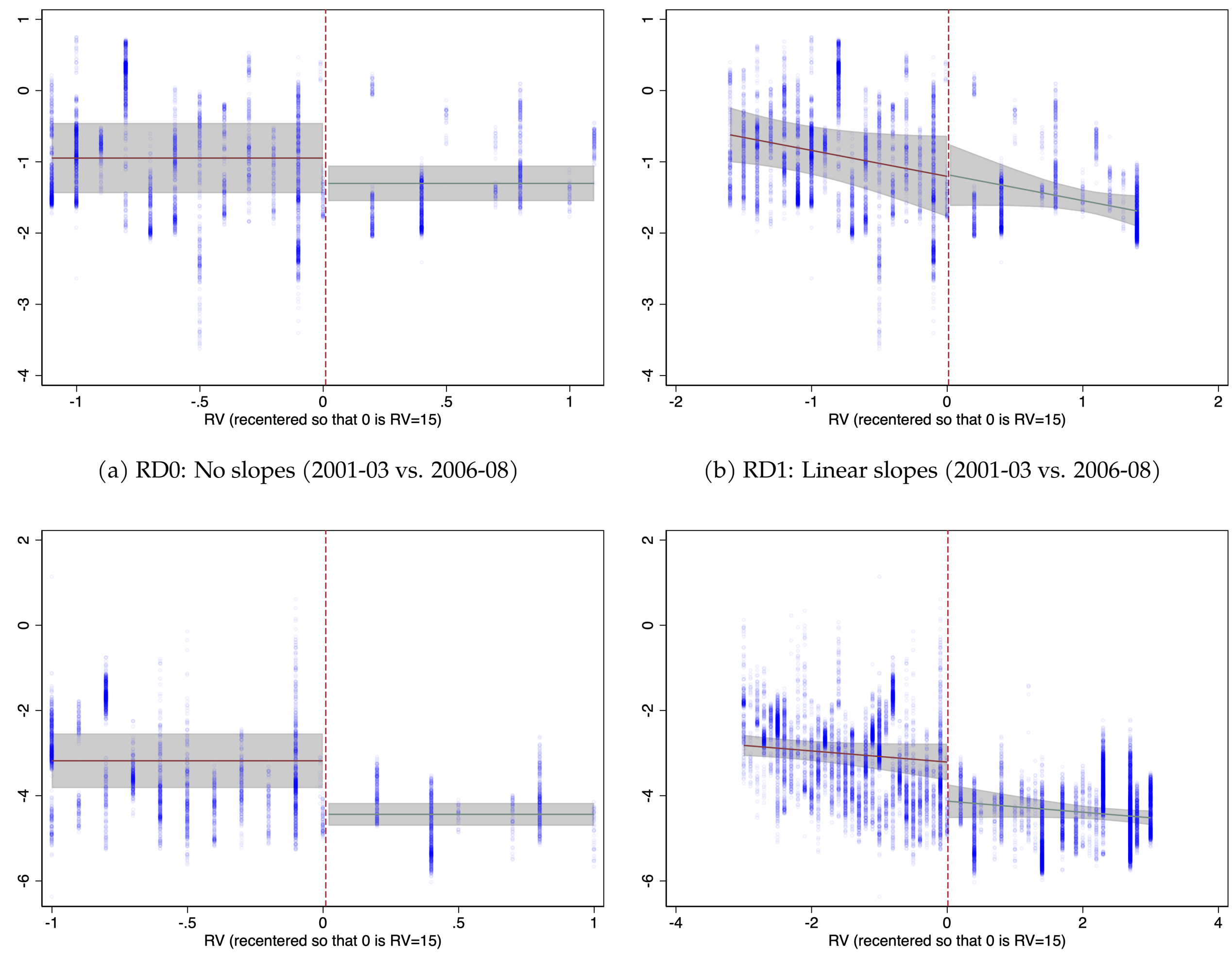


Figure A.9: Regression discontinuity analysis (RD0) - event study analysis

Notes: The figure replicates the event study graph from Panel (b) of Figure 3 with the restricted sample underlying the regression discontinuity estimation shown in Column 6 Table 1. All results based on Meng et al. (2019*b*).



(a) RD0: No slopes (2001-03 vs. 2006-08)

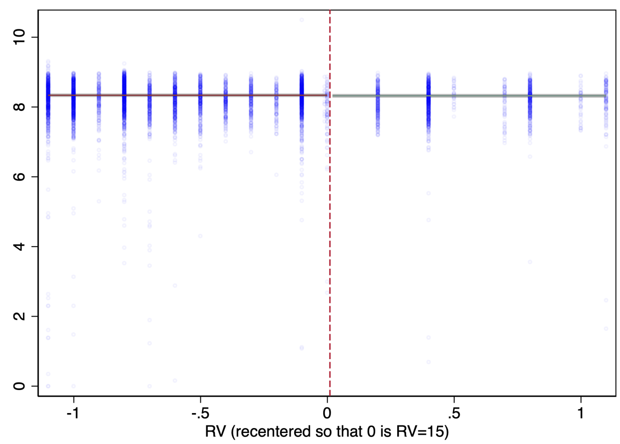
(b) RD1: Linear slopes (2001-03 vs. 2006-08)

(c) RD0: No slopes (2001-03 vs. 2011-13)

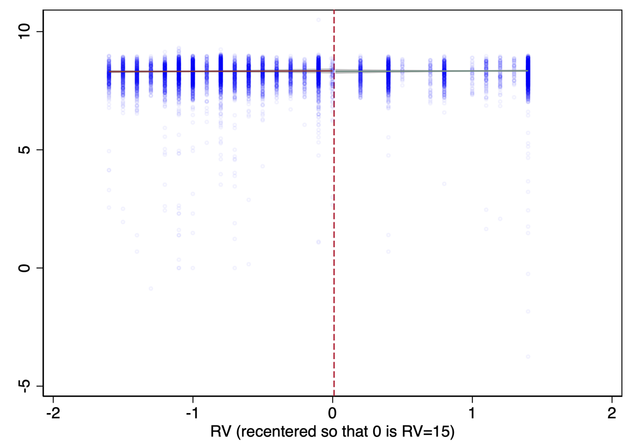
(d) RD1: Linear slopes (2001-03 vs. 2011-13)

Figure A.10: Regression discontinuity – change in $PM_{2.5}$

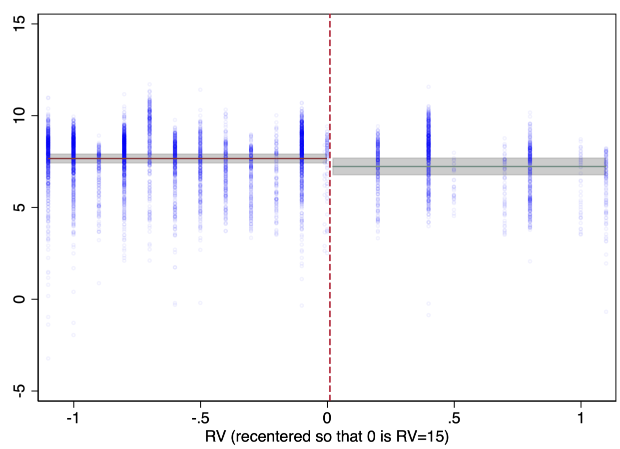
Notes: The figures show visual representations of the regression discontinuity approaches without slopes (RD0) and with linear trends (RD1). Lines are model predictions, shaded areas represent 95% confidence intervals. Horizontal axis shows EPA-registered $PM_{2.5}$ values, re-centered so that 0 is the cutoff (RV=15), and narrowed to the optimal bandwidth. Vertical axis shows the change in $PM_{2.5}$ from 2001-03 to 2006-08 in Panels (a) and (b) or from 2001-03 to 2011-13 in Panels (c) and (d). The jump between the fitted lines at the cutoff is equal to the coefficient estimate in Columns 9 and 10 of Table 1. The figure is based on data from Meng et al. (2019*b*).



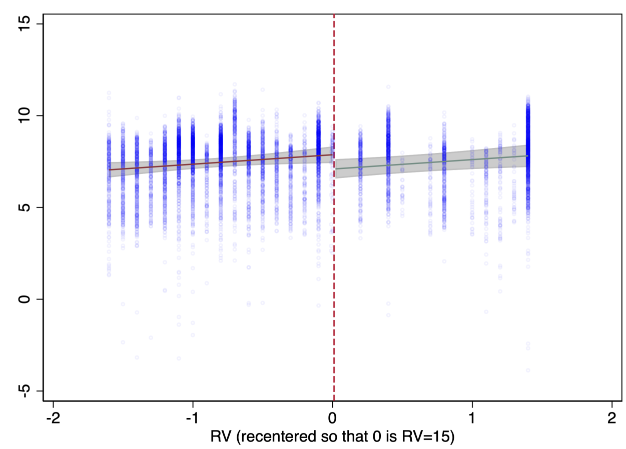
(a) RD0: No slopes (log population)



(b) RD1: Linear slopes (log population)



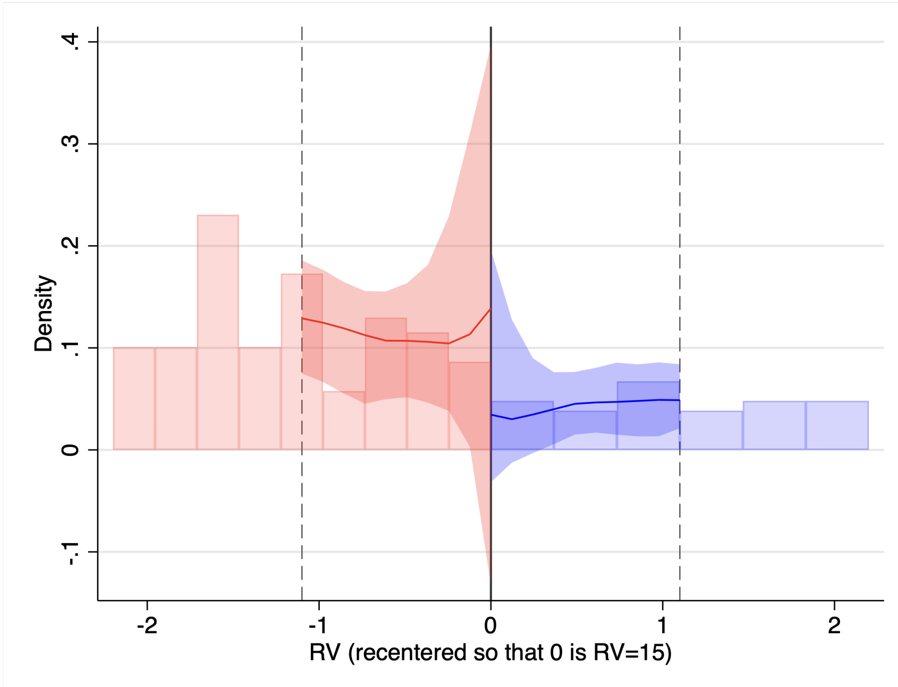
(c) RD0: No slopes (log population density)



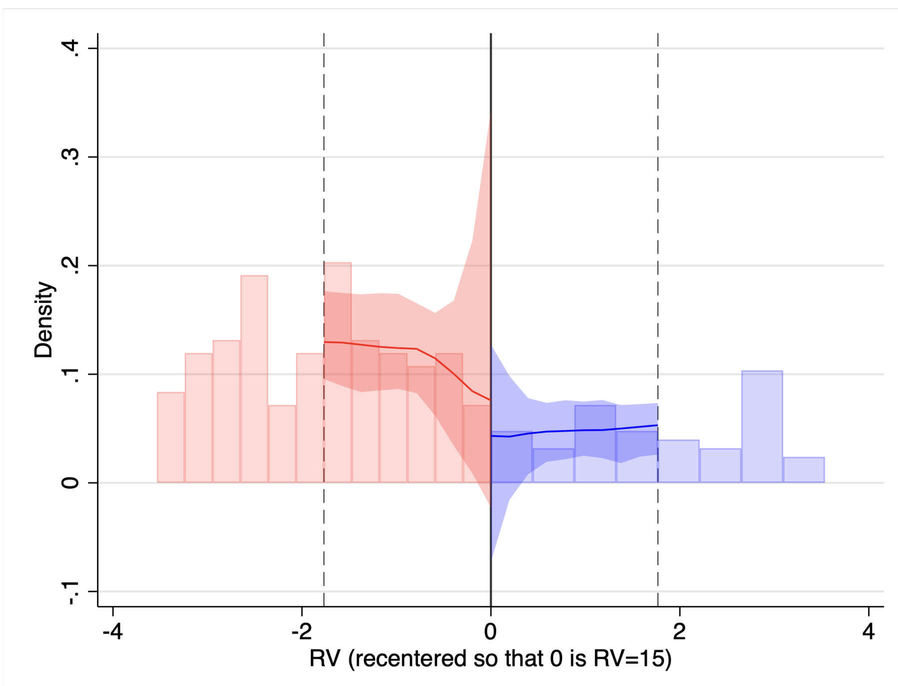
(d) RD1: Linear slopes (log population density)

Figure A.11: Continuity of predetermined population and population density

Notes: The figures show tests for continuity of predetermined covariates applied to the regression discontinuity approaches without slopes (RD0) and with linear trends (RD1). Panels (a) and (b) use 2000 population (log) and Panels (b) and (c) use 2000 population density (log). Lines are model predictions, shaded areas represent 95% confidence intervals. The horizontal axis shows EPA-registered $PM_{2.5}$ values, re-centered so that 0 is the cutoff (RV=15), and narrowed to the optimal bandwidth. Bandwidth selection in this figure is based on data from Meng et al. (2019*b*).



(a) RD0 optimal bandwidth



(b) RD1 optimal bandwidth

Figure A.12: Regression discontinuity – manipulation around threshold

Notes: The figures show tests for manipulation around threshold for regression discontinuity approaches without slopes (RD0) and with linear trends (RD1). The horizontal axis shows EPA-registered $PM_{2.5}$ values at the level of commuting zones, re-centered so that 0 is the cutoff (RV=15), and narrowed to twice the optimal bandwidth. Vertical axis shows density of RV values, in absolute (histogram) and polynomial approximation within the optimal bandwidth (dashed lines) following Calonico et al. (2014) and Calonico et al. (2020). A large discontinuous and significant jump between the fitted lines at the cutoff would indicate manipulation around the threshold. The figure is based on data from Meng et al. (2019*b*).

A.6 Testing for differences in coefficients across models

Table A.5: Testing for equality of coefficients across models

	DiD	DiDwb	M1DiD	M2DiD	RD0
DiDwb	0				
M1	0	0.410			
M2	0	0.424	0.463		
RD0	0.005	0.686	0.867	0.968	
RD1	0.127	0.790	0.919	0.967	0.937

Notes: The table shows p-values from two-sided tests of pairwise equality of coefficients corresponding to Panel (a) in Table 1. To calculate p-values, we cluster-bootstrap estimates for our different estimators. We draw counties (allowing for clustering) with replacement based on two strata (attainment and nonattainment), estimate the different models, and repeat the process 10,000 times and calculate the difference in coefficients for each bootstrap. The p-values correspond to the share of runs where the difference has the opposite sign as the difference in Table 1, multiplied by two to allow for two-sided testing.

A.7 Nonattainment status and changes in $PM_{2.5}$ – Robustness

Table A.6: Nonattainment status and changes in $PM_{2.5}$: *Dropping attainment counties with neighboring county in nonattainment*

	ATT					LATE	
	All Tracts				with RV	Optimal Bandw.	
	DiD	DiDwb	M1DiD	M2DiD	DiD	RD0	RD1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Part A: Effect from 2001-03 to 2006-08							
Panel (a): Homogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-1.57	-0.50	-0.72	-0.67	-1.61	-0.64	-0.77
	(0.34)	(0.15)	(0.080)	(0.100)	(0.35)	(0.29)	(0.40)
Observations	64516	64516	26130	26664	43523	5441	10194
Panel (b): Placebo Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-0.22	-0.10	-0.028	0.037	-0.39	-0.037	-0.25
	(0.12)	(0.12)	(0.12)	(0.13)	(0.14)	(0.21)	(0.34)
Observations	41830	41830	15411	15145	20837	2341	6123
Panel (c): Heterogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	4.72	4.13	1.08	1.93	4.68	3.51	2.80
	(0.81)	(0.83)	(0.23)	(0.33)	(0.82)	(0.76)	(0.61)
NA(x)Baseline	-0.42	-0.36	-0.13	-0.18	-0.42	-0.29	-0.25
	(0.060)	(0.061)	(0.016)	(0.025)	(0.060)	(0.047)	(0.030)
Observations	64516	64516	26130	26664	43523	5441	10194
Implied ATE	-1.57	-1.29	-0.85	-0.83	-1.61	-0.85	-0.98
10th pct	-0.41	-0.30	-0.50	-0.32	-0.46	-0.052	-0.28
90th pct	-3.66	-3.10	-1.50	-1.75	-3.70	-2.30	-2.23
Part B: Effect from 2001-03 to 2011-13							
Panel (d): Homogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-2.52	-0.64	-0.43	-0.50	-2.60	-1.51	-1.78
	(0.27)	(0.11)	(0.13)	(0.14)	(0.28)	(0.44)	(0.61)
Observations	64516	64516	26130	26664	43523	4562	13442
Panel (e): Placebo Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-0.98	-0.050	0.17	0.15	-1.53	0.15	0.26
	(0.15)	(0.14)	(0.16)	(0.16)	(0.16)	(0.19)	(0.28)
Observations	41830	41830	15411	15145	20837	1204	3375
Panel (f): Heterogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	3.74	-0.13	5.06	4.71	3.66	4.20	3.18
	(0.41)	(0.43)	(0.51)	(0.44)	(0.41)	(0.83)	(1.01)
NA(x)Baseline	-0.42	-0.039	-0.39	-0.37	-0.42	-0.40	-0.34
	(0.029)	(0.032)	(0.038)	(0.032)	(0.029)	(0.053)	(0.053)
Observations	64516	64516	26130	26664	43523	4562	13442
Implied ATE	-2.52	-0.72	-0.82	-0.82	-2.60	-1.80	-1.95
10th pct	-1.37	-0.61	0.26	0.20	-1.45	-0.70	-1.01
90th pct	-4.61	-0.92	-2.77	-2.66	-4.69	-3.79	-3.66

Notes: The table shows coefficient estimates for the treatment effect of nonattainment status on the change in $PM_{2.5}$ levels between the pre- and post-treatment periods. All estimations exclude counties that were in attainment but had a neighboring county in nonattainment. Each panel(x)column combination is from a separate regression as described in the text. (1) uses simple DiD, (2) adds controls for baseline $PM_{2.5}$ (2001-03), (3) runs DiD using a sample matched (1-to-1) on baseline $PM_{2.5}$, (4) matches on baseline $PM_{2.5}$, tract population and population density (both 2000), (5) again uses simple DiD but with the limited sample of areas for which an EPA-registered $PM_{2.5}$ value exists, (6) and (7) use the limited sample based on optimal bandwidth selection in a regression discontinuity framework. Standard errors in parentheses are clustered at the county level. All results based on Meng et al. (2019*b*).

