Unwatched Pollution:

The Effect of Intermittent Monitoring on Air Quality

Eric Zou

October 2019

Assistant Professor of Economics
University of Oregon
www.eric-zou.com

Enforcement with bigger, better, cooler data!

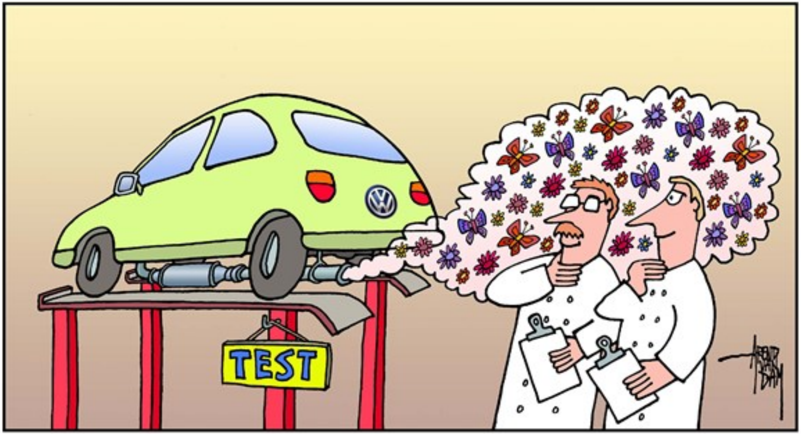
Police drone:
Driving restriction enforcement+

Spy satellite:
International sanction enforcement*

Source: +news.xinhuanet.cn; *U.S. Department of the Treasury

Today: Intermittent monitoring

- **Intermittent monitoring**: widely used cost-reduction tool in environmental enforcement
    - e.g., periodic factory inspections; car exhaust testings
- Problem: works only if **strategic responses** are difficult
    - Polluters don't know about monitoring schedule
    - Polluters don't have capacity to turn off during monitoring and right back on after monitoring



exhaust scandal: faked emissions test

Source: The Cagle Post

This paper

- Retrospective analysis of Clean Air Act's outdoor particulate matter (PM) enforcement

  - **Setting:** every county must show compliance using monitoring data e.g., $PM_{2.5}$ annual mean < 15 ug/m^3, with no days > 35 ug/m^3

  - **Non-compliance**: Extra emission reduction; higher barriers of entry $\Rightarrow$ Significant losses in employment & factory productivity (e.g., Walker, 2013; Greenstone, List and Syverson, 2012)

  - **Intermittent monitoring:** EPA permits many monitors to follow cyclical once-every-six-day (**"1-in-6 day"**) monitoring schedule (Akland, 1972)

- I use satellite measure of air quality to detect strategic responses

This paper

- Retrospective analysis of Clean Air Act's outdoor particulate matter (PM) enforcement
    - **Setting:** every county must show compliance using monitoring data e.g., $PM_{2.5}$ annual mean < 15 ug/m^3, with no days > 35 ug/m^3
    - **Non-compliance**: Extra emission reduction; higher barriers of entry $\Rightarrow$ Significant losses in employment & factory productivity (e.g., Walker, 2013; Greenstone, List and Syverson, 2012)
    - **Intermittent monitoring:** EPA permits many monitors to follow cyclical once-every-six-day (**"1-in-6 day"**) monitoring schedule (Akland, 1972)
- I use satellite measure of air quality to detect strategic responses

This paper

- Retrospective analysis of Clean Air Act's outdoor particulate matter (PM) enforcement
  - **Setting:** every county must show compliance using monitoring data e.g., $PM_{2.5}$ annual mean < 15 ug/m^3, with no days > 35 ug/m^3
  - **Non-compliance**: Extra emission reduction; higher barriers of entry $\Rightarrow$ Significant losses in employment & factory productivity (e.g., Walker, 2013; Greenstone, List and Syverson, 2012)
  - **Intermittent monitoring:** EPA permits many monitors to follow cyclical once-every-six-day (**"1-in-6 day"**) monitoring schedule (Akland, 1972)
- I use satellite measure of air quality to detect strategic responses

This paper

- Retrospective analysis of Clean Air Act's outdoor particulate matter (PM) enforcement
  - **Setting:** every county must show compliance using monitoring data e.g., $PM_{2.5}$ annual mean < 15 ug/m^3, with no days > 35 ug/m^3
  - **Non-compliance**: Extra emission reduction; higher barriers of entry $\Rightarrow$ Significant losses in employment & factory productivity (e.g., Walker, 2013; Greenstone, List and Syverson, 2012)
  - **Intermittent monitoring:** EPA permits many monitors to follow cyclical once-every-six-day (**"1-in-6 day"**) monitoring schedule (Akland, 1972)
- I use satellite measure of air quality to detect strategic responses

Ambient particulate matter (PM) monitoring site

Source: U.S. EPA

2001 Monitoring Schedule

1/6-Day &1/3-Day Monitoring Schedule for TSP, Pb, PM-10, PM-2.5, and VOC

= 1/6 schedule

January

Su	M	Tu	W	Th	F	Sa
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

February

Su	M	Tu	W	Th	F	Sa
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28			

March

Su	M	Tu	W	Th	F	Sa
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	31

April

Su	M	Tu	W	Th	F	Sa
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

May

Su	M	Tu	W	Th	F	Sa
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

June

Su	M	Tu	W	Th	F	Sa
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

July

Su	M	Tu	W	Th	F	Sa
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

August

Su	M	Tu	W	Th	F	Sa
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

September

Su	M	Tu	W	Th	F	Sa
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						

October

Su	M	Tu	W	Th	F	Sa
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

November

Su	M	Tu	W	Th	F	Sa
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

December

Su	M	Tu	W	Th	F	Sa
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

2001 Monitoring Schedule

1/6-Day &1/3-Day Monitoring Schedule for TSP, Pb, PM-10, PM-2.5, and VOC

= 1/6 schedule

January

Su	M	Tu	W	Th	F	Sa
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

February

Su	M	Tu	W	Th	F
				1	2
4	5	6	7	8	9
11	12	13	14	15	16
18	19	20	21	22	23
25	26	27	28		

April

Su	M	Tu	W	Th	F	Sa
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30					

May

Su	M	Tu	W	Th	F	Sa
		1	2	3	4	5
6	7	8	9	10	11	12
13	14	15	16	17	18	19
20	21	22	23	24	25	26
27	28	29	30	31		

Su	M	Tu	W	Th	F	Sa
					1	2
3	4	5	6	7	8	9
10	11	12	13	14	15	16
17	18	19	20	21	22	23
24	25	26	27	28	29	30

July

Su	M	Tu	W	Th	F	Sa
1	2	3	4	5	6	7
8	9	10	11	12	13	14
15	16	17	18	19	20	21
22	23	24	25	26	27	28
29	30	31				

August

Su	M	Tu	W	Th	F	Sa
			1	2	3	4
5	6	7	8	9	10	11
12	13	14	15	16	17	18
19	20	21	22	23	24	25
26	27	28	29	30	31	

September

Su	M	Tu	W	Th	F	Sa
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30						

October

Su	M	Tu	W	Th	F	Sa
	1	2	3	4	5	6
7	8	9	10	11	12	13
14	15	16	17	18	19	20
21	22	23	24	25	26	27
28	29	30	31			

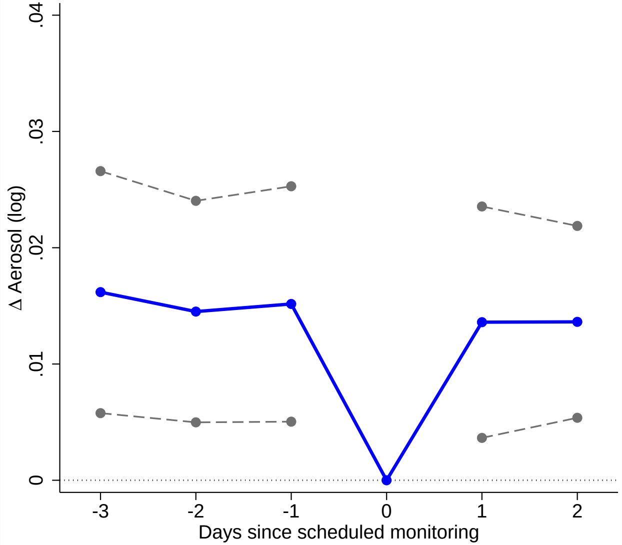
November

Su	M	Tu	W	Th	F	Sa
				1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

December

Su	M	Tu	W	Th	F	Sa
						1
2	3	4	5	6	7	8
9	10	11	12	13	14	15
16	17	18	19	20	21	22
23	24	25	26	27	28	29
30	31					

Preview: The "pollution gap"
Satellite detects less particle pollution when monitoring is on



Notes: N=685,060. Sample spans 2001-2013. Dep var = aerosol optical depth within 10km grid cell containing a 1/6day monitoring site. Dashed lines show 95% CI using SEs clustered at the county level.