Table A.7: Nonattainment status and changes in $PM_{2.5}$: *Dropping PM_{10} nonattainment counties*

	ATT					LATE	
	All Tracts				with RV	Optimal Bandw.	
	DiD	DiDwb	M1DiD	M2DiD	DiD	RD0	RD1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Part A: Effect from 2001-03 to 2006-08							
Panel (a): Homogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-0.91	-0.59	-0.39	-0.27	-0.90	-0.64	0.048
	(0.059)	(0.083)	(0.12)	(0.16)	(0.092)	(0.22)	(0.39)
Observations	60978	60978	22889	22763	37507	9920	9559
Panel (b): Placebo Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-0.33	-0.13	-0.056	0.026	-0.51	-0.053	-0.22
	(0.13)	(0.12)	(0.13)	(0.13)	(0.16)	(0.18)	(0.31)
Observations	45114	45114	18904	18662	21643	3229	4447
Panel (c): Heterogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	0.65	-0.33	1.16	1.29	0.65	3.39	3.94
	(0.28)	(0.32)	(0.30)	(0.31)	(0.29)	(0.59)	(0.62)
NA(x)Baseline	-0.11	-0.019	-0.11	-0.11	-0.11	-0.28	-0.27
	(0.020)	(0.026)	(0.020)	(0.020)	(0.020)	(0.036)	(0.032)
Observations	60978	60978	22889	22763	37507	9920	9559
Implied ATE	-1.01	-0.61	-0.49	-0.36	-1.00	-0.78	-0.14
10th pct	-0.70	-0.56	-0.18	-0.061	-0.70	-0.010	0.61
90th pct	-1.56	-0.71	-1.04	-0.91	-1.55	-2.16	-1.50
Part B: Effect from 2001-03 to 2011-13							
Panel (d): Homogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-2.02	-0.70	-0.55	-0.57	-2.07	-0.79	-1.03
	(0.078)	(0.085)	(0.088)	(0.094)	(0.11)	(0.27)	(0.39)
Observations	60978	60978	22889	22763	37507	3712	12664
Panel (e): Placebo Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-0.93	0.039	0.22	0.15	-1.59	0.090	0.67
	(0.14)	(0.13)	(0.14)	(0.15)	(0.16)	(0.16)	(0.34)
Observations	45114	45114	18904	18662	21643	2133	3932
Panel (f): Heterogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	3.20	-0.98	4.68	4.66	3.16	5.58	4.58
	(0.33)	(0.35)	(0.33)	(0.33)	(0.34)	(0.80)	(0.71)
NA(x)Baseline	-0.37	0.020	-0.37	-0.37	-0.37	-0.43	-0.39
	(0.024)	(0.028)	(0.024)	(0.024)	(0.024)	(0.058)	(0.042)
Observations	60978	60978	22889	22763	37507	3712	12664
Implied ATE	-2.36	-0.68	-0.88	-0.90	-2.40	-0.95	-1.29
10th pct	-1.34	-0.73	0.14	0.12	-1.38	0.25	-0.21
90th pct	-4.21	-0.58	-2.74	-2.75	-4.25	-3.12	-3.24

Notes: The table shows coefficient estimates for the treatment effect of nonattainment status on the change in $PM_{2.5}$ levels between the pre- and post-treatment periods. All estimations exclude areas that were in nonattainment of the PM_{10} NAAQS between 2001-04. Each panel(x)column combination is from a separate regression as described in the text. (1) uses simple DiD, (2) adds controls for baseline $PM_{2.5}$ (2001-03), (3) runs DiD using a sample matched (1-to-1) on baseline $PM_{2.5}$, (4) matches on baseline $PM_{2.5}$, tract population and population density (both 2000), (5) again uses simple DiD but with the limited sample of areas for which an EPA-registered $PM_{2.5}$ values exists, (6) and (7) use the limited sample based on optimal bandwidth selection in a regression discontinuity framework. Standard errors in parentheses are clustered at the county level. All results based on Meng et al. (2019*b*).

Table A.8: Nonattainment status and changes in $PM_{2.5}$: *with flexible state time trends*

	ATT					LATE	
	All Tracts				with RV	Optimal Bandw.	
	DiD	DiDwb	M1DiD	M2DiD	DiD	RD0	RD1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Part A: Effect from 2001-03 to 2006-08							
Panel (a): Homogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-1.11	-0.22	-0.13	-0.31	-1.36	-0.073	0.074
	(0.26)	(0.065)	(0.080)	(0.14)	(0.37)	(0.048)	(0.16)
Observations	72043	72043	28290	28908	47962	7026	10459
Panel (b): Placebo Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-0.20	0.0054	-0.018	0.020	-0.47	0.038	0.076
	(0.098)	(0.087)	(0.073)	(0.083)	(0.14)	(0.15)	(0.20)
Observations	49357	49357	20388	20127	25276	2143	5411
Panel (c): Heterogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	4.76	3.01	2.67	3.45	4.65	2.08	2.57
	(0.59)	(0.57)	(0.22)	(0.40)	(0.54)	(0.18)	(0.19)
NA(x)Baseline	-0.39	-0.24	-0.20	-0.26	-0.39	-0.16	-0.18
	(0.043)	(0.042)	(0.016)	(0.031)	(0.041)	(0.014)	(0.0091)
Observations	72043	72043	28290	28908	47962	7026	10459
Implied ATE	-1.09	-0.61	-0.29	-0.40	-1.16	-0.26	-0.21
10th pct	-0.016	0.053	0.25	0.31	-0.089	0.17	0.30
90th pct	-3.04	-1.82	-1.28	-1.68	-3.09	-1.04	-1.13
Part B: Effect from 2001-03 to 2011-13							
Panel (d): Homogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-1.56	-0.26	-0.097	-0.27	-1.76	0.18	-0.31
	(0.27)	(0.058)	(0.087)	(0.15)	(0.40)	(0.069)	(0.15)
Observations	72043	72043	28290	28908	47962	6137	25856
Panel (e): Placebo Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-0.43	0.017	0.0046	0.040	-0.78	-0.058	0.43
	(0.065)	(0.050)	(0.055)	(0.065)	(0.12)	(0.052)	(0.23)
Observations	49357	49357	20388	20127	25276	1046	4626
Panel (f): Heterogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	5.57	1.83	5.12	5.42	5.68	5.51	4.47
	(0.45)	(0.41)	(0.23)	(0.38)	(0.41)	(0.39)	(0.31)
NA(x)Baseline	-0.47	-0.16	-0.37	-0.39	-0.48	-0.39	-0.34
	(0.034)	(0.031)	(0.016)	(0.030)	(0.031)	(0.030)	(0.020)
Observations	72043	72043	28290	28908	47962	6137	25856
Implied ATE	-1.54	-0.51	-0.40	-0.41	-1.51	-0.28	-0.58
10th pct	-0.23	-0.084	0.61	0.67	-0.19	0.78	0.35
90th pct	-3.90	-1.29	-2.24	-2.34	-3.90	-2.21	-2.27

Notes: The table shows coefficient estimates for the treatment effect of nonattainment status on the change in $PM_{2.5}$ levels between the pre- and post-treatment periods. All estimations include state fixed effects. Each panel(x)column combination is from a separate regression as described in the text. (1) uses simple DiD, (2) adds controls for baseline $PM_{2.5}$ (2001-03), (3) runs DiD using a sample matched (1-to-1) on baseline $PM_{2.5}$, (4) matches on baseline $PM_{2.5}$, tract population and population density (both 2000), (5) again uses simple DiD but with the limited sample of areas for which an EPA-registered $PM_{2.5}$ value exists, (6) and (7) use the limited sample based on optimal bandwidth selection in a regression discontinuity framework. Standard errors in parentheses are clustered at the county level. All results based on Meng et al. (2019*b*).

Table A.9: Nonattainment status and changes in $PM_{2.5}$: *with flexible state time trends and quartile of density time trends*

	ATT					LATE	
	All Tracts				with RV	Optimal Bandw.	
	DiD	DiDwb	M1DiD	M2DiD	DiD	RD0	RD1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Part A: Effect from 2001-03 to 2006-08							
Panel (a): Homogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-0.97	-0.23	-0.14	-0.35	-1.27	-0.11	-0.042
	(0.24)	(0.062)	(0.079)	(0.14)	(0.34)	(0.047)	(0.14)
Observations	71951	71951	28264	28882	47881	7021	10451
Panel (b): Placebo Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-0.12	0.00093	0.026	0.049	-0.44	0.057	0.076
	(0.097)	(0.088)	(0.072)	(0.083)	(0.14)	(0.15)	(0.19)
Observations	49289	49289	20373	20127	25219	2141	5396
Panel (c): Heterogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	4.74	2.81	2.36	3.41	4.66	1.49	1.63
	(0.61)	(0.51)	(0.30)	(0.51)	(0.57)	(0.32)	(0.32)
NA(x)Baseline	-0.39	-0.23	-0.18	-0.25	-0.39	-0.12	-0.12
	(0.046)	(0.038)	(0.022)	(0.039)	(0.043)	(0.023)	(0.020)
Observations	71951	71951	28264	28882	47881	7021	10451
Implied ATE	-1.06	-0.59	-0.27	-0.39	-1.15	-0.24	-0.20
10th pct	0.0036	0.032	0.21	0.30	-0.086	0.078	0.14
90th pct	-2.99	-1.72	-1.15	-1.66	-3.09	-0.82	-0.80
Part B: Effect from 2001-03 to 2011-13							
Panel (d): Homogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-1.30	-0.27	-0.11	-0.34	-1.61	0.089	-0.34
	(0.25)	(0.056)	(0.082)	(0.15)	(0.35)	(0.074)	(0.14)
Observations	71951	71951	28264	28882	47881	6133	25831
Panel (e): Placebo Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-0.29	-0.0060	0.093	0.098	-0.72	-0.15	0.45
	(0.066)	(0.050)	(0.055)	(0.065)	(0.12)	(0.084)	(0.25)
Observations	49289	49289	20373	20127	25219	1044	4612
Panel (f): Heterogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	5.38	1.61	4.52	5.04	5.39	4.01	3.37
	(0.52)	(0.36)	(0.29)	(0.49)	(0.48)	(0.68)	(0.34)
NA(x)Baseline	-0.45	-0.14	-0.33	-0.36	-0.46	-0.28	-0.26
	(0.039)	(0.027)	(0.021)	(0.037)	(0.036)	(0.049)	(0.022)
Observations	71951	71951	28264	28882	47881	6133	25831
Implied ATE	-1.41	-0.49	-0.37	-0.41	-1.48	-0.23	-0.54
10th pct	-0.16	-0.10	0.53	0.59	-0.22	0.55	0.18
90th pct	-3.67	-1.19	-1.99	-2.22	-3.77	-1.64	-1.84

Notes: The table shows coefficient estimates for the treatment effect of nonattainment status on the change in $PM_{2.5}$ levels between the pre- and post-treatment periods. All estimations include state fixed effects and tract population density quartile fixed effects. Each panel(x)column combination is from a separate regression as described in the text. (1) uses simple DiD, (2) adds controls for baseline $PM_{2.5}$ (2001-03), (3) runs DiD using a sample matched (1-to-1) on baseline $PM_{2.5}$, (4) matches on baseline $PM_{2.5}$, tract population and population density (both 2000), (5) again uses simple DiD but with the limited sample of areas for which an EPA-registered $PM_{2.5}$ value exists, (6) and (7) use the limited sample based on optimal bandwidth selection in a regression discontinuity framework. Standard errors in parentheses are clustered at the county level. All results based on Meng et al. (2019*b*).

Table A.10: TWFE estimates in tract-year panel (2000-2015)

	ATT					LATE	
	All Tracts				with RV	Optimal Bandw.	
	DiD	DiDwb	M1DiD	M2DiD	DiD	RD0	RD1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
NA Effect	-1.675	-0.544	-0.469	-0.855	-1.719	-0.794	-0.727
	(0.254)	(0.069)	(0.073)	(0.262)	(0.260)	(0.268)	(0.208)
Observations	1152688	1152688	488400	516464	767392	98192	180448

Notes: The table shows results from a panel regression with two-way fixed effects (TWFE) with a homogeneous treatment effect from nonattainment designations, equivalent to Panels (a) and (d) of Table 1. Data used is a tract-year panel from 2000 to 2015. Estimation includes tract and year fixed effects. Standard errors clustered at the county level in parentheses. All results based on Meng et al. (2019*b*).

A.8 NBP and CAIR as potential confounders

Since nonattainment designations under the NAAQS for $PM_{2.5}$ are not the only air quality policies during our sample period, we may worry about mis-attributing changes in air quality to $PM_{2.5}$ nonattainment designations if those other air quality policies are correlated with nonattainment designations. Specifically, during our sample period, the NO_x Budget Trading Program (NBP) (discussed in e.g. Deschenes et al. 2017, Curtis 2018) and its' successor, the Clean Air Interstate Rule (CAIR) were implemented. The NBP was a cap-and-trade program enacted by twenty eastern states plus DC in 2003/2004 that targeted NO_x emissions from power plants and other large stationary sources. In 2009, the NBP was replaced by the CAIR which expanded geographic coverage and targeted SO_2 and Ozone emissions in addition to NO_x.

As NO_x and SO_2 are precursors to $PM_{2.5}$, and these policies coincide with our study period, we next verify that these are not partially driving our results. To do so, we collect data on all facilities subject to regulation under the NBP and CAIR from the EPA's Clean Air Markets Data Program Facility Attributes Table (EPA 2023). We generate a binary variable indicating if a county contained a facility that became subject to NBP (for the 5-year period ending in 2006-08) or CAIR (for the 10-year period ending in 2011-13), and include this indicator as a control variable in our DiD and DiDwb regressions. Part A of Table A.11 shows the results for NBP and CAIR respectively in Panel (a) and (b). Column 1 and 3 reproduce our baseline coefficients from Table 1. Columns 2 and 4 show that controlling for these programs leaves our $PM_{2.5}$ nonattainment coefficients from Column 1 and 3 largely unchanged, indicating that any potential bias from these other policies is likely minimal.[60]

Instead of testing robustness of our effects to NBP/CAIR controls, we next analyze in Part B of Table A.11 whether the NBP/CAIR estimates suffer from similar bias when ignoring possibly confounding trends in $PM_{2.5}$. First, in Panel (c), we estimate the effect of NBP or CAIR using our $PM_{2.5}$ data. While the estimates in Columns 1 and 2 suggest substantial $PM_{2.5}$ reduction effects, at least for NBP, this effect disappears when controlling for baseline levels of pollution.[61] Second, in Panel (d), we use the data and code from Deschenes et al. (2017*b*) to first replicate their results in Column 1 and 2 (corresponding to their Table 2b Row 9 Columns 4 and 5 – see table notes for details). Once we control for trends based on baseline $PM_{2.5}$, the estimated coefficients fall in corresponding Columns 3 and 4, mirroring the pattern using our data in Panel (c). Deschenes et al. (2017) note that their effects on $PM_{2.5}$ are inconclusive. Once we control for trends, the results are even closer to zero, in line with our main findings. Finally, in Panel (e) we again replicate the results from

[60]The results look almost identical when we control for NBP/CAIR participation at the state level to account for possible spillovers. These results are available from the authors upon request.

[61]Note also that the NBP and CAIR effects from Panel (c) Columns 1 and 2 also disappear when we control for $PM_{2.5}$ nonattainment in Panel (a) and (b) Column 2.

Deschenes et al. (2017) in Columns 1 and 2 but focus on Ozone concentrations instead of $PM_{2.5}$ (corresponding to their Table 2b Row 4 Columns 4 and 5). Interestingly, controlling for trends based on baseline Ozone reduces estimated coefficients only by a small amount, largely confirming the results in Deschenes et al. (2017). This suggest that confounding trends in $PM_{2.5}$ may be particularly severe.