Main findings

- **Strategic responses:** Satellite detects more particle pollution when monitors are off (avg. $= 1.6\%$ gap; "hot-spot" areas $> 8\%$ gap)
- **Sources:** Large gap consistently correlated with presence of certain industries (e.g., wood mills)
- **Coordination:** Evidence of state/local government gaming (e.g., strategic air quality warning)
- **Outcomes:** Health (elderly mortality) and human capital (test scores and crime) impacts that justify upgrading to continuous monitoring

Main findings

- **Strategic responses:** Satellite detects more particle pollution when monitors are off (avg. $= 1.6\%$ gap; "hot-spot" areas $> 8\%$ gap)
- **Sources:** Large gap consistently correlated with presence of certain industries (e.g., wood mills)
- **Coordination:** Evidence of state/local government gaming (e.g., strategic air quality warning)
- **Outcomes:** Health (elderly mortality) and human capital (test scores and crime) impacts that justify upgrading to continuous monitoring

Related literature

- **Detecting polluter cheating using big data** (Oliva, 2015; Reyneart and Sallee, 2017; Vollaard, 2017)
    - What's new here: ambient pollution regulation, U.S. setting

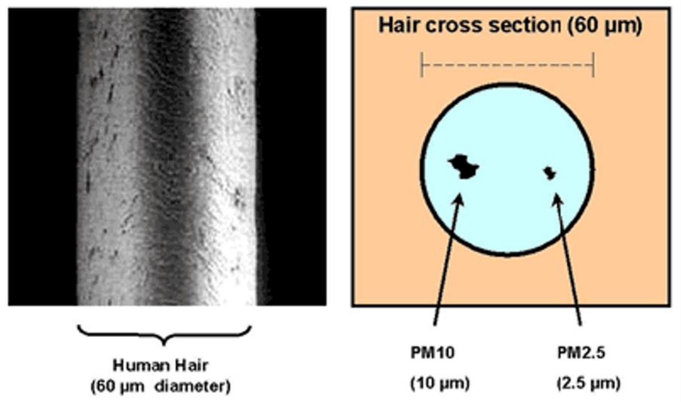
- **Using satellite data in environmental surveillance** (Kittaka et al., 2004; Ruminski et al., 2006; Duncan et al., 2014; Donaldson and Storeygard, 2016)
    - What's new here: informs regulatory decision-making (see also: Grainger, Schreiber and Chang, mimeo)

- **Economics of incomplete regulation, enforcement and monitoring** (e.g., Becker, 1968; Malik, 1990; Fowlie, 2009; Gray and Shimshack, 2011; Duflo et al., 2013; Shimshack, 2014)

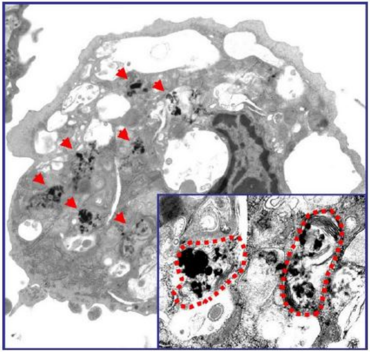
Overview

- Institutional background
- Data
- Identification of pollution gap
- Sources of pollution gap
- Consequences of pollution gap

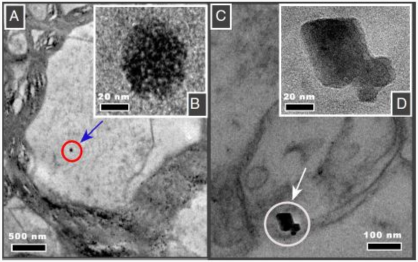
Particulate matter



Particulate matter in lung and brain



PM in lung tissue[+]



PM in brain tissue*

Source: [+]Araujo, et al., *Circulation Research* 2008; *Maher et al., *PNAS* 2016

Particulate matter monitoring

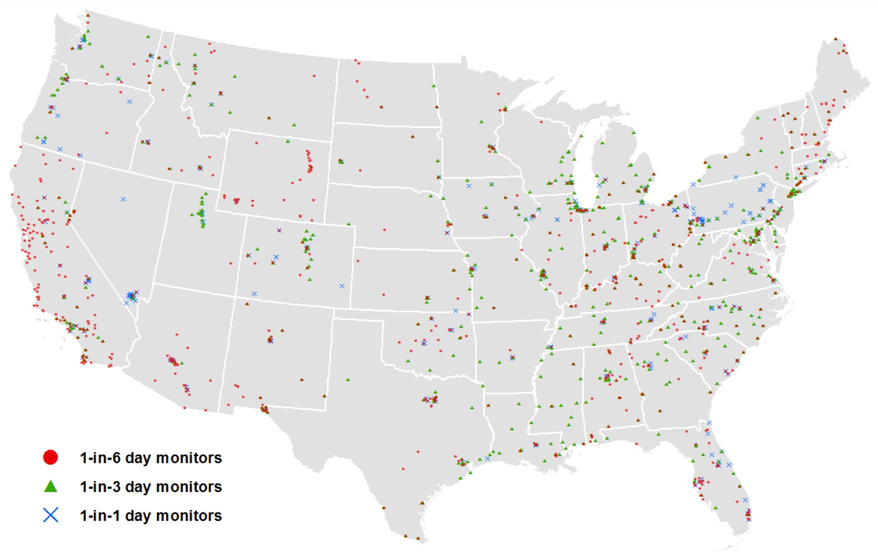
Ex: Annualized per-monitor cost of PM_{10} monitoring (2013$)

	1-in-6 day	1-in-1 day
Capital	$4,434	$5,927
Operating & maintenance	$16,596	$34,985
Total	**$21,030**	**$40,912**

Source: U.S. EPA (1993)

- Annual cost if "upgrading" all 1-in-6 day monitors to 1-in-1 day $\approx$ $12m
- Status quo spending on the entire PM monitoring network $\approx$ $48m
- Vast majority of monitors follow one of three types of schedule: 1-in-6 day (42% of monitors), 1-in-3 day (33%), and 1-in-1 day (22%)

Particulate matter monitor network, 2001



Notes: 2001-2013 avg: 1,750 monitors in 1,240 locations, spanning 640 counties that account for 70% of U.S. population.

Overview

- Institutional background
- Data
- Identification of the pollution gap
- Sources of the pollution gap
- Consequences of the pollution gap

Data

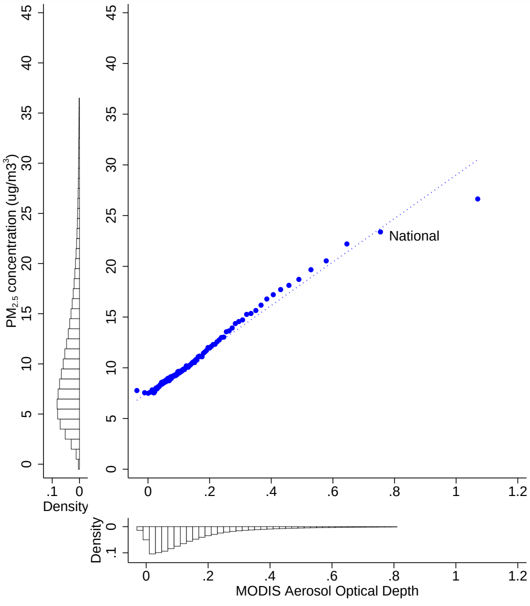
- EPA Air Quality System monitor characteristics (2001-2013)
  - Monitoring schedule
  - Latitude & longitude location

- NASA Moderate Resolution Imaging Spectroradiometer (MODIS, 2001-2013)
  - Measure: **Aerosol** optical depth
  - Captures: atmospheric particles, e.g. nitrates, sulfates, black carbon
  - Resolution: 10km$\times$10km grid - by - daily

$\Rightarrow$ Baseline outcome variable: **aerosol level around a monitor**

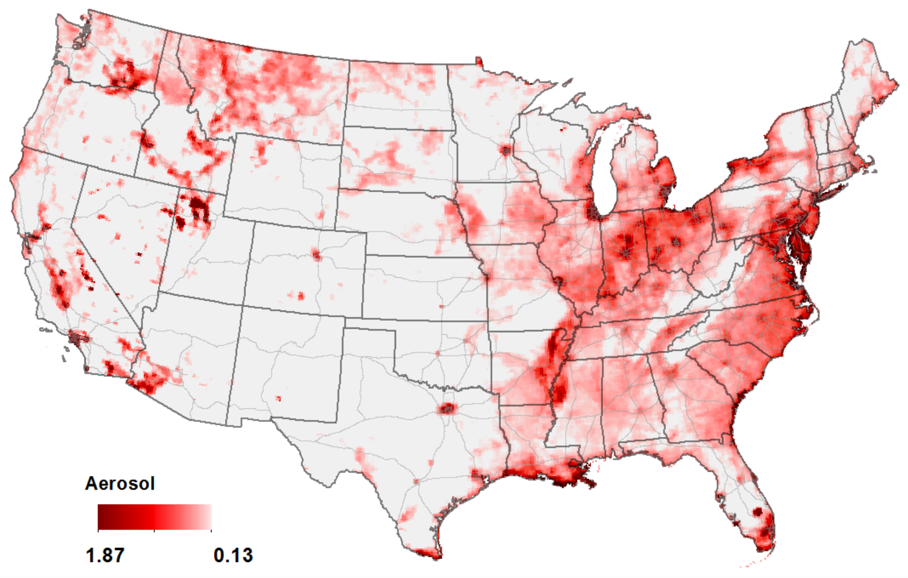
- i.e. aerosol in the 10km$\times$10km grid that contains the monitor

Correlation: $PM_{2.5}$ vs. Aerosol Concentration



Notes: Sample restricts to 10km$\times$10km grid cells that contain $PM_{2.5}$ monitors

Aerosol Concentration, 2001-2013 Grid Level Average



Notes: Map shows 10km×10km lvl 13-yr avg. aerosol optical depth, for cells with above avg. value.

Overview

- Institutional background
- Data
- Identification of pollution gap
- Sources of pollution gap
- Consequences of pollution gap

Identification strategy

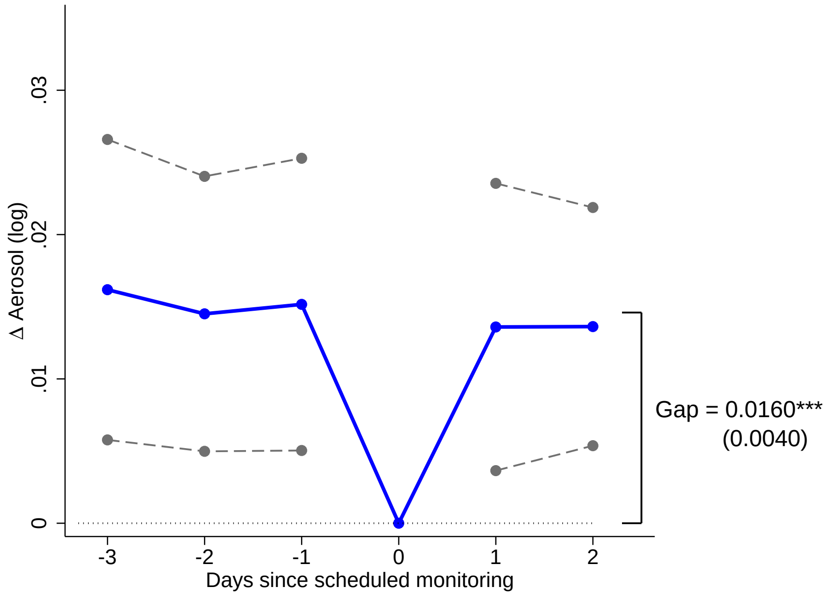
- Simple "off-day" vs. "on-day" pollution comparison

$$\overbrace{Aerosol_{st}}^{\substack{\text{(log) Aerosol lvl} \\ \text{at monitor } s \\ \text{on date } t}} = \beta \cdot \underbrace{1(Off\text{-}days)_t}_{\substack{\text{1 if monitoring} \\ \text{is scheduled-off}}} + \overbrace{T_t}^{\substack{\text{Seasonal} \\ \text{ctrls.}}} + \underbrace{\alpha_s}_{\substack{\text{Monitor} \\ \text{FEs}}} + \overbrace{X_{st}}^{\substack{\text{Weather} \\ \text{ctrls.}}} \gamma + \varepsilon_{st}$$

  - Seasonal controls: year, month-of-year, day-of-week fixed effects
  - Weather controls: daily temperature, precipitation, wind conditions
  - Standard errors clustered at the county level

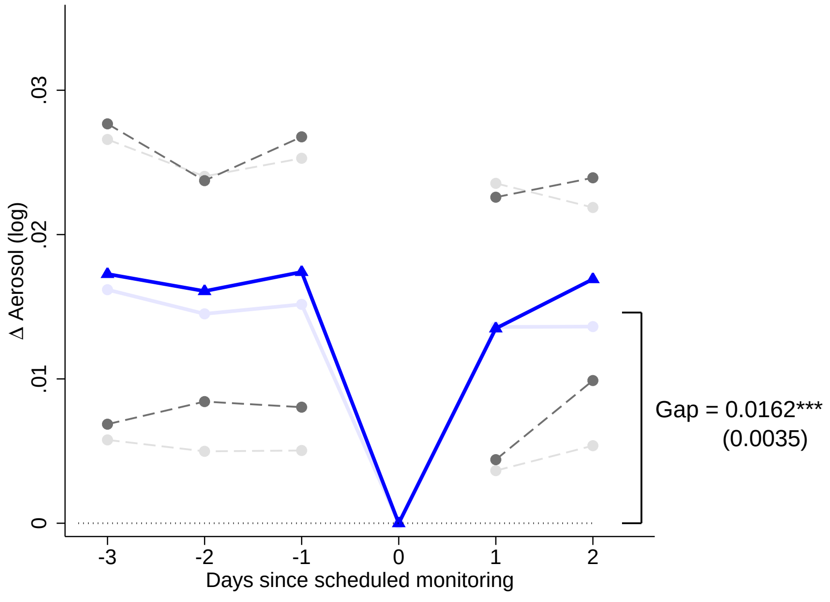
- **Identification**: nothing affects air quality on a 1-in-6 day basis, except for the 1-in-6 day monitoring schedule

The pollution gap: No-control specification



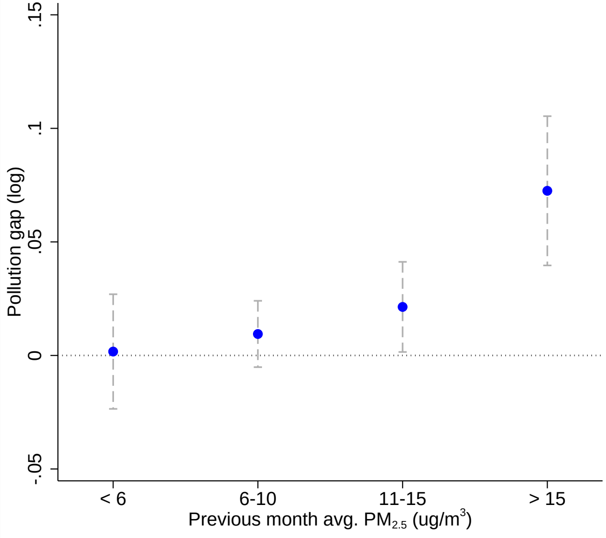
Notes: N=685,060. Dep var = aerosol optical depth within 10km grid cell containing a 1/6day monitoring site. Dashed lines show 95% CI using SEs clustered at the county level.