Table A.11: Controlling for potential confounding by contemporaneous policy changes and exploring NBP/CAIR effects

	(DiD) (1)	(DiD) (2)	(DiDwb) (3)	(DiDwb) (4)
Part A: Robustness controlling for NBP or CAIR				
Panel (a): NBP enacted 2003/04 (period ending 2006-08)				
Nonattainment	-1.47	-1.55	-0.49	-0.56
	(0.34)	(0.48)	(0.098)	(0.095)
NBP		0.18		0.17
		(0.36)		(0.22)
Observations	72043	72043	72043	72043
Panel (b): CAIR enacted 2009 (period ending 2011-13)				
Nonattainment	-2.35	-2.39	-0.56	-0.60
	(0.27)	(0.29)	(0.096)	(0.097)
CAIR		0.22		0.18
		(0.20)		(0.078)
Observations	72043	72043	72043	72043
Part B: Focusing on NBP or CAIR without/with controlling for trends				
Panel (c): NBP or CAIR				
NBP	-0.49		0.0034	
	(0.22)		(0.21)	
CAIR		-0.20		0.11
		(0.28)		(0.085)
Observations	72043	72043	72043	72043
Panel (d): $PM_{2.5}$ – replicating Deschenes et al. (2017)				
NBP x Post x Summer	-0.45	-1.03	0.19	-0.55
	(0.32)	(0.27)	(0.31)	(0.36)
Observations	4172	4172	4172	4172
Panel (e): Ozone 8-hour value (ppb) – replicating Deschenes et al. (2017)				
NBP x Post x Summer	-3.38	-3.37	-3.01	-3.06
	(0.56)	(0.54)	(0.54)	(0.56)
Observations	2352	2352	2352	2352

Notes: Part A tests robustness of our nonattainment estimates to controlling for NBP or CAIR status. Column 1 shows the baseline DiD results identical to Column 1 in Table 1. Column 3 shows the baseline DiDwb results identical to Column 2 in Table 1. Columns 2 and 4 add indicators for counties containing at least one facility that was subject to regulation under NBP or CAIR as control variables in Panel (a) and (b) respectively. Part B focuses on NBP and CAIR and tests robustness to controlling for trends based on baseline pollution (DiDwb). Panel (c) uses our approach and $PM_{2.5}$ data from Meng et al. (2019*b*), and shows results for DiD in Columns 1 and 2 for NBP and CAIR, where the endlines are 2006-08 and 2011-13 respectively. Columns 3 and 4 add the controls (DiDwb) to Columns 1 and 2. In Panel (d) and (e), we use the data and code from Deschenes et al. (2017*b*) to first replicate their results in Columns 1 and 2, focusing on the NBP with data from 2001-2007. Panel (d) and (e) Columns 1 and 2 correspond to their Table 2b, Row 9 and Row 4, Columns 4 and 5 respectively. The analysis is for panel data at the county-by-year-by-season level and both columns include county-by-season, summer-by-year, and county-by-year fixed effects as well as detailed weather controls. Column 1 is weighted by emission/pollution monitors, and Column 2 by population. Columns 3 and 4 (DiDwb) add year dummies interacted with baseline pollution and seasons to Columns 1 and 2 respectively. Panel (d) uses their $PM_{2.5}$ data as outcome and control for baseline trend, and Panel (e) uses their Ozone data as outcome and control for baseline trend. Standard errors in parentheses are clustered at the county level for Panels (a) to (c) and clustered at the state by season level for Panel (d) and (e).

A.9 Addressing uncertainty in pollution data

To address possible non-classical measurement error in the $PM_{2.5}$ reanalysis data, we show three robustness test.

The first tests relies on the uncertainty data in van Donkelaar et al. (2021*b*). For each grid-point, this data not only contains the estimated $PM_{2.5}$ concentration, but also information of the uncertainty around this estimate due to local geo-physical characteristics or distance to monitors. We drop 30% of Census tracts with the largest uncertainty average over our sample period and normalized by its mean.[62] To avoid mixing $PM_{2.5}$ data sources, we rely on pollution data from van Donkelaar et al. (2021*b*) exclusively for this first exercise, so the most comparable baseline table that retains all observations is Appendix Table A.24, which corresponds to main Table 1 based on data from Meng et al. (2019*b*). Part A of Table A.12 shows that our estimates are robust to dropping 30% of tracts that have the most uncertain $PM_{2.5}$ data estimates.

Second, since areas with ground-based air pollution monitors likely have less uncertainty in the $PM_{2.5}$ data, we drop all counties that neither contain a monitor themselves nor have a neighboring county with a monitor, using data from Meng et al. (2019*b*). Part B of Table A.12 shows that our results are robust to dropping these counties.

Third, in our most restrictive approach in Table A.13, we rely exclusively on data from ground-based pollution monitors from EPA (2022*a*). This severely reduces our observations, but even in this restrictive version, the patters of our main results and the bias of naive DiD are robust.

[62]That is we first take the average uncertainty and the average $PM_{2.5}$ estimate for each Census tract across our sample period (where we use tract data from population-weighted Census block estimates as in our main paper). We then calculate a normalized measure of uncertainty by dividing the uncertainty from the previous step by the average from the previous step for each tract. We then drop the 30% of tracts with the highest value of normalized uncertainty.

Table A.12: Dropping areas with higher uncertainty in pollution measurements

	ATT					LATE	
	All Tracts				with RV	Optimal Bandw.	
	DiD	DiDwb	M1DiD	M2DiD	DiD	RD0	RD1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Part A: Dropping 30% of tracts with highest uncertainty measures							
Panel (a): Homogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-1.46	-0.29	-0.40	-0.40	-1.47	-0.59	-0.76
	(0.36)	(0.12)	(0.18)	(0.21)	(0.38)	(0.36)	(0.41)
Observations	50452	50452	25499	26061	33582	4499	6166
Panel (b): Placebo Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-0.24	-0.086	-0.053	0.022	-0.58	-0.21	-0.35
	(0.12)	(0.11)	(0.12)	(0.12)	(0.18)	(0.30)	(0.47)
Observations	29672	29672	17636	17314	13547	1117	2233
Panel (c): Heterogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	5.27	3.50	2.13	3.79	5.37	3.77	3.02
	(0.83)	(0.86)	(0.41)	(0.81)	(0.83)	(0.87)	(0.94)
NA(x)Baseline	-0.44	-0.29	-0.18	-0.29	-0.45	-0.30	-0.25
	(0.059)	(0.063)	(0.028)	(0.058)	(0.059)	(0.052)	(0.052)
Observations	50452	50452	25499	26061	33582	4499	6166
Implied ATE	-1.39	-0.92	-0.54	-0.56	-1.37	-0.81	-0.73
10th pct	-0.17	-0.10	-0.047	0.24	-0.13	0.033	-0.042
90th pct	-3.61	-2.39	-1.43	-2.01	-3.61	-2.33	-1.98
Part B: Only counties (incl. neighboring counties) with monitor							
Panel (d): Homogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-1.53	-0.28	-0.41	-0.32	-1.56	-0.34	-0.016
	(0.37)	(0.16)	(0.20)	(0.21)	(0.37)	(0.32)	(0.42)
Observations	47821	47821	24267	24197	44008	6180	9928
Panel (e): Placebo Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-0.39	-0.096	0.0027	-0.048	-0.50	-0.14	-0.29
	(0.15)	(0.13)	(0.15)	(0.16)	(0.15)	(0.21)	(0.29)
Observations	26991	26991	15636	15634	23178	2113	5121
Panel (f): Heterogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	4.82	3.58	1.80	1.97	4.79	4.11	3.70
	(0.80)	(0.84)	(0.32)	(0.34)	(0.81)	(0.87)	(0.68)
NA(x)Baseline	-0.42	-0.30	-0.16	-0.16	-0.42	-0.31	-0.26
	(0.059)	(0.064)	(0.020)	(0.021)	(0.059)	(0.054)	(0.035)
Observations	47821	47821	24267	24197	44008	6180	9928
Implied ATE	-1.47	-0.97	-0.54	-0.46	-1.50	-0.54	-0.20
10th pct	-0.32	-0.14	-0.11	-0.014	-0.34	0.32	0.51
90th pct	-3.57	-2.49	-1.32	-1.27	-3.59	-2.08	-1.51

Notes: Part A is equivalent to Appendix Table A.24 based on pollution data from van Donkelaar et al. (2021*b*), after dropping 30% of tracts that have the most uncertain $PM_{2.5}$ predictions. Part B is equivalent to Table 1 based on data from Meng et al. (2019*b*), after dropping all counties that do not contain a monitor and also do not have a neighboring county with a monitor. The table shows coefficient estimates for the treatment effect of nonattainment status on the change in $PM_{2.5}$ levels between the pre- and post-treatment periods. Each panel(x)column combination is from a separate regression as described in the text. (1) uses simple DiD, (2) adds controls for baseline $PM_{2.5}$ (2001-03), (3) runs DiD using a sample matched (1-to-1) on baseline $PM_{2.5}$, (4) matches on baseline $PM_{2.5}$, tract population and population density (both 2000), (5) again uses simple DiD but with the limited sample of areas for which an EPA-registered $PM_{2.5}$ value exists, (6) and (7) use the limited sample based on optimal bandwidth selection in a regression discontinuity framework. Standard errors in parentheses are clustered at the county level.

Table A.13: Nonattainment status and changes in $PM_{2.5}$ using EPA monitor data

	ATT					LATE	
	All Tracts				with RV	Optimal Bandw.	
	DiD	DiDwb	M1DiD	M2DiD	DiD	RD0	RD1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel (a): Homogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-1.25	-0.012	-0.10	-0.12	-1.22	-0.15	-0.084
	(0.24)	(0.16)	(0.24)	(0.31)	(0.24)	(0.26)	(0.48)
Observations	667	667	279	268	596	36	86
Panel (b): Placebo Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-0.53	-0.072	-0.37	-0.058	-0.56	-0.63	-0.53
	(0.11)	(0.16)	(0.25)	(0.25)	(0.11)	(0.57)	(0.52)
Observations	431	431	239	244	360	23	113
Panel (c): Heterogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	5.29	3.70	2.07	2.09	5.32	4.54	6.82
	(1.22)	(1.24)	(0.61)	(0.64)	(1.22)	(2.04)	(2.12)
NA(x)Baseline	-0.42	-0.27	-0.15	-0.15	-0.42	-0.31	-0.46
	(0.084)	(0.087)	(0.039)	(0.038)	(0.084)	(0.13)	(0.14)
Observations	667	667	279	268	596	36	86
Implied ATE	-1.25	-0.59	-0.25	-0.29	-1.22	-0.36	-0.43
10th pct	-0.015	0.22	0.19	0.16	0.019	0.57	0.94
90th pct	-2.90	-1.68	-0.84	-0.88	-2.86	-1.59	-2.25

Notes: The table shows coefficient estimates for the treatment effect of nonattainment status on the change in $PM_{2.5}$ levels between the pre- and post-treatment periods. Each panel(x)column combination is from a separate regression as described in the text. (1) uses simple DiD, (2) adds controls for baseline $PM_{2.5}$ (2001-03), (3) runs DiD using a sample matched (1-to-1) on baseline $PM_{2.5}$, (4) matches on baseline $PM_{2.5}$, tract population and population density (both 2000), (5) again uses simple DiD but with the limited sample of areas for which an EPA-registered $PM_{2.5}$ value exists, (6) and (7) use the limited sample based on optimal bandwidth selection in a regression discontinuity framework. Standard errors in parentheses are clustered at the county level. All results based on monitor data from EPA (2022*a*).

A.10 Synthetic Control Estimates

Instead of our MDiD approach combining matching and DiD, an alternative approach can be based on Synthetic Controls (SC). The traditional SC method is designed for a context with few treated units and a small 'donor pool' of control units (Abadie et al. 2010). In our context, however, we have many treated and control units, and therefore use MDiD as on of our primary alternatives. Nevertheless, we extend the SC methods to our setting and provide two sets of results based on synthetic counterfactuals for additional robustness analysis.

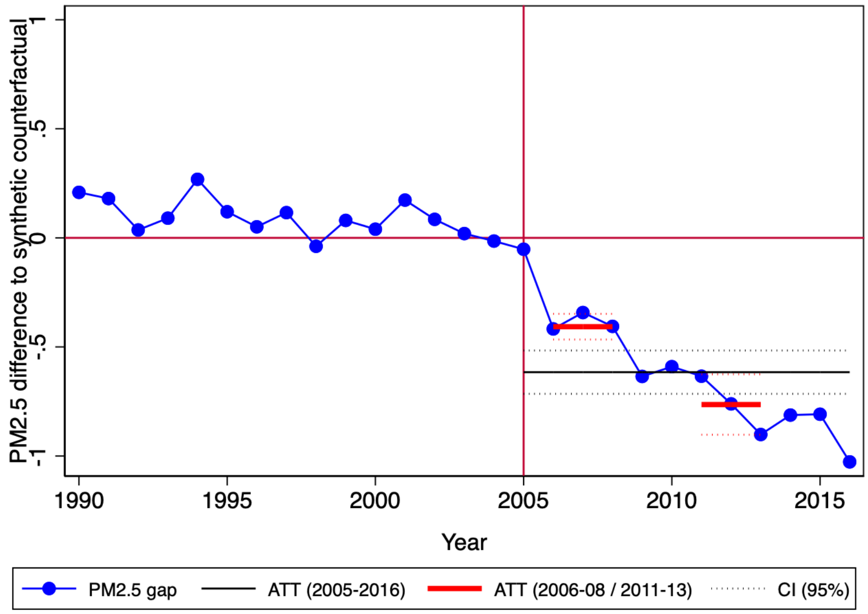
First, we use the recently proposed Synthetic Difference-in-Differences (SDiD) estimator to estimate average treatment effects (Arkhangelsky et al. 2021). In addition to unit weights chosen to closely replicate the average treated unit before treatment as in MDiD, SDiD uses time weights in the pre-treatment period to reduce variation in time trends among control units. SDiD is implemented as a weighted DiD with unit fixed effects and has been shown to perform equally well or better than traditional SC or DiD in common settings (Arkhangelsky et al. 2021). We show county level results based on SDiD in Figure A.13a using annual $PM_{2.5}$ concentrations based on Meng et al. (2019*b*) during the 1990-2004 pre-treatment period as predictor variables.[63] The blue line shows that the gap between the average nonattainment county and the weighted control units is small until 2005, but diverges in the expected direction after 2005. The estimated ATT for the full post-treatment period (2005-2016) shows a 0.62 $\mu g/m^3$ reduction in $PM_{2.5}$ (black line). To compare the SDiD estimates to our main analysis in Table 1 we focus on the same post-treatment time periods. The two red lines in Figure A.13a show that the SDiD estimates are very similar to our main estimates, with an ATT of 0.41 until 2006-08 and an ATT of 0.76 until 2011-13, confirming robustness of our main results.

While traditional SC estimation is not suitable for estimating average treatment effects in our setting, it offers another approach to heterogeneity analysis. We construct synthetic counterfactuals for each of the 208 nonattainment counties based on 1990-2004 $PM_{2.5}$ levels, each time limiting the 'donor pool' to attainment counties from the same Census division.[64] We then compare the change between 2001-03 and 2006-08 between each of the 208 nonattainment counties and its' synthetic counterfactual visualized by the red markers in Figure A.13b. This shows a population weighted average effect of 0.62 (red dotted line), and shows heterogeneity that increases the treatment effect with baseline pollution as captured by EPA-registered $PM_{2.5}$ values, in line with our main results. For a placebo exercise, we also run a SC analysis for each of the 339 attainment counties that have an EPA-registered $PM_{2.5}$ value below 15. The blue markers in Figure A.13b show the results. Reassuringly,

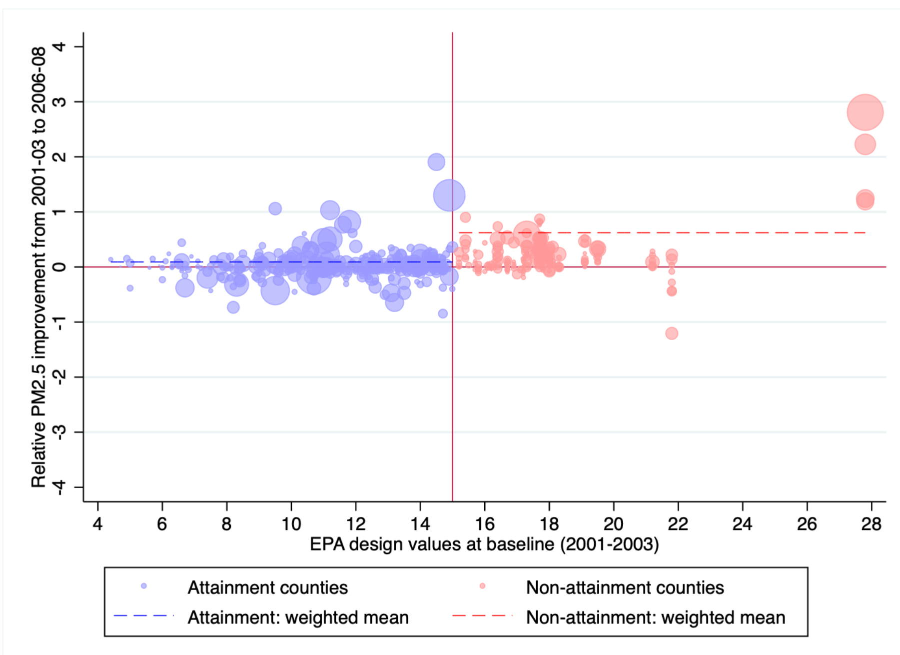
[63]Standard error calculations proposed by Arkhangelsky et al. (2021) are cluster-robust at the level of treated units, i.e. counties in our setting. County $PM_{2.5}$ levels are population-weighted averages across tracts.

[64]The nine Census Divisions are New England, Middle Atlantic, East North Central, West North Central, South Atlantic, East South Central, West South Central, Mountain Division and Pacific Division.

and in contrast to Figure 4, there is no visible association between $PM_{2.5}$ improvements and EPA-registered $PM_{2.5}$ values for these placebo attainment counties, as these are evaluated against their own synthetic counterfactuals.