The pollution gap: Full-control specification



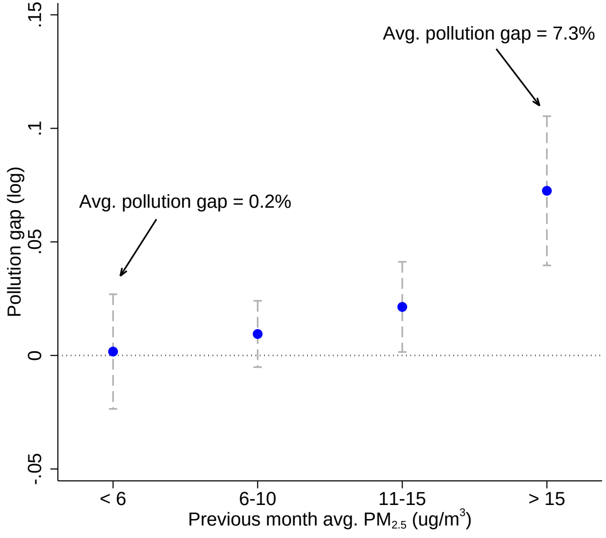
Notes: N=685,060. Dep var = aerosol optical depth within 10km grid cell containing a 1/6day monitoring site. Dashed lines show 95% CI using SEs clustered at the county level.

Heterogeneity by previous month's average $PM_{2.5}$



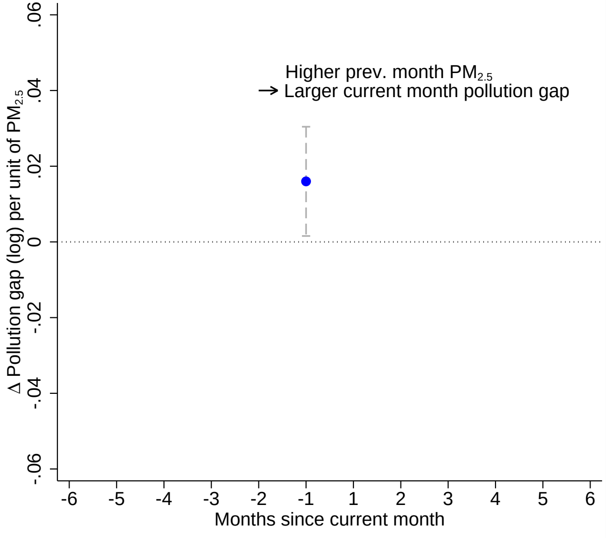
Notes: Interaction of the pollution gap with bins of realized $PM_{2.5}$ in the past month, controlling for interactions with other five lags and all six leads. Regression includes fixed effects dummies (site, year, month-of-year, and day-of-week) and weather controls. Dashed range bars plot 95% confidence intervals constructed using standard errors clustered at the county level.

Heterogeneity by previous month's average $PM_{2.5}$



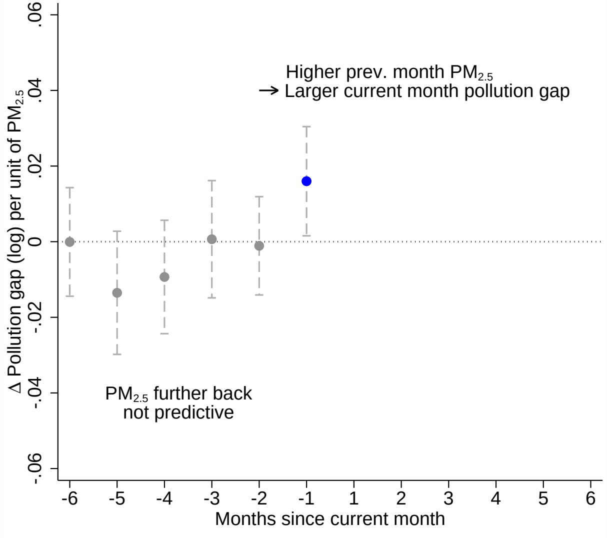
Notes: Interaction of the pollution gap with bins of realized $PM_{2.5}$ in the past month, controlling for interactions with other five lags and all six leads. Regression includes fixed effects dummies (site, year, month-of-year, and day-of-week) and weather controls. Dashed range bars plot 95% confidence intervals constructed using standard errors clustered at the county level.

Heterogeneity by leads & lags of monthly average $PM_{2.5}$



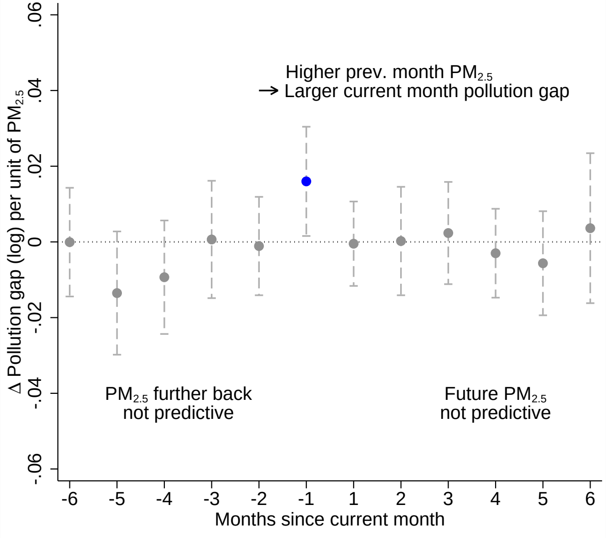
Notes: Interaction of the 1-in-6 day pollution gap with six lags and six leads of the county's realized $PM_{2.5}$ concentration. Regression includes fixed effects dummies (site, year, month-of-year, and day-of-week) and weather controls. Dashed range bars plot 95% confidence intervals constructed using standard errors clustered at the county level.

Heterogeneity by leads & lags of monthly average $PM_{2.5}$



Notes: Interaction of the 1-in-6 day pollution gap with six lags and six leads of the county's realized $PM_{2.5}$ concentration. Regression includes fixed effects dummies (site, year, month-of-year, and day-of-week) and weather controls. Dashed range bars plot 95% confidence intervals constructed using standard errors clustered at the county level.

Heterogeneity by leads & lags of monthly average $PM_{2.5}$



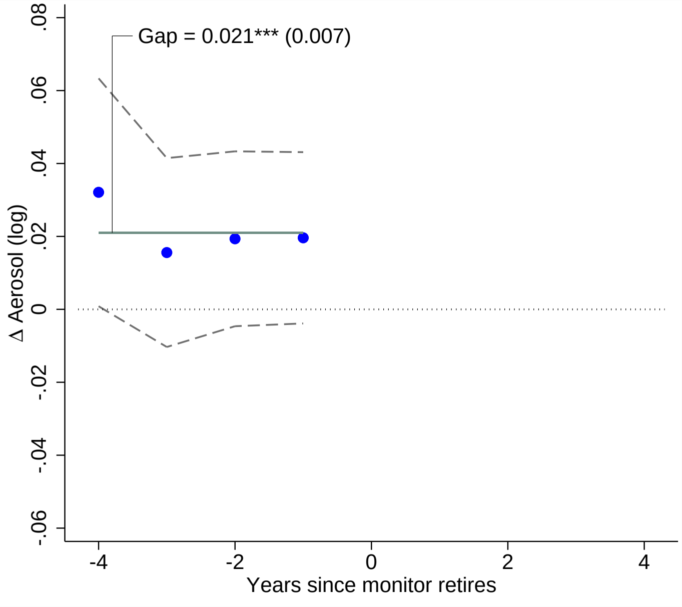
Notes: Interaction of the 1-in-6 day pollution gap with six lags and six leads of the county's realized $PM_{2.5}$ concentration. Regression includes fixed effects dummies (site, year, month-of-year, and day-of-week) and weather controls. Dashed range bars plot 95% confidence intervals constructed using standard errors clustered at the county level.

Placebo Tests

- **Identification assumption:** no 1-in-6 day monitoring policy, no 1-in-6 day pollution gap

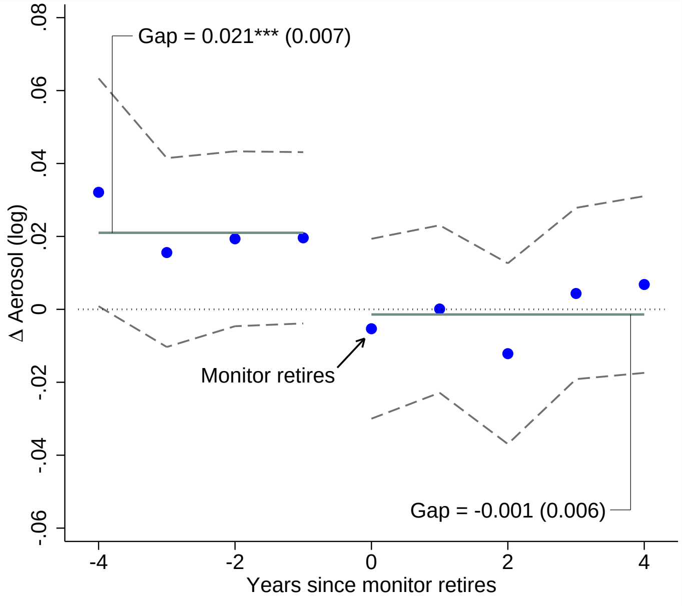
- **Placebo test idea:** Explore variation in monitoring frequency ...
    - ... across areas: regions that operate everyday monitoring
    - ... over time: retirement of 1-in-6 day monitors

Placebo test: 1-in-6 day monitor retirement



Notes: N=403,959. Dashed lines show 95% CI using SEs clustered at the county level.

Placebo test: 1-in-6 day monitor retirement



Notes: N=403,959. Dashed lines show 95% CI using SEs clustered at the county level.

Overview

- Institutional background
- Data
- Identification of pollution gap
- Sources of pollution gap
- Consequences of pollution gap

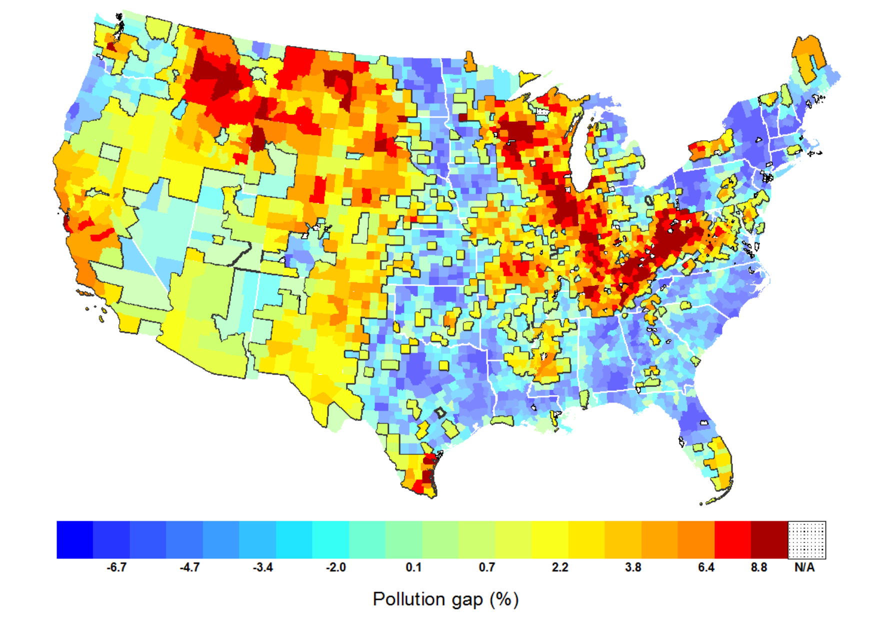
Sources of pollution gap

- Previous section documents avg pollution gap at a typical 1-in-6 day monitor

- This section: concrete evidence on sources of gaming
  - Regions?
  - Industries?
  - Coordination?

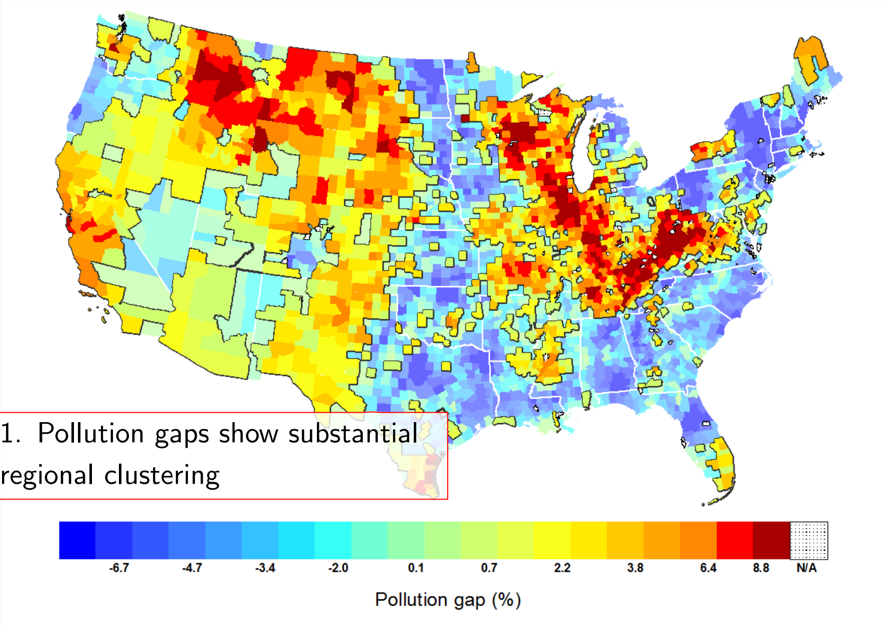
- Discover sources of gaming using existing data on polluters
  - Census: County Business Patterns
  - Polluter registries: National Emissions Inventory; Toxic Release Inventory

County level pollution gap estimates



Notes: Each county-lvl reg contains ≈35,000 obs at 10×10km cell×daily lvl. "N/A" = reg with fewer < 6,000 obs.

County level pollution gap estimates

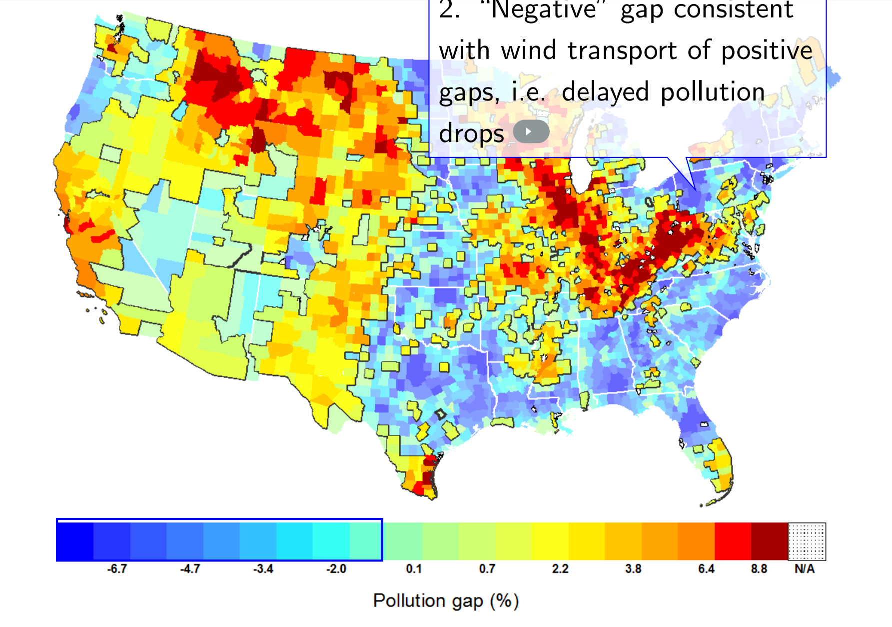


1. Pollution gaps show substantial regional clustering

Notes: Each county-lvl reg contains ≈35,000 obs at 10×10km cell×daily lvl. "N/A" = reg with fewer < 6,000 obs.