(a) Synthetic Difference-in-Differences Estimation



(b) Unit-wise Synthetic Control Estimation (2001-03 vs. 2006-08)

Figure A.13: Synthetic Control Estimates

Notes: These figures are based on county level synthetic control estimates for 3,109 counties, 208 of which are in nonattainment of the $PM_{2.5}$ NAAQS from 2005. Panel (a) implements Synthetic Difference-in-Differences estimation following Arkhangelsky et al. (2021). The black line shows the estimated average treatment effect on the treated (ATT) for the full post-treatment period (2005-16). The red lines show the ATT for the three-year average periods used in the main analysis (2006-08 and 2011-13). Dashed lines indicate 95% confidence intervals based on standard errors that allow for correlation within county clusters. Panel (b) shows unit-wise Synthetic Control estimates of the $PM_{2.5}$ improvement between 2001-03 and 2006-08 for the 208 nonattainment counties evaluated against their synthetic counterfactual in red, and 339 attainment counties that have an EPA-registered $PM_{2.5}$ value (below 15), each relative to their unit-wise synthetic counterfactual in blue. Bubble size indicates county population in 2010 and dashed lines are population-weighted means. Based on data from Meng et al. (2019*b*).

A.11 Heterogeneous $PM_{2.5}$ nonattainment treatment effect by previous PM_{10} nonattainment status

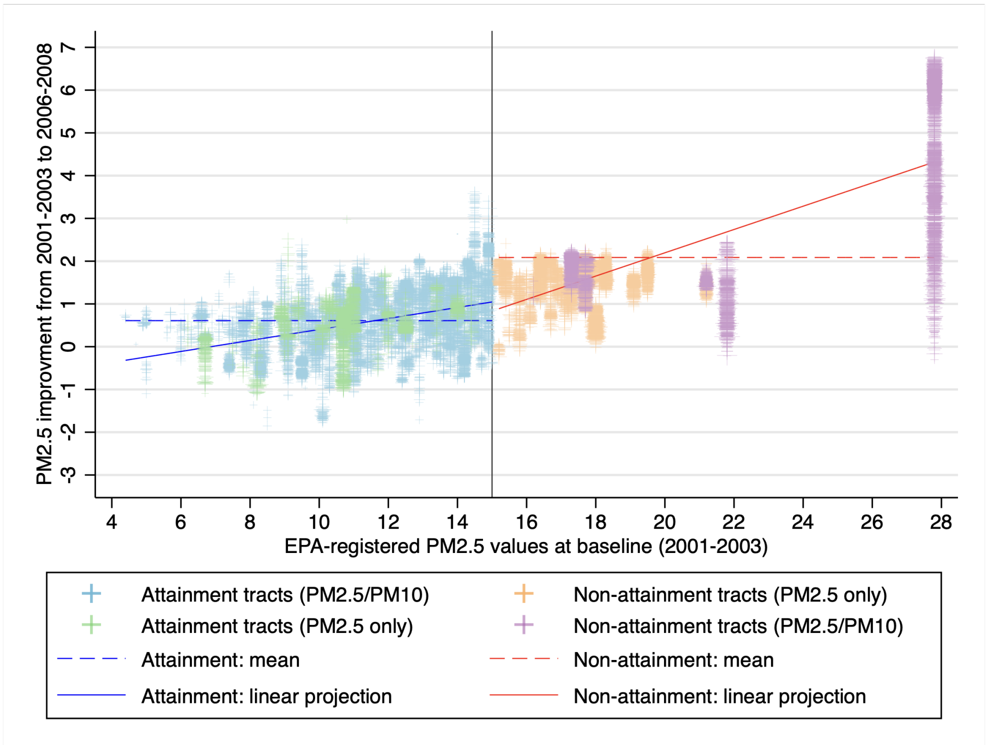


Figure A.14: Improvement in tract $PM_{2.5}$ averages and $PM_{2.5}/PM_{10}$ nonattainment status

Notes: The figure shows the improvement in $PM_{2.5}$ averages at the tract level between two periods, 2001-2003 and 2006-2008. The size of the markers reflect tract level populations. The $PM_{2.5}$ improvements are plotted against the EPA-registered $PM_{2.5}$ values of each attainment/nonattainment area, each of which usually comprises multiple counties and tracts. The dashed line plots the average $PM_{2.5}$ improvement for tracts in nonattainment and attainment areas separately, weighted by tract population, equivalent to the standard DiD estimate. The solid lines plot the linear projection of tract level $PM_{2.5}$ improvements on the EPA-registered $PM_{2.5}$ values of the nonattainment and attainment areas separately, weighted by tract population. In addition, the purple markers indicate nonattainment areas that are also in nonattainment of the previous PM_{10} regulation, and the green markers indicate attainment areas that are in nonattainment of the previous PM_{10} regulation. Based on data from Meng et al. (2019*b*).

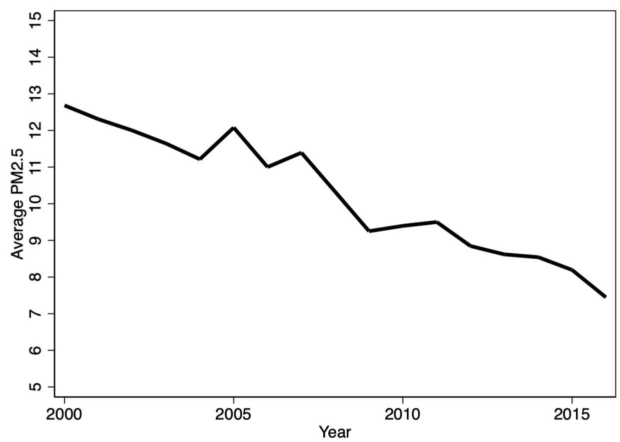
Table A.14: Heterogeneous treatment effects by previous PM_{10} status

	ATT					LATE	
	All Tracts				with RV	Optimal Bandw.	
	DiD	DiDwb	M1DiD	M2DiD	DiD	RD0	RD1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel (a): Homogeneous Treatment Effect: from 2001-03 to 2006-08							
PM2.5 NA	-0.91	-0.26	-0.32	-0.19	-0.90	-0.35	-0.022
w/o prev. PM10 NA	(0.059)	(0.14)	(0.16)	(0.18)	(0.092)	(0.29)	(0.40)
PM2.5 NA	-2.78	-1.30	-1.29	-1.63	-2.79	0	0.63
w. prev. PM10 NA	(0.72)	(0.34)	(0.19)	(0.31)	(0.72)	(.)	(0.65)
Observations	72043	72043	28291	28909	47962	7026	10459
Panel (b): Homogeneous Treatment Effect: from 2001-03 to 2011-13							
PM2.5 NA	-2.02	-0.63	-0.56	-0.60	-2.07	-1.20	-1.16
w/o prev. PM10 NA	(0.078)	(0.093)	(0.091)	(0.098)	(0.11)	(0.38)	(0.36)
PM2.5 NA	-3.76	-0.57	-1.71	-2.11	-3.72	0	0.69
w. prev. PM10 NA	(0.71)	(0.25)	(0.26)	(0.30)	(0.71)	(.)	(2.17)
Observations	72043	72043	28291	28909	47962	6137	25856

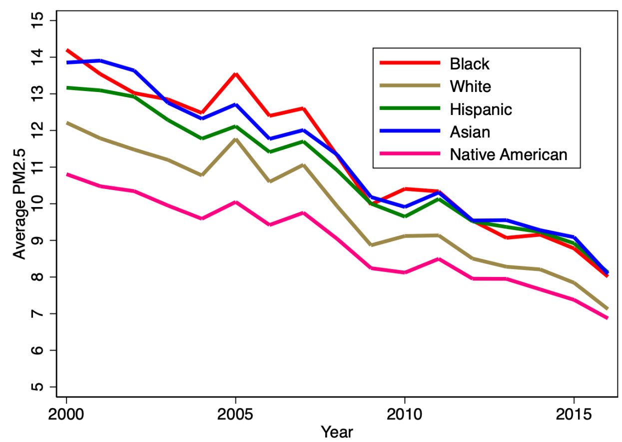
Notes: The table shows coefficient estimates for the treatment effect of nonattainment status on the change in $PM_{2.5}$ levels between the pre- and post-treatment periods. We allow heterogeneous treatment effects by previous PM_{10} nonattainment status, as in Equation (7). All regressions control for trends based on PM_{10} nonattainment status, so the shown coefficients are the heterogeneous marginal effects of $PM_{2.5}$ nonattainment status, as in Figure 7. Each panel(x)column combination is from a separate regression as indicated: (1) uses simple DiD, (2) adds controls for baseline $PM_{2.5}$ (2001-03), (3) runs DiD using a sample matched (1-to-1) on baseline $PM_{2.5}$, (4) matches on baseline $PM_{2.5}$, tract population and population density (both 2000), (5) again uses simple DiD but with the limited sample of areas for which an EPA-registered $PM_{2.5}$ value exists, (6) and (7) use the limited sample based on optimal bandwidth selection in a regression discontinuity framework. Standard errors in parentheses are clustered at the county level. All results based on Meng et al. (2019*b*).

A.12 Replication of $PM_{2.5}$ exposure levels in Jbaily et al. (2022)

Our main analysis uses the same $PM_{2.5}$ data (Meng et al. 2019*b*) that is also used by Jbaily et al. (2022) to document pollution exposure disparities across income and racial groups in the US. We replicate the relevant average exposure levels from their paper in Figure A.15. Panel (a) shows the population-weighted $PM_{2.5}$ exposure across all residents and panel (b) shows averages by racial groups. Reassuringly, both the levels and changes over time are virtually identical to those shown in panels (a) and (b) of Extended Data Fig. 1 in Jbaily et al. (2022).



(a) All residents



(b) By racial group

Figure A.15: Replication of $PM_{2.5}$ levels in Jbaily et al. (2022)

Notes: The Figure replicates population-weighted $PM_{2.5}$ exposure levels by population groups as shown in Extended Data Fig. 1 in Jbaily et al. (2022). Results are based on Meng et al. (2019*b*) and tract level population counts.

A.13 Comparison to recent analysis in Currie et al. (2023)

Part of our analysis of 2005 nonattainment effects, particularly the regulation's impact on $PM_{2.5}$ exposure gaps between Black and White residents, is closely related to the recent contribution by Currie et al. (2023) — henceforth CVW.[65] In the below, we first show that we can replicate some of the main findings of CVW despite using only our publicly available data.[66] Thereafter, we highlight the main differences between our and their approaches and discrepancies in data.

A. Replication of headline results in CVW (Currie et al. 2023)

To replicate the results of CVW, we rely on pollution data from Di et al. (2021) as in their analysis. While CVW use samples of individuals from the long form 2000 Census and the American Community Survey (ACS) from 2001, we use the population counts that account for the universe of individuals from Census records, and linearly interpolate between the 2000, 2010, and 2020 Census records. As in CVW, we assign pollution to individuals based on Census blocks. We run the entire replication analysis at the Census block level. Importantly, we also use their assignment into treatment status for the purposes of replication, which we discuss in more detail in the next section.

Currie et al. (2023) begin by showing that Black Americans are exposed to substantially higher levels of $PM_{2.5}$ than White Americans, and that this gap has narrowed over time. We show in Table A.15 that the average $PM_{2.5}$ exposure levels and the Black-White gap closely, although not exactly, replicate the numbers reported in Table 2 of CVW. The numbers in Column 1 and 3 are particularly similar, as these are both based on the 2000 Census without including data from the American Community Survey (ACS).[67]

Turning to the 2005 NAAQS nonattainment designations, CVW first show an event study of the regulation on $PM_{2.5}$ concentrations. Figure A.16 shows that we can replicate their event study almost exactly. Note that there is no pre-trend in this Figure, which is due to CVW assigning a subset of treated units into the control group as we discuss in more detail in the next section. We next replicate their baseline average treatment effects in Table A.16. With year and county fixed effects, we estimate an ATE of -1.2 $\mu g/m^3$, almost identical to the estimate of -1.230 reported by CVW. Similarly, when adding state-year fixed effects, our estimate falls to -0.76 (compared to -0.737 in CVW).

Finally, CVW ask how much of the reduction in the Black-White exposure gap between 2005 and 2015 can be accounted for by the 2005 nonattainment designations. To account for effect heterogene-

[65] We are grateful for the authors of CVW for helpful discussions, especially Reed Walker.

[66] Their individual level data is indispensable for their analysis of contributions of individual level income to exposure gaps. We only focus on their main results here.

[67] CVW use the 2000 Census long form which is a subset of the 2000 Census.

Table A.15: Replication of Table 2 in CVW

	Actual 2000 Exposure	Actual 2015 Exposure	Counterfactual 2015 using 2000 locations
Panel (a): Original numbers reported in CVW			
White	12.96	8.25	8.22
Black	14.52	8.79	8.89
B-W Difference	1.56	0.54	0.67
Chg. in B-W Diff	0.00	-1.02	-0.89
Panel (b): Replication using our data			
White	12.90	8.12	8.17
Black	14.53	8.80	8.90
B-W Difference	1.63	0.68	0.73
Chg. in B-W Diff	0.00	-0.95	-0.90

Notes: Panel (a) restates Table 2 from Currie et al. (2023). Panel (b) replicates those numbers using our data. Columns 1 and 2 report average $PM_{2.5}$ exposure levels, using block level population weights that are linearly interpolated between 2010 and 2020. Column 3 uses constant 2000 population weights instead. Pollution data is from Di et al. (2021).

ity, they estimate RIF-Quantile treatment effects in 19 pollution 'vigintiles', separately for Black and White residents. We replicate the RIF-Quantile regressions in Figure A.17, which again closely resembles Figure 8 in CVW. We replicate the counterfactual gap accounting based on these regression results in Table A.17. Again, the results closely resemble those reported in Table 4 of CVW. Our overall actual change in the gap (0.47 vs. 0.59 $\mu g/m^3$) and counterfactual change in the gap (0.18 vs 0.23 $\mu g/m^3$) are similar but slightly smaller. Yet, using our publicly available Census based data recovers virtually the same contribution of the CAA nonattainment areas to the reduction in the Black-White pollution gap (61.1%) as the ACS-based individual level sample in CVW (61.2%).

Throughout this section, we used the same classification of treated and control units as in CVW. The small differences in results are due to differences between our (interpolated) block level population counts and the individual level survey sample used in CVW. We next discuss these differences further.

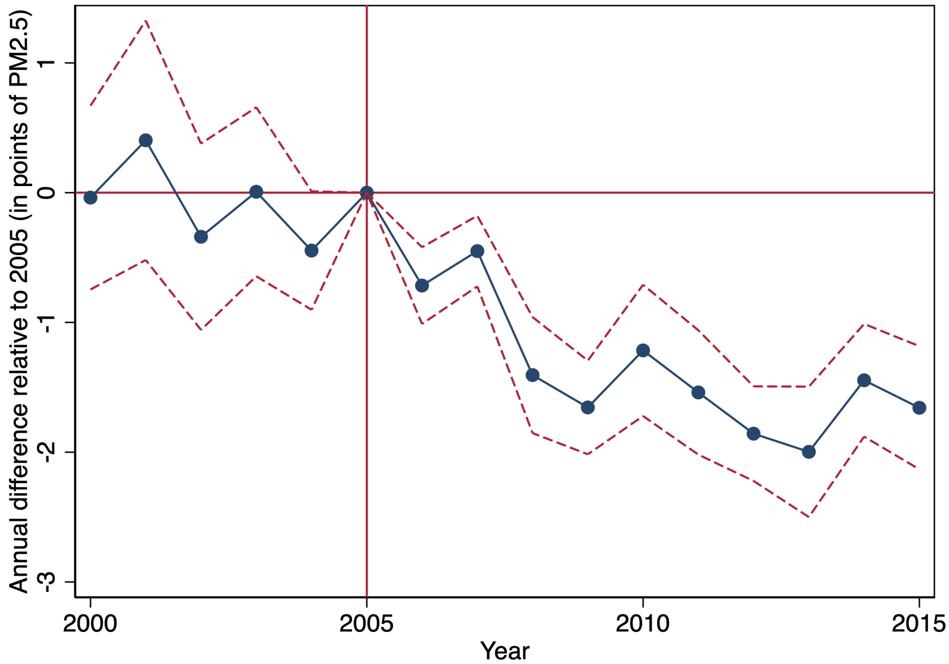


Figure A.16: Replication of Figure 6 in CVW

Notes: This figure replicates Figure 6 in Currie et al. (2023) using the block level data from our paper. The graph shows an event study plotting the coefficients from nonattainment areas as defined by CVW interacted with year dummies. The regression model controls for county fixed effects and year fixed effects. The regression is weighted by block level population counts, linearly interpolated between 2000, 2010 and 2020, and errors are clustered by commuting zone. Pollution data is from Di et al. (2021).

Table A.16: Replication of Table 3 in CVW

	(1)	(2)
Panel (a): Original numbers reported in CVW		
$PM_{2.5}$ NA	-1.230	-0.727
	(0.335)	(0.080)
Observations	32,360,000	32,360,000
Panel (b): Our data		
PM2.5 NA	-1.20	-0.76
	(0.40)	(0.078)
Observations	108,583,670	108,583,670
County FE	Yes	Yes
Year FE	Yes	No
State-Year FE	No	Yes

Notes: The table replicates difference-in-differences estimates of the average treatment effect from nonattainment designations shown in Table 3 in Currie et al. (2023) using our data at the block level with block population weights. Column 1 replicates original Column 1, Column 2 replicates original Column 5. Pollution data is from Di et al. (2021).