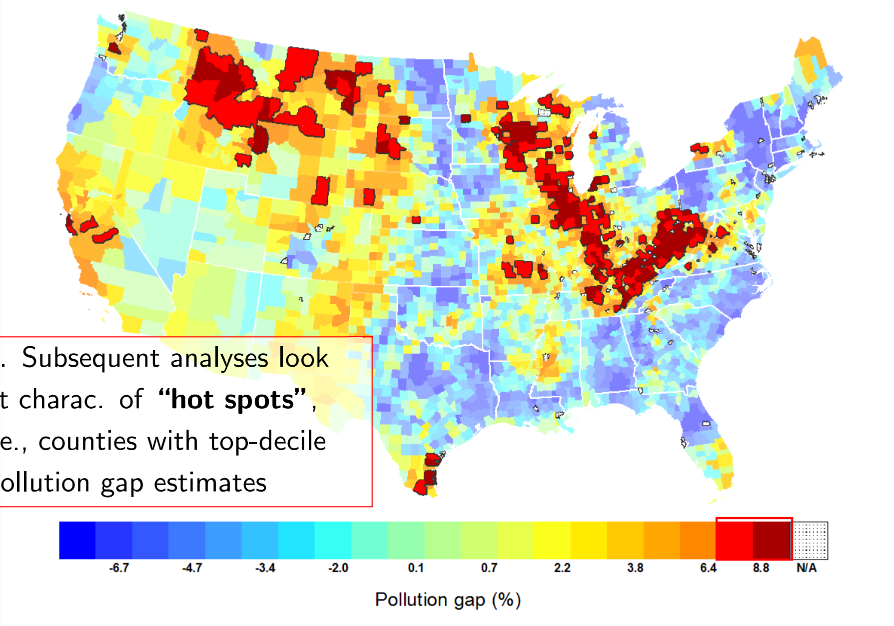
County level pollution gap estimates

2. "Negative" gap consistent with wind transport of positive gaps, i.e. delayed pollution drops



Notes: Each county-lvl reg contains ≈35,000 obs at 10×10km cell×daily lvl. "N/A" = reg with fewer < 6,000 obs.

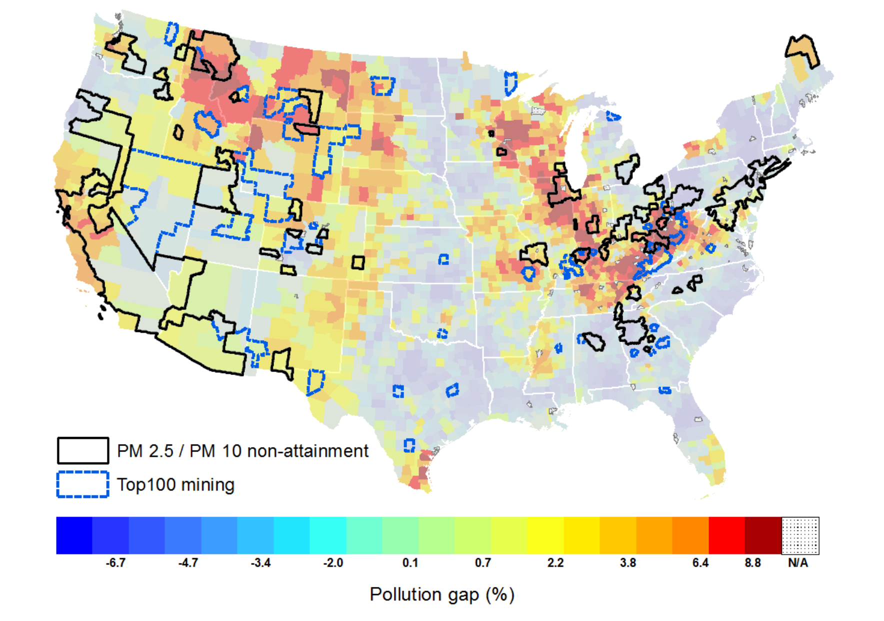
County level pollution gap estimates



3. Subsequent analyses look at charac. of **"hot spots"**, i.e., counties with top-decile pollution gap estimates

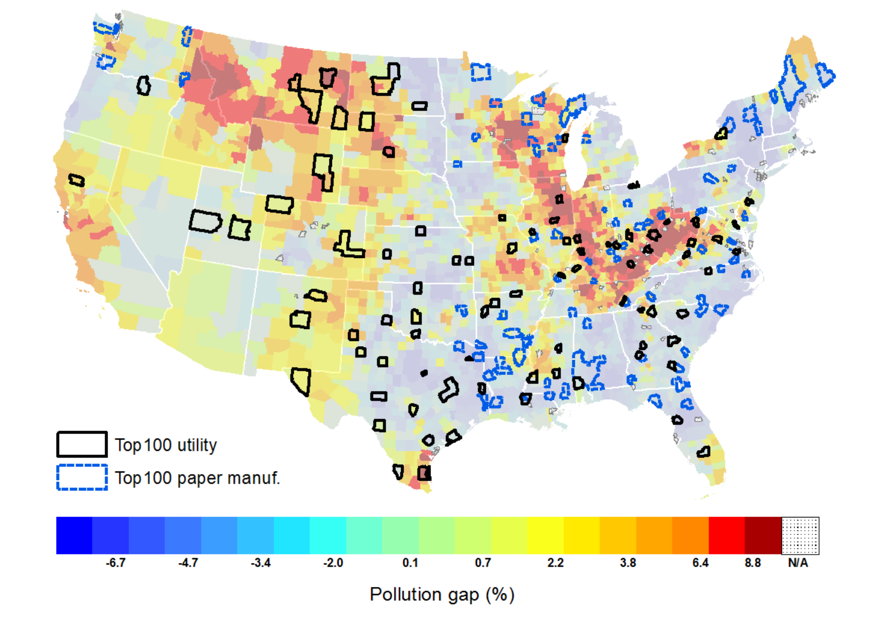
Notes: Each county-lvl reg contains ≈35,000 obs at 10×10km cell×daily lvl. "N/A" = reg with fewer < 6,000 obs.

Example: Strong cross-sectional correlates



Notes: "Top100" defined in terms of industry's employment relative to county total (County Business Pattern, 2001-2013).

Example: Weak cross-sectional correlates



Notes: "Top100" defined in terms of industry's employment relative to county total (County Business Pattern, 2001-2013).

Summary: Industry correlates of “hot-spot” counties

- Not causal effects of industries

- But, highlight robust industry correlates of “hot-spot” counties across specifications changes in ...
  - ... **geo restriction:** all counties → <50 miles hot-spot counties
  - ... **industry ctrls:** polluting industries only → all industries
  - ... **sources of variation:** national cross-section → state FEs
  - ... **model sparsity:** OLS → LASSO

- Consistent winner of horse races:
  1. Wood product manuf.
     (“runner-ups”: mining, chemical product manuf.)
  2. Highway

Summary: Industry correlates of "hot-spot" counties

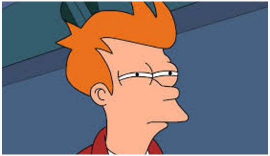
- Not causal effects of industries
- But, highlight robust industry correlates of "hot-spot" counties across specifications changes in ...
  - ... **geo restriction:** all counties → <50 miles hot-spot counties
  - ... **industry ctrls:** polluting industries only → all industries
  - ... **sources of variation:** national cross-section → state FEs
  - ... **model sparsity:** OLS → LASSO
- Consistent winner of horse races:
  1. Wood product manuf.
     ("runner-ups": mining, chemical product manuf.)
  2. Highway (!)