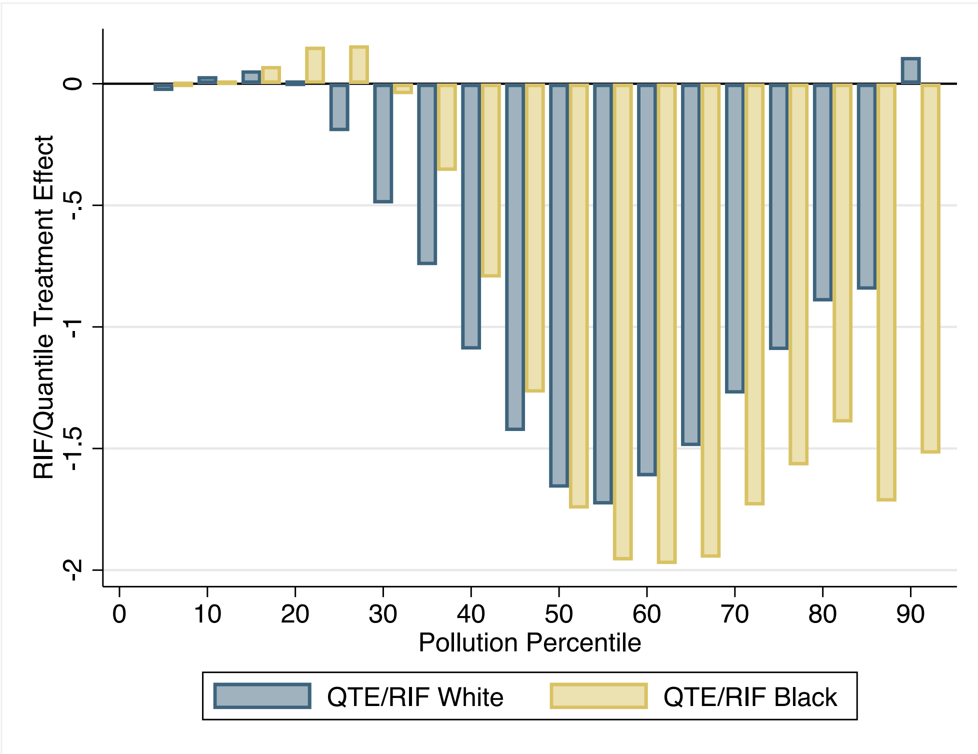


Figure A.17: Replication of Figure 8 in CVW

Notes: This figure replicates Figure 8 in Currie et al. (2023) using the block level data from our paper. It plots regression coefficients from 38 separate regressions, 19 for each race, where the dependent variable consists of the RIF-Quantile transformation of the respective $PM_{2.5}$ vigintile (indicated by the x-axis). The regression model controls for county fixed effects and state-by-year fixed effects. Regressions are weighted by block level population counts, linearly interpolated between 2000, 2010 and 2020, and errors are clustered by commuting zone. Pollution data from Di et al. (2021).

Table A.17: Replication of Table 4 in CVW

$PM_{2.5}$ Quantile Bin	Actual $PM_{2.5}$ in 2005	Actual $PM_{2.5}$ in 2015	White Counterfactual $PM_{2.5}$ in 2015 Without CAA	Black Counterfactual $PM_{2.5}$ in 2015 Without CAA
5	5.38	4.22	4.22	4.22
10	7.94	5.58	5.58	5.58
15	8.97	6.22	6.21	6.22
20	9.7	6.71	6.7	6.7
25	10.36	7.11	7.11	7.09
30	10.91	7.45	7.49	7.42
35	11.43	7.75	7.85	7.76
40	11.92	8.01	8.19	8.11
45	12.36	8.25	8.52	8.48
50	12.74	8.47	8.88	8.87
55	13.1	8.69	9.21	9.3
60	13.46	8.89	9.53	9.76
65	13.82	9.09	9.81	10.05
70	14.18	9.29	10.04	10.39
75	14.55	9.52	10.2	10.55
80	14.95	9.78	10.4	10.82
85	15.34	10.13	10.62	11.12
90	15.85	10.71	11.19	11.85
95	17.35	12.55	12.51	12.97

Main Counterfactual incl. 2005-2015 Mobility Responses		
	Original numbers reported in CVW	**Our data**
2005 Actual B-W Gap	1.20	1.16
2015 Counterfactual B-W Gap	0.97	0.98
Counterfactual Chg in B-W Gap	-0.23	-0.18
Actual Chg in B-W Gap	-0.59	-0.47
% Attributable to CAA	61.2	61.1

Notes: The table replicates Table 4 from Currie et al. (2023) using our block level data. Population counts are linearly interpolated between 2000, 2010 and 2020 to approximate the approach in Currie et al. (2023), who follow individuals in their data as they move across locations. Counterfactuals are calculated as the actual $PM_{2.5}$ levels in 2015 minus the RIF-Quantile treatment effects of nonattainment (applied in proportion to the population share living in nonattainment areas), separately for each vigintile and for each racial group. Pollution data is from Di et al. (2021).

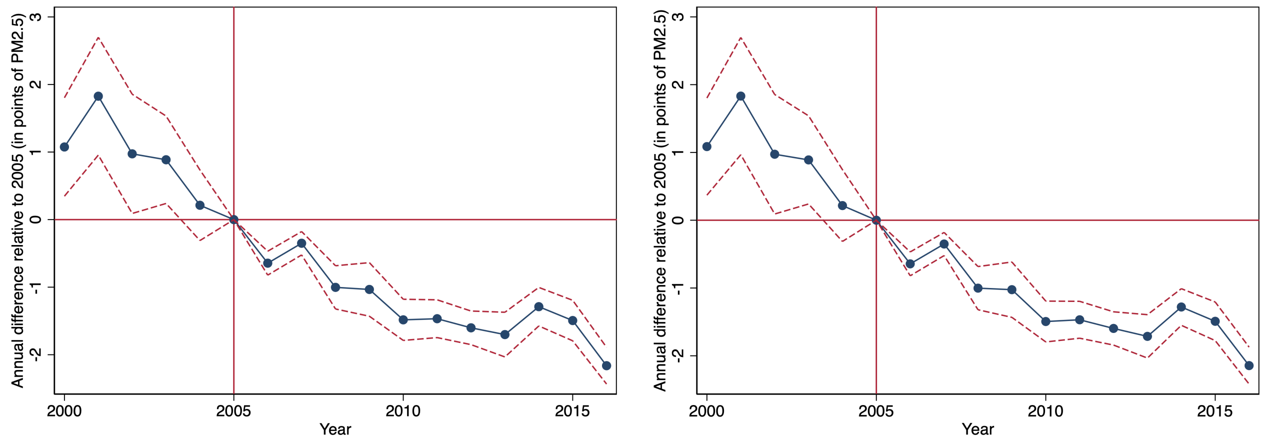
B. *Explaining differences compared to CVW* (*Currie et al. 2023*)

There are several differences between our approach and that of CVW, yet there are only two differences that are important: treatment assignment and controlling for baseline trends. We first briefly discuss minor data discrepancies that make no difference for the main findings before we turn to the two important differences.

We have shown in the previous section that using our publicly available Census data recovers virtually the same estimated nonattainment effects on pollution and contribution of the CAA nonattainment areas to narrowing the Black-White exposure gap. This is reassuring and shows that any differences due to using publicly available data vs. individual level American Community Survey (ACS) samples are negligible, especially because pollution is assigned to individuals at the Census block level in both approaches. Nevertheless, we briefly list some of the data differences and use the event study to illustrate that they do not matter for this analysis.[68] First, CVW use a sample based on the Census long form as well as the 1% ACS sample. Our data is constructed from the full Census population. If samples are random, we should recover the same estimates in expectation. Second, CVW incorporate year-to-year mobility through the annual ACS samples while we use fixed 2010 location in our main analysis, or interpolated block populations by race using the 2000, 2010 and 2020 Census in the preceding section or for robustness in Figure A.1, for example. The event study graph is virtually indistinguishable when using interpolated vs. constant 2010 population as shown in Figure A.18a, compared to Figure A.21. Third, CVW use 2000 block boundaries while we use 2010 block boundaries. We also aggregate to the tract level using block population weights, which should, however, be equivalent to running the regression at the block level for the purposes of the event study. Figure A.18b shows that the event study is virtually unchanged if we use 2000 block boundaries and run the analysis at the block level, compared to Figure A.21.

While the differences between our data and that used in CVW appear negligible, there are two important differences. The first key difference is the assignment into treatment or control of those areas that are in $PM_{2.5}$ nonattainment, but have also been in PM_{10} nonattainment previously. In CVW, all such areas are assigned into the control group, sometimes known as switcher approach. As these areas do not switch into nonattainment from being in attainment of a previous NAAQS, so are 'merely' in nonattainment of an additional NAAQS ($PM_{2.5}$), one may expect that these areas experience a lower treatment effect from the additional nonattainment assignment. A switcher approach that assigns these areas into the control group assumes a treatment effect of zero for these areas. In Figure 7 we test this assumption and show, however, that the treatment effect for these areas

[68]Note that individual data would be required to assess the contribution of individual level factors to the exposure gap, which, however explain little as shown in CVW.



(a) Interpolated population weights 2000, 2010, 2020

(b) 2000 block borders and block level analysis

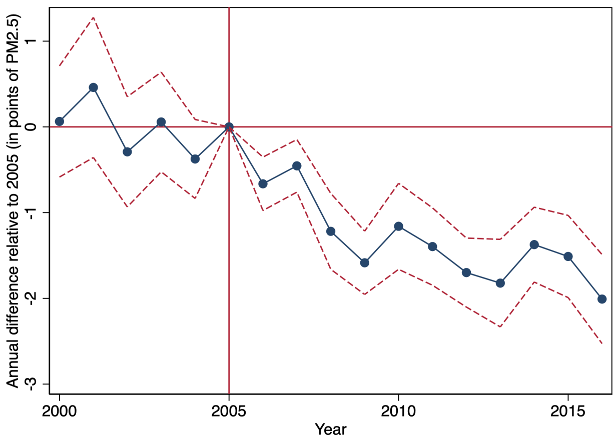
Figure A.18: Robustness of differences in pre-trends in event study

Notes: The figure replicates the event study graph from Panel (b) of Figure A.21. Panel (a) uses population weights that are interpolated between the 2000, 2010 and 2020 Census, using the IPUMS NHGIS crosswalk, instead of constant population weights at the 2010 level. Panel (b) uses borders and population counts from the 2000 Census instead of the 2010 Census. In addition, the analysis is at the Census block level, rather than pre-aggregating to the Census tract level using Census block weights as in our main analysis (the results are equivalent using either). Results are based on Di et al. (2021). Standard errors are clustered at the county level and 95% confidence intervals are shown.

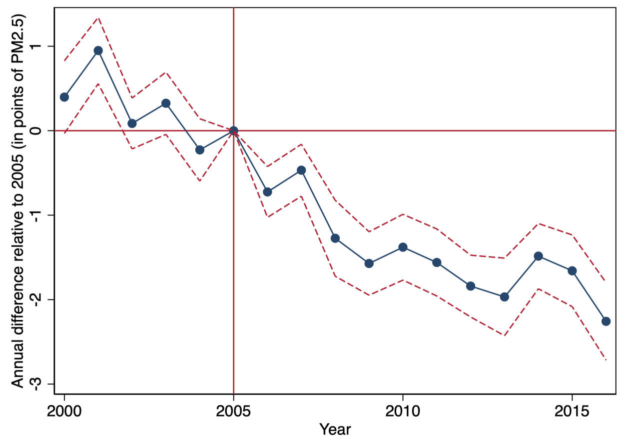
is – if anything – larger than for those areas that switched from PM_{10} attainment to $PM_{2.5}$ nonattainment. Using the treatment assignment of CVW, we can replicate their event study with insignificant pre-trends as shown in Figure A.19a (see also Figure A.16 above). This is intuitive, as $PM_{2.5}$ nonattainment areas that were also in nonattainment for PM_{10} tend to be more polluted and, as we show in our Figure 3a, also likely to exhibit the largest pre-trends, thus assigning them into the control group eliminates the pre-trends on average. If we instead drop these areas entirely, the pre-trends reappear (Figure A.19b). Note that a second, but minor difference in treatment assignment is that CVW assign entire commuting zones (CZ) into nonattainment treatment as long as a county within the CZ is in nonattainment, while we use EPA defined nonattainment areas based on air regions (i.e. the nonattainment counties). Figures A.19c and A.19d show that using CZ instead of counties based on EPA air regions have no discernible implications for the event study.

The second key difference is that we control for pollution trends based on baseline pollution as discussed in detail in our main paper. We next rerun some of the main estimations in CVW but additionally controlling for baseline pollution, similar to our DIDwb approach.[69] First, the estimated nonattainment effect is much smaller (even zero in one specification) as shown in Columns 3 and 4 of Table A.18 (Column 1 and 2 replicate the results in Table A.16 Panel b). This is in line with our main findings that when ignoring such trends, a naive DiD approach overestimates the nonat-

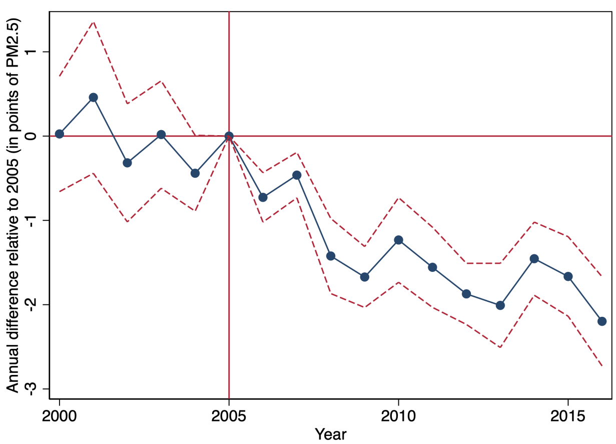
[69] That is we control for baseline pollution in 2000 interacted with year dummies in all of their panel regressions.



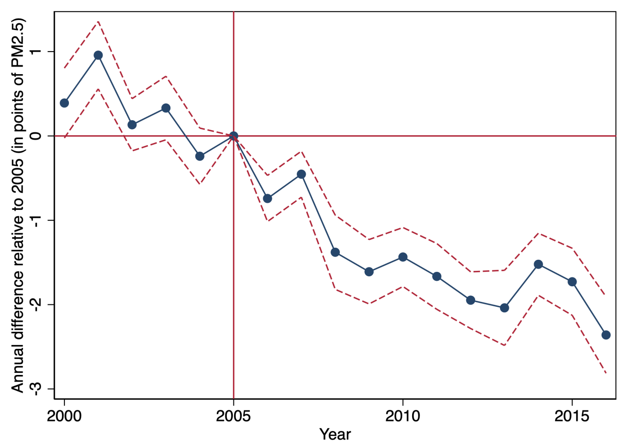
(a) Assigning double treated into control (county level)



(b) Dropping double treated (county level)



(c) Assigning double treated into control (CZ level)



(d) Dropping double treated (CZ level)

Figure A.19: Event study assigning previously treated units into the control group or dropping them

Notes: The figure replicates the event study graph from Panel (b) of Figure A.21 to facilitate comparison with the event study in Currie et al. (2023). Panel (a) assigns all nonattainment counties that were also in nonattainment with the earlier 1990 PM_{10} status in 2001-2004 into the control group (20 counties). Panel (b) instead drops these 20 counties. Panel (c) and (d) repeat these analysis of (a) and (b) respectively, but at the commuting zone level. Commuting zones in nonattainment, where all counties were previously also in nonattainment (i.e. did not switch into nonattainment), are assigned into the control group in Panel (c). These are 6 commuting zones, including, e.g. Los Angeles. In Panel (d), these commuting zones are instead dropped. Panel (b) and (d) look similar when we additionally drop counties from the control group that are in nonattainment with the PM_{10} standard, but in attainment with the $PM_{2.5}$ standard (this drops 71 counties instead of 20 counties, and is the sample we use for Table A.7). Standard errors are clustered at the commuting zone level in all four panels. Pollution data based on Di et al. (2021).

tainment effects significantly. When we use our treatment assignment instead (Columns 4-8), and control for baseline pollution, we recover effects similar as in our main analysis. These are naturally all slightly larger than the corresponding effects based on the CVW treatment assignment, where the most polluting areas with the largest effect are assigned into the control group as discussed in the previous paragraph. Second, turning to the estimation of the contribution of CAA nonattainment designations to narrowing the Black-White exposure gap, Table A.19 shows that controlling for trends based on baseline pollution lowers the estimated contribution from 61.1% (Column 1) to 18.6% (Column 2) using the same RIF analysis and CVW treatment assignment as in the previous replication section. When we use our treatment assignment instead, we find an overestimated contribution of 115.1% (Column 3) versus 22.5% with controls for trends (Column 4), similar to the pattern in our main paper. Our estimated effect in Column 4 aligns closely with our main findings in Table A.25 (i.e. the version of Table 2 that uses Di et al. (2021) data).[70] The main insight is that irrespective of using CVW or our treatment assignment, controlling for secular trends based on baseline pollution significantly reduces the estimated CAA contribution to the narrowing Black-White exposure gap, in this case by a factor of around 3-4. The large upward bias from ignoring such trends dominates the downward bias from using an approach that assigns some treated units already in nonattainment into the control group.