Indirect evidence of local government coordination

- Pollution gap correlates with **previous month's** PM readings
- Pollution gap disappears in **same year** when 1-in-6 day monitors retire
- Pollution gap shows up near **highways**

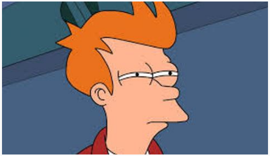
Indirect evidence of local government coordination

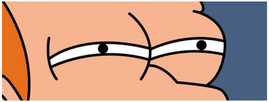
- Pollution gap correlates with **previous month's** PM readings
- Pollution gap disappears in **same year** when 1-in-6 day monitors retire
- Pollution gap shows up near **highways**



Indirect evidence of local government coordination

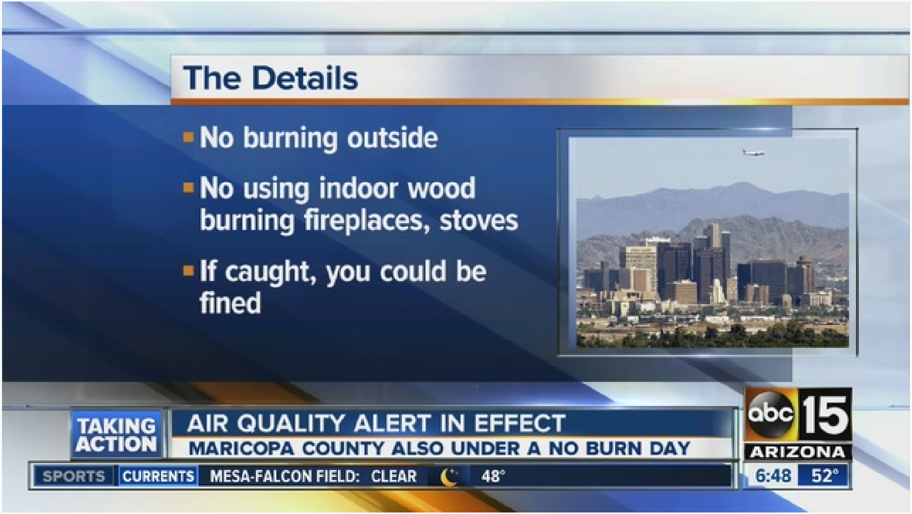
- Pollution gap correlates with **previous month's** PM readings
- Pollution gap disappears in **same year** when 1-in-6 day monitors retire
- Pollution gap shows up near **highways**





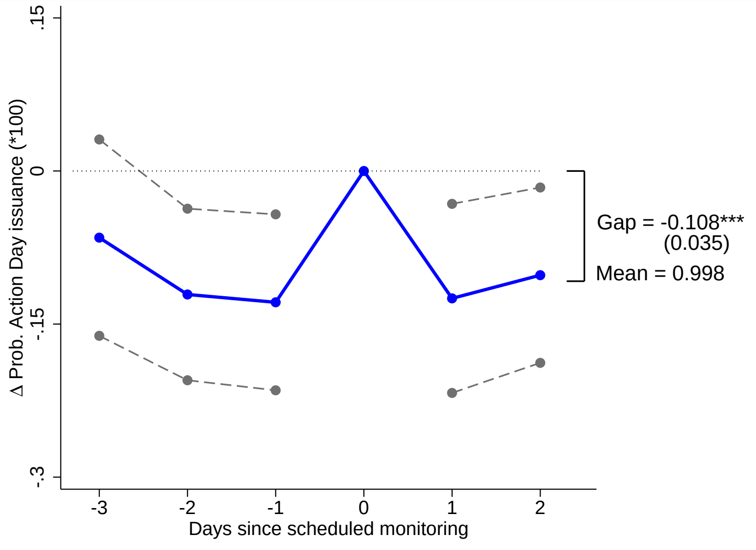
Strategic "Action Days" warnings

- A public warning infrastructure voluntarily adopted by local governments
- Use mass media to advise citizens "take actions" to spare the air, when pollution is expected to be high
  - Avoid burning outside
  - Reduce vehicle idling; carpool; use public transportation
  - Energy conservation
- Effective in influencing outdoor activities and traffic use (Neidell, 2007; Cutter and Neidell, 2009; Graff Zivin and Neidell, 2009)


The Details
No burning outside
No using indoor wood burning fireplaces, stoves
If caught, you could be fined
TAKING ACTION
AIR QUALITY ALERT IN EFFECT
MARICOPA COUNTY ALSO UNDER A NO BURN DAY
abc 15
ARIZONA
SPORTS
CURRENTS
MESA-FALCON FIELD: CLEAR
48°
6:48
52°


HIGH POLLUTION
ADVISORY
CARPOOL - USE BUS
USE 10
WATKINS

"Action Day" warnings, by 1-in-6 day PM monitoring cycle



Notes: N=624,663. Sample spans 2004-2013. Dep var is dummy for Action Day issuance at the CBSA×day level, adjusted for consecutive issuance of alerts. Results similar with or without controls. Dashed lines show 95% CI using SEs clustered at the CBSA level.

Strategic “Action Day” warnings: Additional evidence

- State heterogeneity
- More strategic warnings in non-attainment areas
- Warnings effective in manipulating air quality ($>$ 6% pollution gap with warning)
- But, strategic warnings may not explain the entire pollution gap ($\approx$ 1.2% pollution gap without warning, or in regions with no “Action Day” programs)

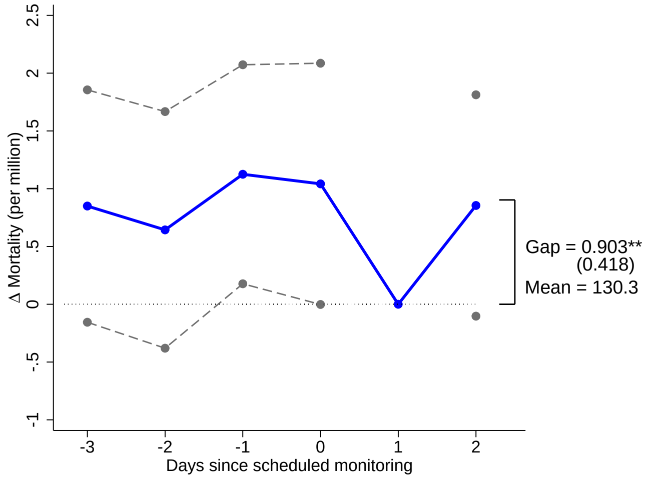
Overview

- Institutional background
- Data
- Identification of pollution gap
- Sources of pollution gap
- Consequences of pollution gap

Mortality effects of the pollution gap

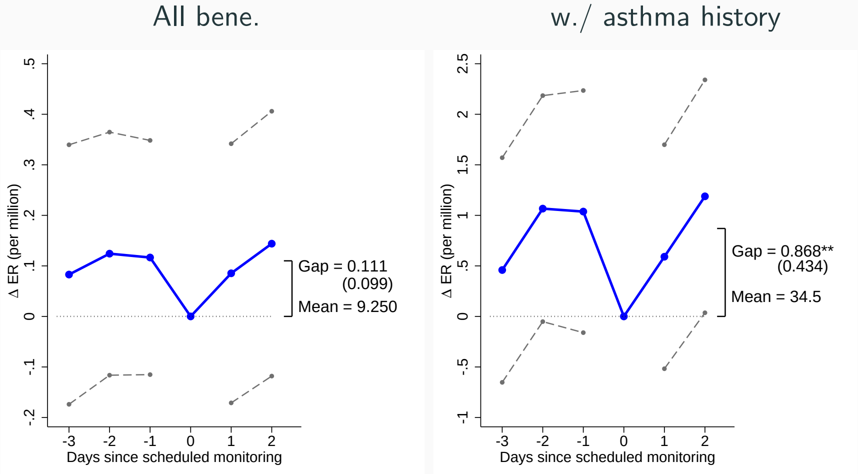
- In on-going work with Nolan Miller and David Molitor, we test for mortality consequence of intermittent monitoring
- Data: daily elderly (age 65+) mortality rate, constructed from Medicare administrative records on the universe of beneficiaries from 2001 - 2011

Medicare pop. mortality rate, by 1-in-6 day PM monitoring cycle



Notes: N=432,825. Sample spans 2001-2011. Sample includes all counties that monitor PM on a 1-in-6 day basis. Day 0 corresponds to the monitoring day. Mortality rate on day 1 is normalized to zero. Results similar with or without controls. Dashed lines show 95% CI using SEs clustered at the county level.