Table A.18: Extended replication of Table 3 in CVW

	CVW treatment assignment				Our treatment assignment			
	DiD		DiDwb		DiD		DiDwb	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PM2.5 NA CVW	-1.20	-0.76	-0.17	0.0064				
	(0.40)	(0.078)	(0.41)	(0.12)				
PM2.5 NA SS					-2.12	-1.49	-0.66	-0.20
					(0.54)	(0.45)	(0.16)	(0.10)
Observations	108,583,670				108,583,670			
County FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No	Yes	No	Yes	No
State-Year FE	No	Yes	No	Yes	No	Yes	No	Yes

Notes: The table replicates difference-in-differences estimates of the average treatment effect from nonattainment designations shown in Table 3 in Currie et al. (2023) using our data and a panel regression at the block-by-year level. Column 1 replicates original Column 1, Column 2 replicates original Column 5. Columns 3 and 4 control for baseline $PM_{2.5}$ separately in each year. Columns 5-8 repeat the analysis but use our treatment assignment instead of the CVW treatment assignment. Pollution data is from Di et al. (2021).

[70] For our main findings we allow for simple linear heterogeneity by baseline pollution and race, rather than the RIF approach used here. In Table A.25 based on Di et al. (2021) and DiDwb, we find a contribution of 24% versus the 22.5% estimated using RIF and a slightly different time window.

Table A.19: Extended replication of Table 4b in CVW

Main Counterfactual incl. 2005-2015 Mobility Responses				
	CVW	CVW-wb	SS	SS-wb
	(1)	(2)	(3)	(4)
2005 Actual B-W Gap	1.16	1.16	1.16	1.16
2015 Counterfactual B-W Gap	.98	.78	1.23	.8
Counterfactual Chg in B-W Gap	-.18	-.38	.07	-.36
Actual Chg in B-W Gap	-.47	-.47	-.47	-.47
% Attributable to CAA	61.1	18.6	115.1	22.5
2005 NA Treatment	Switcher	Switcher	All	All
Baseline Control (DiDwb)	No	Yes	No	Yes

Notes: The table shows an extended replication of Table 4 from Currie et al. (2023) using our block level data. Column 1 shows the same RIF-based replication as in Table A.17. Column 2 adds controls for baseline $PM_{2.5}$ in each RIF-Quantile regression. Population counts are linearly interpolated between 2000, 2010 and 2020 to approximate the approach in Currie et al. (2023), who follow individuals in their data as they move across locations. Counterfactuals are calculated as the actual $PM_{2.5}$ levels in 2015 minus the RIF-Quantile treatment effects of nonattainment (applied in proportion to the population share living in nonattainment areas), separately for each vigintile and for each racial group. Pollution data is from Di et al. (2021).

A.14 Counterfactual pollution disparities with constant 2010 population

Table A.20: Pollution disparities - counterfactual gap analysis with constant 2010 population

Panel (a): Black-White Pollution Gap

	$PM_{2.5}$ exposure		Black-White Gap		Contribution of CAA (in %) [homogeneous effect]					
Period	Black	White	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	13.14	11.49	1.65							
2006-2008	12.11	10.53	1.58	-0.07	282	93	78	76	68	4
2011-2013	9.65	8.64	1.01	-0.64	52	13	10	12	28	25

	$PM_{2.5}$ exposure		Black-White Gap		Contribution of CAA (in %) [heterogeneous effect]					
Period	Black	White	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	13.14	11.49	1.65							
2006-2008	12.11	10.53	1.58	-0.07	413	307	155	193	201	122
2011-2013	9.65	8.64	1.01	-0.64	67	14	31	31	49	38

	$PM_{2.5}$ exposure		Black-White Gap		Contribution of CAA (in %) [+race interactions]					
Period	Black	White	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	13.14	11.49	1.65							
2006-2008	12.11	10.53	1.58	-0.07	216	110	124	128	26	-24
2011-2013	9.65	8.64	1.01	-0.64	71	18	43	41	52	47

Panel (b): Urban-Rural Pollution Gap

	$PM_{2.5}$ exposure		Urban-Rural Gap		Contribution of CAA (in %) [homogeneous effect]					
Period	Urban	Rural	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	12.41	10.18	2.23							
2006-2008	11.22	9.61	1.60	-0.63	61	20	17	17	15	1
2011-2013	9.28	7.78	1.49	-0.74	83	20	16	20	44	39

	$PM_{2.5}$ exposure		Urban-Rural Gap		Contribution of CAA (in %) [heterogeneous effect]					
Period	Urban	Rural	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	12.41	10.18	2.23							
2006-2008	11.22	9.61	1.60	-0.63	85	63	32	39	40	23
2011-2013	9.28	7.78	1.49	-0.74	103	23	45	45	74	57

	$PM_{2.5}$ exposure		Urban-Rural Gap		Contribution of CAA (in %) [+urban interactions]					
Period	Urban	Rural	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	12.41	10.18	2.23							
2006-2008	11.22	9.61	1.60	-0.63	82	60	37	42	40	24
2011-2013	9.28	7.78	1.49	-0.74	104	24	46	47	74	61

Notes: Left columns show average $PM_{2.5}$ exposure of Black, White, Urban and Rural populations, and difference between groups, as derived from Census block level pollution concentrations and population counts. Right columns show contribution of CAA nonattainment designations in 2005 based on counterfactual calculations that factor out nonattainment treatment effects as estimated in Columns 1-4, 6, and 7 of Table 1. Population data is from the 2010 Census, held fixed across all years. Pollution data is from Meng et al. (2019*b*).

A.15 Event study for house price trends

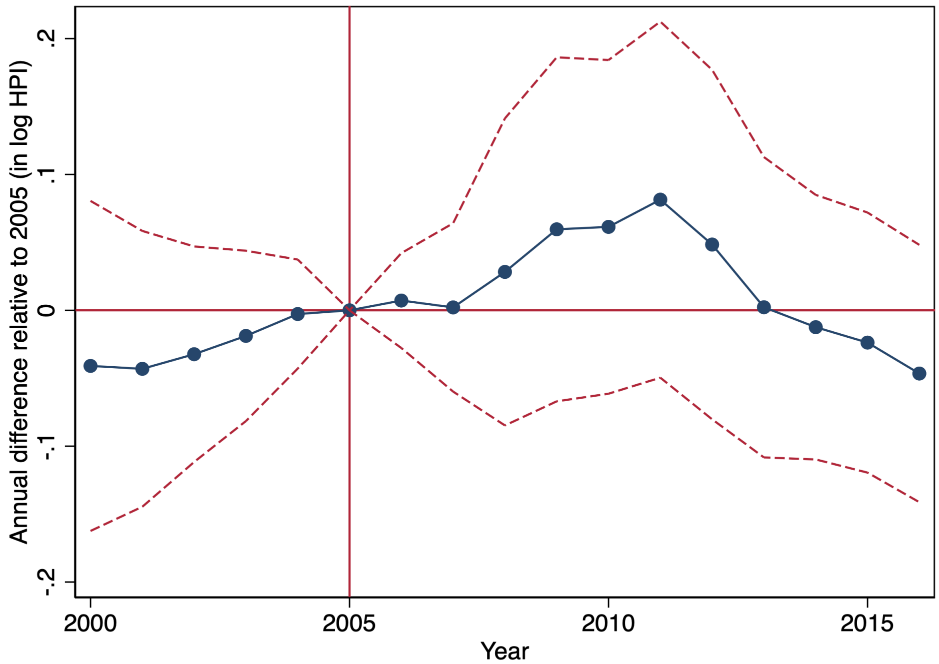


Figure A.20: Event study of house price growth nonattainment vs. attainment areas

Notes: The figure shows an event study plotting the average difference in (log) house prices between $PM_{2.5}$ nonattainment and attainment areas, normalized to 0 in 2005, as predicted from a DiDwb specification that allows for heterogeneous treatment effects by previous PM_{10} nonattainment status and baseline $PM_{2.5}$ levels in 2001-03 as in Table 3. Shown are the average treatment effects at the mean, and 95% confidence intervals are based on standard errors clustered at the county level. Pollution data is from Meng et al. (2019*b*).

A.16 Results for house price changes with commuting zone fixed effects

Table A.21: Pollution damages - instrumental variable comparison (with CZ FE)

	OLS	DiD-IV	DiDwb-IV	M1DiD-IV	M2DiD-IV	RD0-IV	RD1-IV
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel (a): Effect of $PM_{2.5}$ ***increases*** *on house price index growth 2001-03 to 2006-08*							
ΔPM2.5	-0.019	-0.025	-0.040	-0.037	-0.039	-0.053	-0.086
	(0.0053)	(0.0076)	(0.013)	(0.013)	(0.0097)	(0.024)	(0.034)
Observations	54483	54483	54483	21080	21602	5086	7936
K-P F statistic		107.2	9.67	92.7	82.9	37.4	98.3
Elasticity	-0.23	-0.30	-0.48	-0.44	-0.46	-0.64	-1.03
Panel (b): Effect of $PM_{2.5}$ ***increases*** *on house price index growth 2001-03 to 2011-13*							
ΔPM2.5	-0.022	-0.039	-0.14	-0.029	-0.042	-0.0060	-0.050
	(0.0099)	(0.0090)	(0.028)	(0.014)	(0.014)	(0.010)	(0.012)
Observations	54332	54332	54332	20990	21517	4495	19034
K-P F statistic		278.4	17.0	296.2	276.1	303.8	288.1
Elasticity	-0.27	-0.47	-1.73	-0.35	-0.50	-0.071	-0.60

Notes: The dependent variable is the change in the logarithm of the house price index. ΔPM2.5 is the change in $PM_{2.5}$ since 2001-03 in $\mu g/m^3$, instrumented by CAA nonattainment status for $PM_{2.5}$, allowing for heterogeneous effects in the instrument by previous PM_{10} nonattainment status and by baseline $PM_{2.5}$ levels in 2001-03. First-stage specifications in Columns 2-7 correspond to Columns 1-4, 6, and 7 in Table 1, with commuting zone fixed effects added. Standard errors in parentheses are clustered at the county level. Pollution data is from Meng et al. (2019*b*).

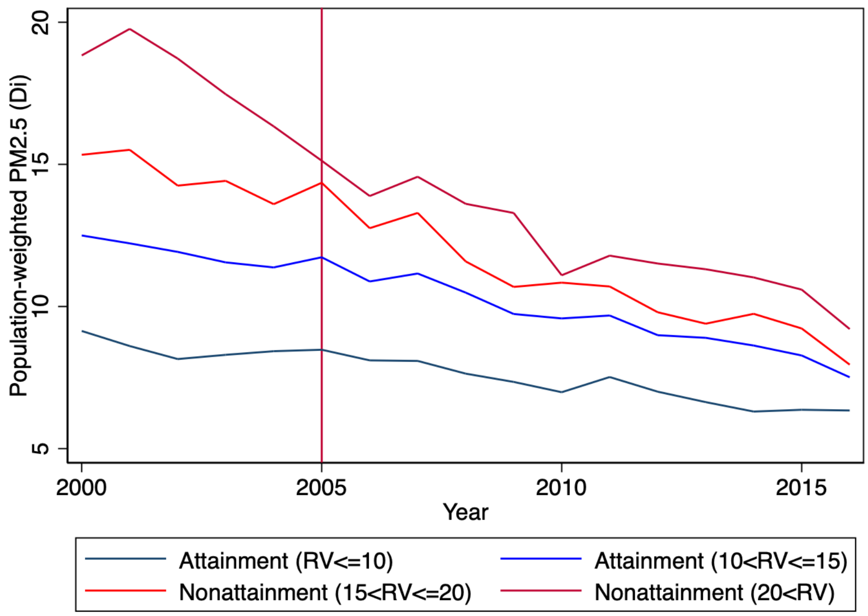
A.17 Reduced form results for house price changes

Table A.22: Reduced form effect of NA on HPI - instrumental variable comparison

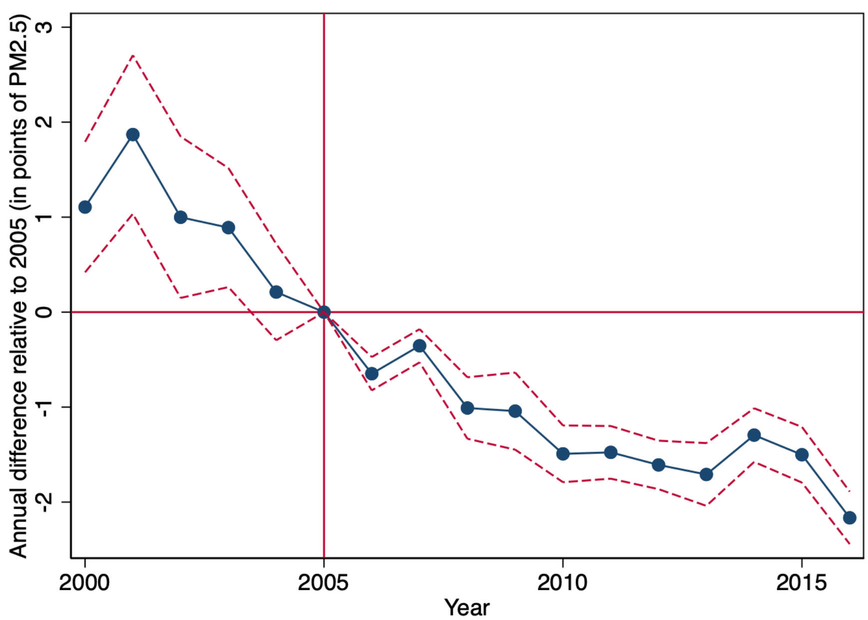
	DiD-RF	DiDwb-RF	M1DiD-RF	M2DiD-RF	RD0-RF	RD1-RF
	(1)	(2)	(3)	(4)	(5)	(6)
Panel (a): Effect of NA on house price index growth 2001-03 to 2006-08						
NA Effect	0.057	0.142	0.057	0.045	0.087	0.008
	(0.025)	(0.030)	(0.038)	(0.044)	(0.066)	(0.120)
Observations	54529	54529	21152	21693	5087	7937
Panel (b): Effect of NA on house price index growth 2001-03 to 2011-13						
NA Effect	-0.018	0.017	-0.027	-0.023	0.062	0.118
	(0.022)	(0.024)	(0.025)	(0.026)	(0.052)	(0.057)
Observations	54378	54378	21062	21608	4496	19035

Notes: The table shows reduced form estimates from a simplified version of our instruments that only includes $PM_{2.5}$ nonattainment (NA) and the interaction with baseline $PM_{2.5}$. Average treatment effects are calculated as linear combination of coefficient estimates for the NA dummy and NA interacted with baseline $PM_{2.5}$, evaluated at the mean. Standard errors in parentheses are clustered at the county level. Pollution data is from Meng et al. (2019*b*).

A.18 Replication of Tables 1-3 and Figures 3-5 in main paper with alternative $PM_{2.5}$ data



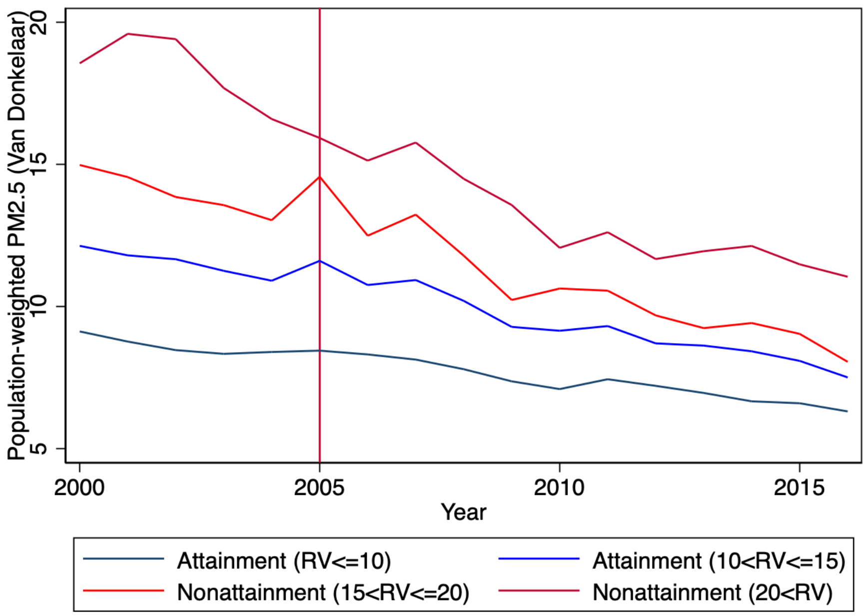
(a) Evolution of $PM_{2.5}$ grouped by EPA RV grouping



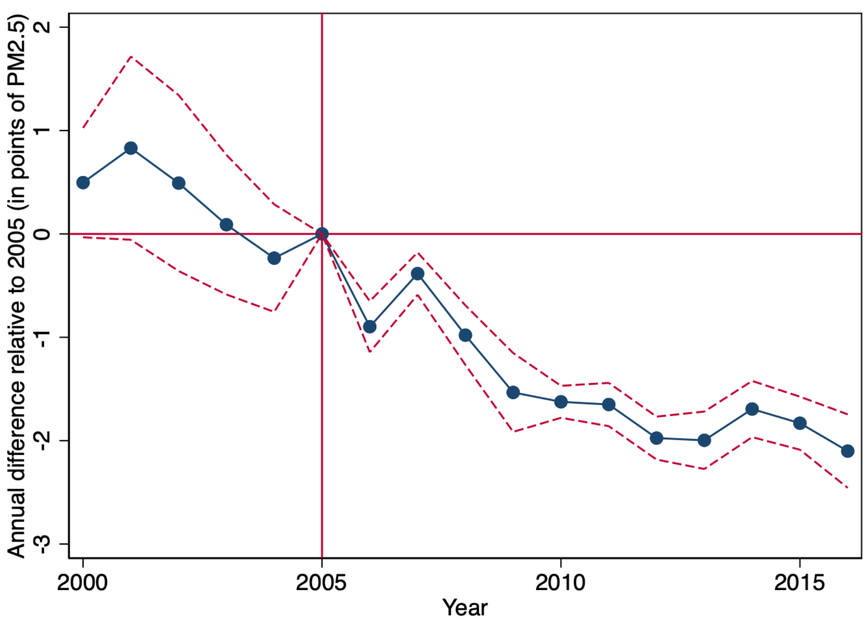
(b) Event study (annual nonattainment-attainment differences in $PM_{2.5}$)

Figure A.21: Trends in $PM_{2.5}$ and event study analysis using Di et al. (2021)

Notes: Panel (a) shows the change in $PM_{2.5}$ averages at the tract level (population-weighted) over time. Each line represents a different bin of EPA-registered $PM_{2.5}$ values assigned to each attainment/nonattainment area, each of which usually comprises multiple counties and tracts. Panel (b) shows coefficient estimates from a regression that includes a treatment dummy interacted with years, controlling for year fixed effects. The dotted blue line shows point estimates and the dashed red lines show 95% confidence intervals based on standard errors that are cluster-robust at the level of counties. Both Panels are based on data from Di et al. (2021).



(a) Evolution of $PM_{2.5}$ grouped by EPA RV grouping



(b) Event study (annual nonattainment-attainment differences in $PM_{2.5}$)

Figure A.22: Trends in $PM_{2.5}$ and event study analysis using van Donkelaar et al. (2021*b*)

Notes: Panel (a) shows the change in $PM_{2.5}$ averages at the tract level (population-weighted) over time. Each line represents a different bin of EPA-registered $PM_{2.5}$ values assigned to each attainment/nonattainment area, each of which usually comprises multiple counties and tracts. Panel (b) shows coefficient estimates from a regression that includes a treatment dummy interacted with years, controlling for year fixed effects. The dotted blue line shows point estimates and the dashed red lines show 95% confidence intervals based on standard errors that are cluster-robust at the level of counties. Both Panels are based on data from van Donkelaar et al. (2021*b*).

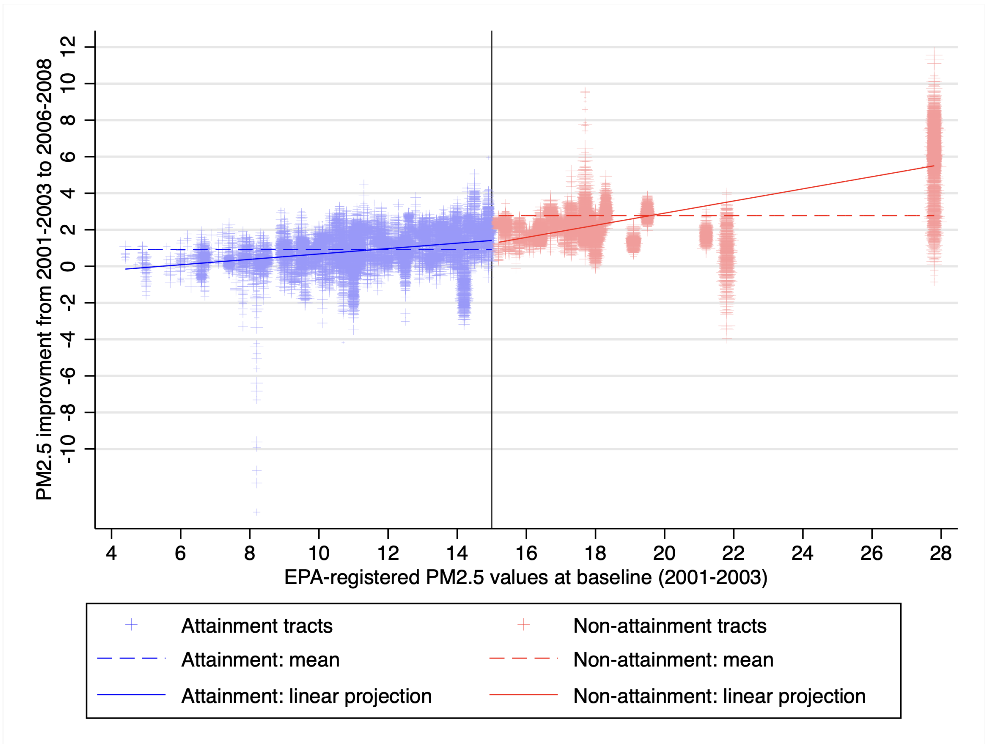


Figure A.23: Improvement in tract $PM_{2.5}$ averages and EPA-registered $PM_{2.5}$ values using Di et al. (2021)

Notes: The figure shows the improvement in $PM_{2.5}$ averages at the tract level between two periods, 2001-2003 and 2006-2008. The size of the markers reflect tract level populations. The $PM_{2.5}$ improvements are plotted against the EPA-registered $PM_{2.5}$ values of each attainment/nonattainment area, each of which usually comprises multiple counties and tracts. The dashed line plots the average $PM_{2.5}$ improvement for tracts in nonattainment and attainment areas separately, weighted by tract population. The solid lines plot the linear projection of tract level $PM_{2.5}$ improvements on the EPA-registered $PM_{2.5}$ values of the nonattainment and attainment areas separately, weighted by tract population. Based on data from Di et al. (2021).

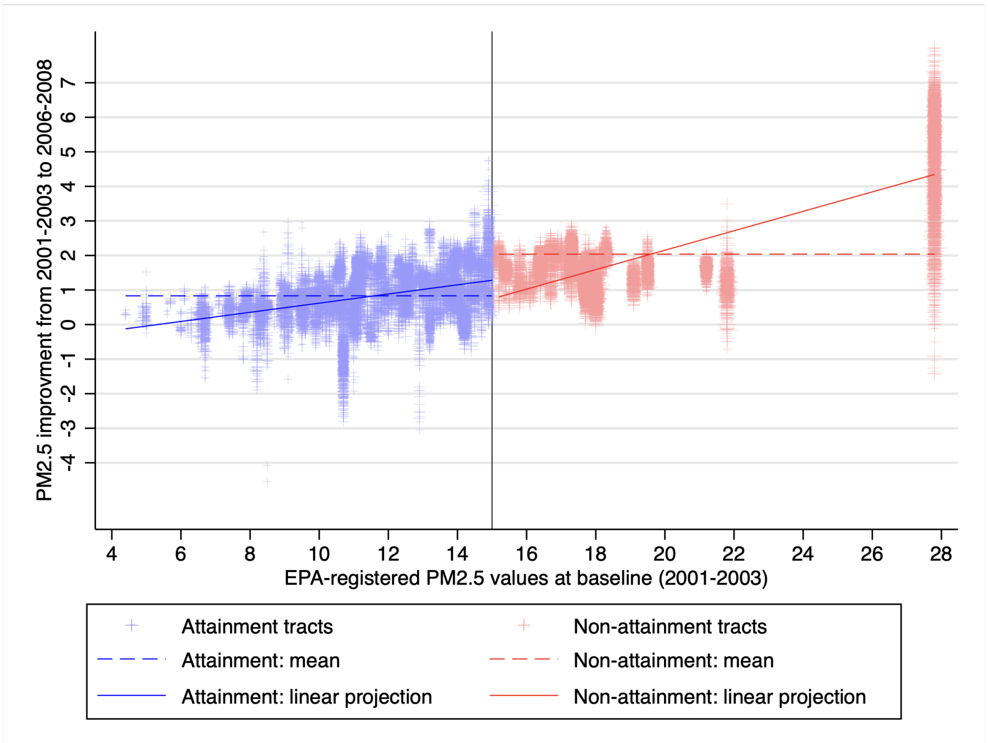


Figure A.24: Improvement in tract $PM_{2.5}$ averages and EPA-registered $PM_{2.5}$ values using van Donkelaar et al. (2021*b*)

Notes: The figure shows the improvement in $PM_{2.5}$ averages at the tract level between two periods, 2001-2003 and 2006-2008. The size of the markers reflect tract level populations. The $PM_{2.5}$ improvements are plotted against the EPA-registered $PM_{2.5}$ values of each attainment/nonattainment area, each of which usually comprises multiple counties and tracts. The dashed line plots the average $PM_{2.5}$ improvement for tracts in nonattainment and attainment areas separately, weighted by tract population. The solid lines plot the linear projection of tract level $PM_{2.5}$ improvements on the EPA-registered $PM_{2.5}$ values of the nonattainment and attainment areas separately, weighted by tract population. Based on data from van Donkelaar et al. (2021*b*).

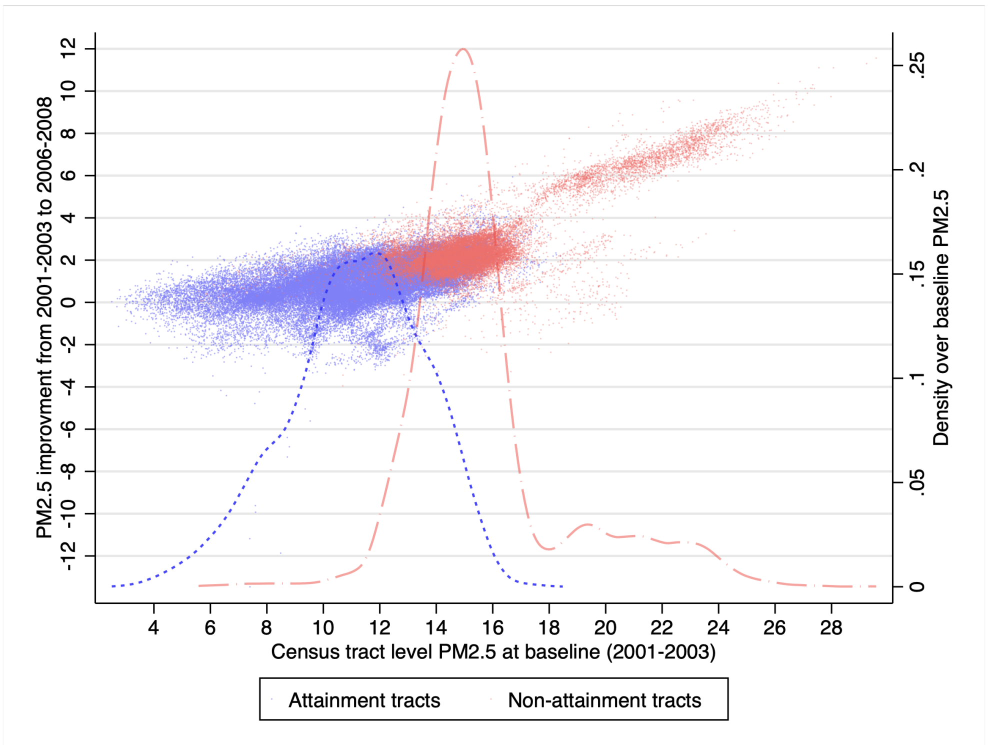


Figure A.25: Improvement in tract $PM_{2.5}$ averages and baseline $PM_{2.5}$ levels using Di et al. (2021)

Notes: The markers in the figure show the improvement in $PM_{2.5}$ averages at the tract level between two periods, 2001-2003 and 2006-2008. The $PM_{2.5}$ improvements are plotted against the baseline $PM_{2.5}$ levels of each tract, using two different colors for tracts in nonattainment and attainment areas. The kernel density (right axis) shows the overlap between nonattainment and attainment tracts in terms of baseline $PM_{2.5}$, weighted by tract population. The figure is based on data from Di et al. (2021).

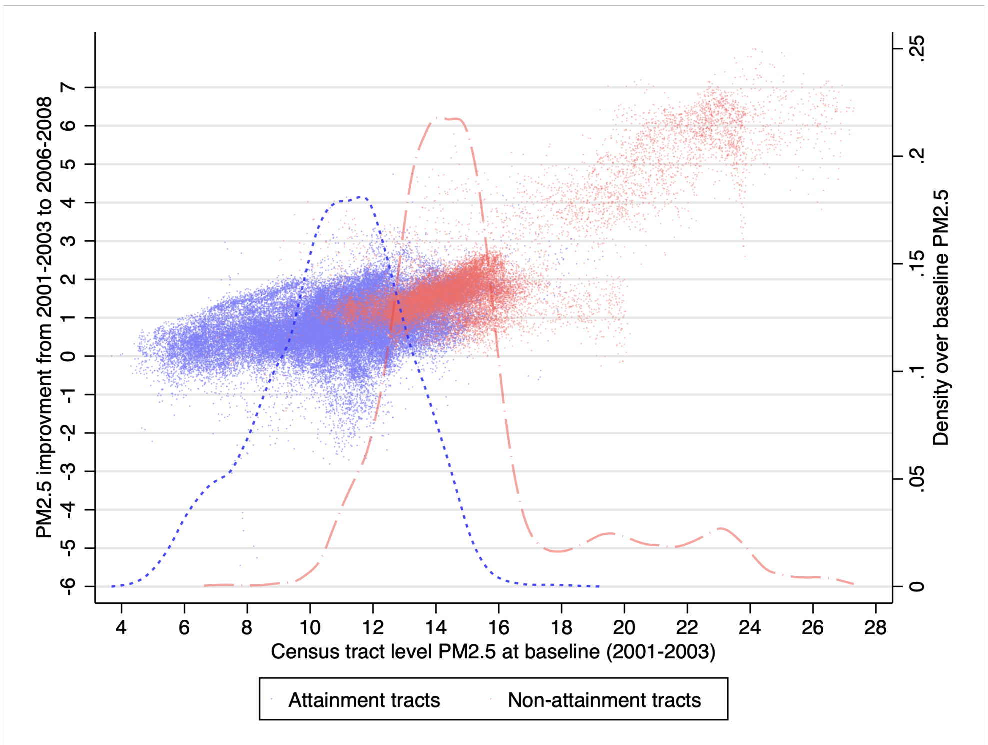


Figure A.26: Improvement in tract $PM_{2.5}$ averages and baseline $PM_{2.5}$ levels using van Donkelaar et al. (2021*b*)

Notes: The markers in the figure show the improvement in $PM_{2.5}$ averages at the tract level between two periods, 2001-2003 and 2006-2008. The $PM_{2.5}$ improvements are plotted against the baseline $PM_{2.5}$ levels of each tract, using two different colors for tracts in nonattainment and attainment areas. The kernel density (right axis) shows the overlap between nonattainment and attainment tracts in terms of baseline $PM_{2.5}$, weighted by tract population. The figure is based on data from van Donkelaar et al. (2021*b*).

Table A.23: Nonattainment status and changes in $PM_{2.5}$ using Di et al. (2021)

	ATT					LATE	
	All Tracts				with RV	Optimal Bandw.	
	DiD	DiDwb	M1DiD	M2DiD	DiD	RD0	RD1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Part A: Effect from 2001-03 to 2006-08							
Panel (a): Homogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-1.92	-0.49	-0.32	-0.74	-1.86	-0.84	-0.31
	(0.38)	(0.13)	(0.12)	(0.40)	(0.39)	(0.58)	(0.52)
Observations	72043	72043	27827	29932	47962	5234	12738
Panel (b): Placebo Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-0.46	-0.0090	-0.014	0.15	-0.56	-0.12	0.077
	(0.20)	(0.20)	(0.20)	(0.20)	(0.21)	(0.23)	(0.63)
Observations	49357	49357	20460	20068	25276	3280	4626
Panel (c): Heterogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	6.27	4.11	2.42	7.41	6.33	4.03	4.76
	(0.83)	(0.85)	(0.75)	(0.87)	(0.83)	(1.20)	(1.06)
NA(x)Baseline	-0.52	-0.33	-0.19	-0.52	-0.52	-0.32	-0.34
	(0.059)	(0.062)	(0.052)	(0.061)	(0.059)	(0.068)	(0.061)
Observations	72043	72043	27827	29932	47962	5234	12738
Implied ATE	-1.92	-1.05	-0.50	-0.75	-1.86	-1.04	-0.65
10th pct	-0.61	-0.22	-0.027	0.56	-0.55	-0.23	0.22
90th pct	-4.19	-2.48	-1.30	-3.01	-4.13	-2.45	-2.15
Part B: Effect from 2001-03 to 2011-13							
Panel (d): Homogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-2.85	-0.50	-0.23	-0.93	-2.94	-0.52	-0.73
	(0.39)	(0.11)	(0.11)	(0.42)	(0.41)	(0.22)	(0.37)
Observations	72043	72043	27827	29932	47962	3743	10459
Panel (e): Placebo Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-0.91	0.26	0.33	0.45	-1.50	0.32	1.14
	(0.18)	(0.15)	(0.18)	(0.19)	(0.21)	(0.37)	(0.63)
Observations	49357	49357	20460	20068	25276	2143	4807
Panel (f): Heterogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	6.32	0.98	4.17	8.20	6.24	6.12	4.87
	(0.83)	(0.84)	(0.70)	(0.87)	(0.83)	(1.63)	(1.06)
NA(x)Baseline	-0.59	-0.11	-0.30	-0.58	-0.59	-0.44	-0.37
	(0.058)	(0.060)	(0.049)	(0.060)	(0.058)	(0.11)	(0.067)
Observations	72043	72043	27827	29932	47962	3743	10459
Implied ATE	-2.85	-0.69	-0.52	-0.94	-2.94	-0.77	-0.87
10th pct	-1.38	-0.42	0.23	0.53	-1.46	0.34	0.052
90th pct	-5.39	-1.15	-1.82	-3.48	-5.48	-2.68	-2.46

Notes: The table shows coefficient estimates for the treatment effect of nonattainment status on the change in $PM_{2.5}$ levels between the pre- and post-treatment periods. Each panel(x)column combination is from a separate regression as described in the text. (1) uses simple DiD, (2) adds controls for baseline $PM_{2.5}$ (2001-03), (3) runs DiD using a sample matched (1-to-1) on baseline $PM_{2.5}$, (4) matches on baseline $PM_{2.5}$, tract population and population density (both 2000), (5) again uses simple DiD but with the limited sample of areas for which an EPA-registered $PM_{2.5}$ value exists, (6) and (7) use the limited sample based on optimal bandwidth selection in a regression discontinuity framework. Standard errors in parentheses are clustered at the county level. All results based on Di et al. (2021).

Table A.24: Nonattainment status and changes in $PM_{2.5}$ using van Donkelaar et al. (2021*b*)

	ATT					LATE	
	All Tracts				with RV	Optimal Bandw.	
	DiD	DiDwb	M1DiD	M2DiD	DiD	RD0	RD1
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Part A: Effect from 2001-03 to 2006-08							
Panel (a): Homogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-1.22	-0.13	-0.20	-0.63	-1.21	-0.26	0.043
	(0.33)	(0.092)	(0.16)	(0.36)	(0.34)	(0.26)	(0.35)
Observations	72043	72043	28311	29808	47962	7026	15683
Panel (b): Placebo Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	-0.38	-0.15	-0.15	-0.071	-0.52	-0.071	0.029
	(0.12)	(0.13)	(0.13)	(0.14)	(0.17)	(0.44)	(0.50)
Observations	49357	49357	20285	20056	25276	1046	4626
Panel (c): Heterogeneous Treatment Effect: from 2001-03 to 2006-08							
Nonattainment	4.76	3.47	2.52	5.34	4.78	2.04	2.89
	(0.58)	(0.59)	(0.61)	(0.61)	(0.59)	(1.33)	(0.70)
NA(x)Baseline	-0.39	-0.28	-0.19	-0.39	-0.39	-0.16	-0.20
	(0.044)	(0.045)	(0.045)	(0.044)	(0.044)	(0.093)	(0.042)
Observations	72043	72043	28311	29808	47962	7026	15683
Implied ATE	-1.22	-0.70	-0.38	-0.64	-1.21	-0.39	-0.17
10th pct	-0.12	0.063	0.15	0.46	-0.11	0.054	0.39
90th pct	-3.18	-2.07	-1.33	-2.60	-3.16	-1.19	-1.17
Part B: Effect from 2001-03 to 2011-13							
Panel (d): Homogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-2.34	-0.15	-0.072	-0.92	-2.42	-0.29	-0.54
	(0.39)	(0.090)	(0.14)	(0.45)	(0.40)	(0.100)	(0.35)
Observations	72043	72043	28311	29808	47962	3743	12997
Panel (e): Placebo Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	-0.98	0.048	0.095	0.15	-1.58	0.97	1.09
	(0.16)	(0.14)	(0.18)	(0.19)	(0.18)	(0.43)	(0.58)
Observations	49357	49357	20285	20056	25276	676	3280
Panel (f): Heterogeneous Treatment Effect: from 2001-03 to 2011-13							
Nonattainment	5.75	0.60	4.95	7.15	5.67	1.10	3.10
	(0.50)	(0.51)	(0.69)	(0.56)	(0.51)	(0.53)	(0.60)
NA(x)Baseline	-0.53	-0.058	-0.35	-0.53	-0.53	-0.098	-0.26
	(0.037)	(0.039)	(0.051)	(0.038)	(0.037)	(0.037)	(0.035)
Observations	72043	72043	28311	29808	47962	3743	12997
Implied ATE	-2.34	-0.27	-0.41	-0.93	-2.42	-0.39	-0.87
10th pct	-0.86	-0.11	0.57	0.56	-0.93	-0.11	-0.14
90th pct	-4.99	-0.56	-2.17	-3.57	-5.06	-0.87	-2.17

Notes: The table shows coefficient estimates for the treatment effect of nonattainment status on the change in $PM_{2.5}$ levels between the pre- and post-treatment periods. Each panel(x)column combination is from a separate regression as described in the text. (1) uses simple DiD, (2) adds controls for baseline $PM_{2.5}$ (2001-03), (3) runs DiD using a sample matched (1-to-1) on baseline $PM_{2.5}$, (4) matches on baseline $PM_{2.5}$, tract population and population density (both 2000), (5) again uses simple DiD but with the limited sample of areas for which an EPA-registered $PM_{2.5}$ value exists, (6) and (7) use the limited sample based on optimal bandwidth selection in a regression discontinuity framework. Standard errors in parentheses are clustered at the county level. All results based on van Donkelaar et al. (2021*b*).

Table A.25: Pollution disparities - counterfactual gap analysis using Di et al. (2021)

Panel (a): Black-White Pollution Gap

	$PM_{2.5}$ exposure		Black-White Gap		Contribution of CAA (in %) [homogeneous effect]					
Period	Black	White	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	13.48	12.32	1.16							
2006-2008	11.91	10.91	1.00	-0.16	172	44	28	66	75	28
2011-2013	9.63	9	0.63	-0.53	77	14	6	25	14	20

	$PM_{2.5}$ exposure		Black-White Gap		Contribution of CAA (in %) [heterogeneous effect]					
Period	Black	White	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	13.48	12.32	1.16							
2006-2008	11.91	10.91	1.00	-0.16	171	93	44	65	92	57
2011-2013	9.63	9	0.63	-0.53	77	18	14	25	20	23

	$PM_{2.5}$ exposure		Black-White Gap		Contribution of CAA (in %) [+race interactions]					
Period	Black	White	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	13.48	12.32	1.16							
2006-2008	11.91	10.91	1.00	-0.16	140	62	54	35	-55	-46
2011-2013	9.63	9	0.63	-0.53	83	24	31	31	-6	-2

Panel (b): Urban-Rural Pollution Gap

	$PM_{2.5}$ exposure		Urban-Rural Gap		Contribution of CAA (in %) [homogeneous effect]					
Period	Urban	Rural	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	12.94	11.51	1.43							
2006-2008	11.27	10.54	0.73	-0.70	72	18	12	28	31	12
2011-2013	9.33	8.64	0.69	-0.74	101	18	8	33	19	26

	$PM_{2.5}$ exposure		Urban-Rural Gap		Contribution of CAA (in %) [heterogeneous effect]					
Period	Urban	Rural	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	12.94	11.51	1.43							
2006-2008	11.27	10.54	0.73	-0.70	85	47	23	41	47	33
2011-2013	9.33	8.64	0.69	-0.74	115	27	25	47	37	39

	$PM_{2.5}$ exposure		Urban-Rural Gap		Contribution of CAA (in %) [+urban interactions]					
Period	Urban	Rural	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	12.94	11.51	1.43							
2006-2008	11.27	10.54	0.73	-0.70	91	53	30	47	49	33
2011-2013	9.33	8.64	0.69	-0.74	119	31	30	51	38	42

Notes: Left columns show average $PM_{2.5}$ exposure of Black, White, Urban and Rural populations, and difference between groups, as derived from Census block level pollution concentrations and population counts. Right columns show contribution of CAA nonattainment designations in 2005 based on counterfactual calculations that factor out nonattainment treatment effects as estimated in Columns 1-4, 6, and 7 of Table 1. Population data is from the 2000, 2010 and 2020 waves of the US Census, linearly interpolated for years in between. Pollution data is from Di et al. (2021).

Table A.26: Pollution disparities - counterfactual gap analysis using van Donkelaar et al. (2021*b*)

Panel (a): Black-White Pollution Gap

	$PM_{2.5}$ exposure		Black-White Gap		Contribution of CAA (in %) [homogeneous effect]					
Period	Black	White	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	13.03	11.88	1.15							
2006-2008	11.8	10.73	1.07	-0.08	228	25	37	118	48	-8
2011-2013	9.38	8.74	0.64	-0.51	66	4	2	26	8	15

	$PM_{2.5}$ exposure		Black-White Gap		Contribution of CAA (in %) [heterogeneous effect]					
Period	Black	White	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	13.03	11.88	1.15							
2006-2008	11.8	10.73	1.07	-0.08	218	124	67	110	69	27
2011-2013	9.38	8.74	0.64	-0.51	64	8	10	24	11	24

	$PM_{2.5}$ exposure		Black-White Gap		Contribution of CAA (in %) [+race interactions]					
Period	Black	White	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	13.03	11.88	1.15							
2006-2008	11.8	10.73	1.07	-0.08	174	80	51	64	191	21
2011-2013	9.38	8.74	0.64	-0.51	68	12	14	28	5	39

Panel (b): Urban-Rural Pollution Gap

	$PM_{2.5}$ exposure		Urban-Rural Gap		Contribution of CAA (in %) [homogeneous effect]					
Period	Urban	Rural	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	12.53	11.15	1.38							
2006-2008	11.19	10.23	0.96	-0.42	77	8	12	40	16	-3
2011-2013	9.17	8.26	0.91	-0.47	132	9	4	52	16	30

	$PM_{2.5}$ exposure		Urban-Rural Gap		Contribution of CAA (in %) [heterogeneous effect]					
Period	Urban	Rural	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	12.53	11.15	1.38							
2006-2008	11.19	10.23	0.96	-0.42	92	55	31	55	31	18
2011-2013	9.17	8.26	0.91	-0.47	151	18	35	71	25	58

	$PM_{2.5}$ exposure		Urban-Rural Gap		Contribution of CAA (in %) [+urban interactions]					
Period	Urban	Rural	(levels)	(change)	DiD	DIDwb	M1DiD	M2DiD	RD0	RD1
2001-2003	12.53	11.15	1.38							
2006-2008	11.19	10.23	0.96	-0.42	95	58	36	59	33	24
2011-2013	9.17	8.26	0.91	-0.47	148	15	33	68	27	56

Notes: Left columns show average $PM_{2.5}$ exposure of Black, White, Urban and Rural populations, and difference between groups, as derived from Census block level pollution concentrations and population counts. Right columns show contribution of CAA nonattainment designations in 2005 based on counterfactual calculations that factor out nonattainment treatment effects as estimated in Columns 1-4, 6, and 7 of Table 1. Population data is from the 2000, 2010 and 2020 waves of the US Census, linearly interpolated for years in between. Pollution data is from van Donkelaar et al. (2021*b*).

Table A.27: Pollution damages - instrumental variable comparison using Di et al. (2021)

	OLS (1)	DiD-IV (2)	DiDwb-IV (3)	M1DiD-IV (4)	M2DiD-IV (5)	R0-IV (6)	R1-IV (7)
Panel (a): Effect of $PM_{2.5}$ ***increases*** *on house price index growth 2001-03 to 2006-08*							
ΔPM2.5	-0.028	-0.048	-0.21	-0.048	-0.081	-0.083	0.25
	(0.014)	(0.0099)	(0.024)	(0.042)	(0.0075)	(0.093)	(0.060)
Observations	54529	54529	54529	20959	22631	3882	9729
K-P F statistic		90.1	17.6	11.9	48.6	8.51	47.5
Elasticity	-0.35	-0.60	-2.62	-0.60	-1.03	-1.04	3.21
Panel (b): Effect of $PM_{2.5}$ ***increases*** *on house price index growth 2001-03 to 2011-13*							
ΔPM2.5	-0.0037	-0.010	-0.15	0.033	-0.034	-0.025	0.055
	(0.0087)	(0.011)	(0.023)	(0.037)	(0.012)	(0.11)	(0.013)
Observations	54378	54378	54378	20867	22557	2965	7911
K-P F statistic		189.9	17.3	17.5	71.7	40.8	258.2
Elasticity	-0.046	-0.13	-1.89	0.42	-0.43	-0.31	0.69

Notes: The dependent variable is the change in the logarithm of the house price index. ΔPM2.5 is the change in $PM_{2.5}$ since 2001-03 in $\mu g/m^3$, instrumented by CAA nonattainment status for $PM_{2.5}$, allowing for heterogeneous effects in the instrument by previous PM_{10} nonattainment status and by baseline $PM_{2.5}$ levels in 2001-03. First-stage specifications in Columns 2-7 correspond to Columns 1-4, 6, and 7 in Table A.23. Standard errors in parentheses are clustered at the county level. Pollution data is from Di et al. (2021).

Table A.28: Pollution damages - instrumental variable comparison using van Donkelaar et al. (2021*b*)

	OLS (1)	DiD-IV (2)	DiDwb-IV (3)	M1DiD-IV (4)	M2DiD-IV (5)	R0-IV (6)	R1-IV (7)
Panel (a): Effect of $PM_{2.5}$ ***increases*** *on house price index growth 2001-03 to 2006-08*							
ΔPM2.5	-0.039	-0.067	-0.19	-0.032	-0.089	0.13	0.090
	(0.019)	(0.012)	(0.039)	(0.056)	(0.0086)	(0.044)	(0.089)
Observations	54529	54529	54529	21287	22442	5087	11963
K-P F statistic		96.4	23.9	264.9	152.8	41.7	42.9
Elasticity	-0.47	-0.81	-2.31	-0.39	-1.09	1.55	1.09
Panel (b): Effect of $PM_{2.5}$ ***increases*** *on house price index growth 2001-03 to 2011-13*							
ΔPM2.5	-0.0062	-0.012	-0.18	0.033	-0.032	-0.10	0.054
	(0.010)	(0.012)	(0.054)	(0.031)	(0.013)	(0.099)	(0.014)
Observations	54378	54378	54378	21199	22363	2965	9902
K-P F statistic		408.9	18.1	55.0	202.3	908.5	1038.9
Elasticity	-0.075	-0.14	-2.24	0.40	-0.39	-1.26	0.66

Notes: The dependent variable is the change in the logarithm of the house price index. ΔPM2.5 is the change in $PM_{2.5}$ since 2001-03 in $\mu g/m^3$, instrumented by CAA nonattainment status for $PM_{2.5}$, allowing for heterogeneous effects in the instrument by previous PM_{10} nonattainment status and by baseline $PM_{2.5}$ levels in 2001-03. First-stage specifications in Columns 2-7 correspond to Columns 1-4, 6, and 7 in Table A.24. Standard errors in parentheses are clustered at the county level. Pollution data is from van Donkelaar et al. (2021*b*).

REFERENCES

Abadie, A., Diamond, A. & Hainmueller, J. (2010), 'Synthetic control methods for comparative case studies: Estimating the effect of california's tobacco control program', *Journal of the American statistical Association* **105**(490), 493–505.

Arkhangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W. & Wager, S. (2021), 'Synthetic difference-in-differences', *American Economic Review* **111**(12), 4088–4118.

Calonico, S., Cattaneo, M. D. & Farrell, M. H. (2020), 'Optimal bandwidth choice for robust bias-corrected inference in regression discontinuity designs', *The Econometrics Journal* **23**(2), 192–210.

Calonico, S., Cattaneo, M. D. & Titiunik, R. (2014), 'Robust nonparametric confidence intervals for regression-discontinuity designs', *Econometrica* **82**(6), 2295–2326.

Chay, K. Y. & Greenstone, M. (2005), 'Does air quality matter? evidence from the housing market', *Journal of Political Economy* **113(2)**, 376–424.

Currie, J., Voorheis, J. & Walker, R. (2023), 'What caused racial disparities in particulate exposure to fall? new evidence from the clean air act and satellite-based measures of air quality', *American Economic Review* **113**(1), 71–97.

Curtis, E. M. (2018), 'Who loses under cap-and-trade programs? the labor market effects of the NOx budget trading program', *Review of Economics and Statistics* **100**(1), 151–166.

Deschenes, O., Greenstone, M. & Shapiro, J. S. (2017*a*), 'Defensive investments and the demand for air quality: Evidence from the NOx budget program', *American Economic Review* **107**(10), 2958–2989.

Deschenes, O., Greenstone, M. & Shapiro, J. S. (2017*b*), 'Replication data for: Defensive investments and the demand for air quality: Evidence from the NOx budget program', American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor], `http://doi.org/10.3886/E112938V1`.

Di, Q., Wei, Y., Shtein, A., Hultquist, C., Xing, X., Amini, H., Shi, L., Kloog, I., Silvern, R., Kelly, J. et al. (2021), 'Daily and annual PM2.5 concentrations for the contiguous United States, 1-km grids, v1 (2000–2016) [dataset]', Palisades, New York: NASA Socioeconomic Data and Applications Center (SEDAC), `https://doi.org/10.7927/0rvr-4538`, accessed February 2, 2022.

EPA (2022), *AirData Annual Summary Data*, U.S. Environmental Protection Agency. Available via https://www.epa.gov/outdoor-air-quality-data. Accessed February 1, 2022.

EPA (2023), *Clean Air Markets Program Data*, U.S. Environmental Protection Agency. Washington, DC: Office of Atmospheric Protection, Clean Air Markets Division. Available from EPA's Air Markets Program Data web site: https://campd.epa.gov. Accessed February 2, 2023.

Jbaily, A., Zhou, X., Liu, J., Lee, T.-H., Kamareddine, L., Verguet, S. & Dominici, F. (2022), 'Air pollution exposure disparities across us population and income groups', *Nature* **601**(7892), 228–233.

Meng, J., Li, C., Martin, R. V., van Donkelaar, A., Hystad, P. & Brauer, M. (2019), 'Historical PM2.5 across North America [dataset]', Washington University of St. Louis Atmospheric Composition Analysis Group, `https://sites.wustl.edu/acag/datasets/historical-pm2-5-across-north-america/`, accessed February 2, 2022.

van Donkelaar, A., Hammer, M. S., Bindle, L., Brauer, M., Brook, J. R., Garay, M. J., Hsu, N. C., Kalashnikova, O. V., Kahn, R. A., Lee, C. et al. (2021), 'Surface pm2.5 data (v5.gl.03) [dataset]', Washington University of St. Louis Atmospheric Composition Analysis Group, `https://sites.wustl.edu/acag/datasets/surface-pm2-5/`, accessed February 2, 2022.