Asthma emergency room visits, by 1-in-6 day PM monitoring cycle



Notes: N=432,825. Sample spans 2001-2011. Sample includes all counties that monitor PM on a 1-in-6 day basis. Day 0 corresponds to the monitoring day. ER rate on day 0 is normalized to zero. Results similar with or without controls. Dashed lines show 95% CI using SEs clustered at the county level.

Mortality analysis: Additional evidence and takeaway

- Similar effects with or without controls
- No evidence of mortality effects in 1-in-1 day counties
- **EPA-style cost-benefit calculation suggests mortality costs exceed savings from intermittent monitoring**

Conclusion

- Retrospective analysis of a decades-old intermittent monitoring rule in ambient PM enforcement
  - Satellite-based evidence of strategic gaming against monitoring intermittency
  - Data-driven detection of potential sources
  - Illustration of local government coordination
  - Public health justification for upgrading to everyday monitoring
- Advanced monitoring in environmental regulation
  - e.g., EPA's **"Next Generation Compliance"** initiative (Giles, 2013)

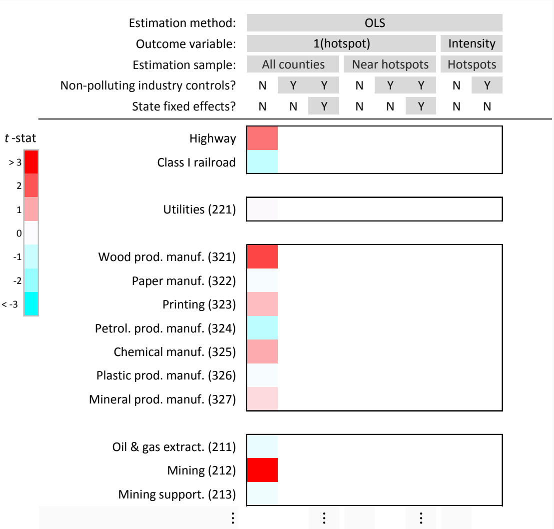
University of Oregon Master's Degree Program in Economics

- 1-year intensive training in applied economics and data science
- Ideal preparation for careers in data science, consulting, professional economics
  - Cutting edge approaches to causal inference and big data tools
  - Open source software for statistical computing
- Also, great for students planning to go on for a Ph.D.
  - 27 tenured/tenure-track faculty members
  - 20 elective courses to choose from, almost every field
- More info: ericzou@uoregon.edu

Extra Slides

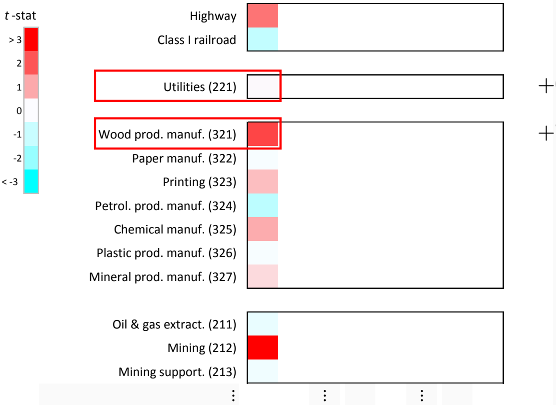
1. Monitoring compliance
2. Frequency selection
3. Neighboring grids
4. NAAQS interaction
5. HHI interaction
6. NAAQS & HHI interaction
7. Continuous $PM_{2.5}$ data
8. 1-in-3 day effects
9. Near-road responses

Industry correlates of pollution gap hot spots



Industry correlates of pollution gap hot spots

Estimation method:	OLS							
Outcome variable:	1(hotspot)						Intensity	
Estimation sample:	All counties			Near hotspots			Hotspots	
Non-polluting industry controls?	N	Y	Y	N	Y	Y	N	Y
State fixed effects?	N	N	Y	N	N	Y	N	N



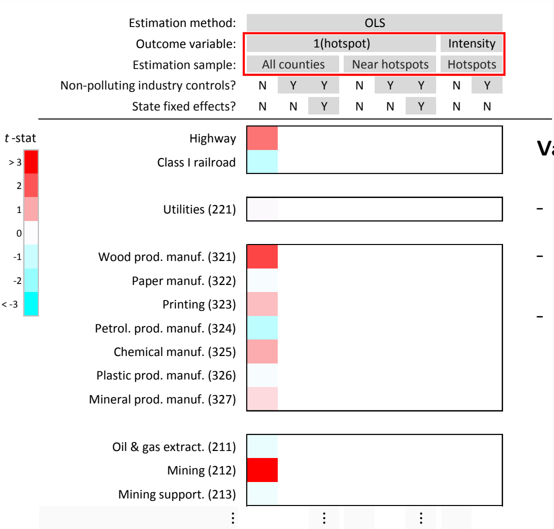
⋮ ⋮ ⋮

$+0.1$ pp (SE $= 1.7$, $\underline{t{=}0.1}$)

$+7.0$ pp (SE $= 1.7$, $\underline{t{=}4.0}$)

⋮ ⋮ ⋮

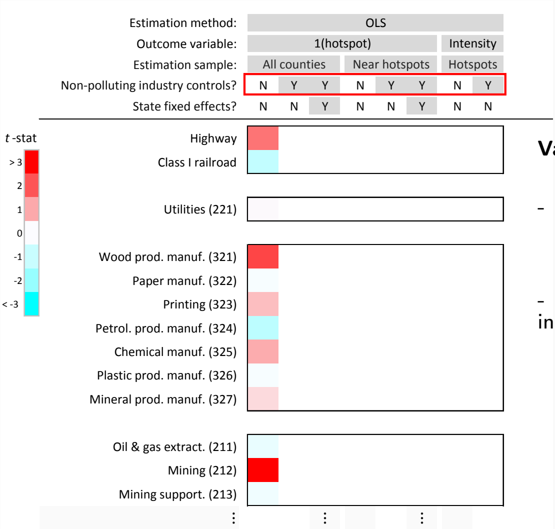
Industry correlates of pollution gap hot spots



Vary sample restrictions:

- All counties ($N \approx$ 3,000)
- Counties $<$ 50 miles from hot spots ($N \approx$ 900)
- Hot-spot counties ($N \approx$ 300; dep. var.$=$ intensity)

Industry correlates of pollution gap hot spots

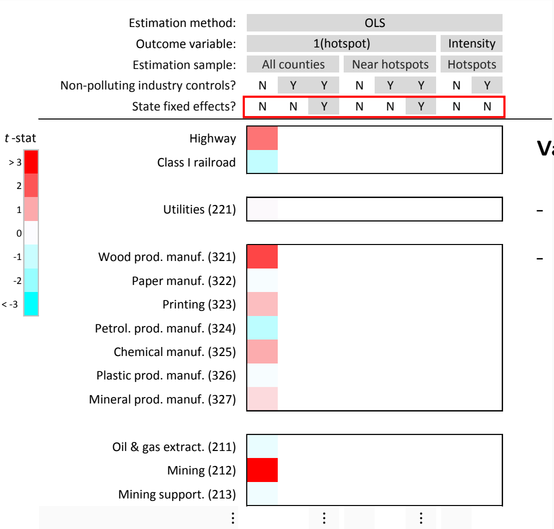


Vary industry controls:

- "Polluting" industries only ($>$ 1% of total PM emission $N_{\text{industry}} = 33$)

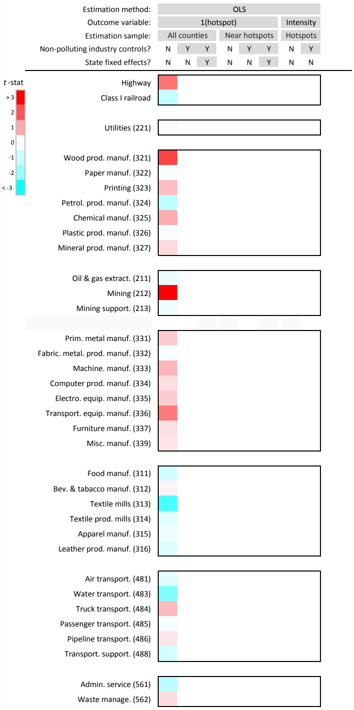
- Further controlling for other industries ($N_{\text{industry}} = 92$)

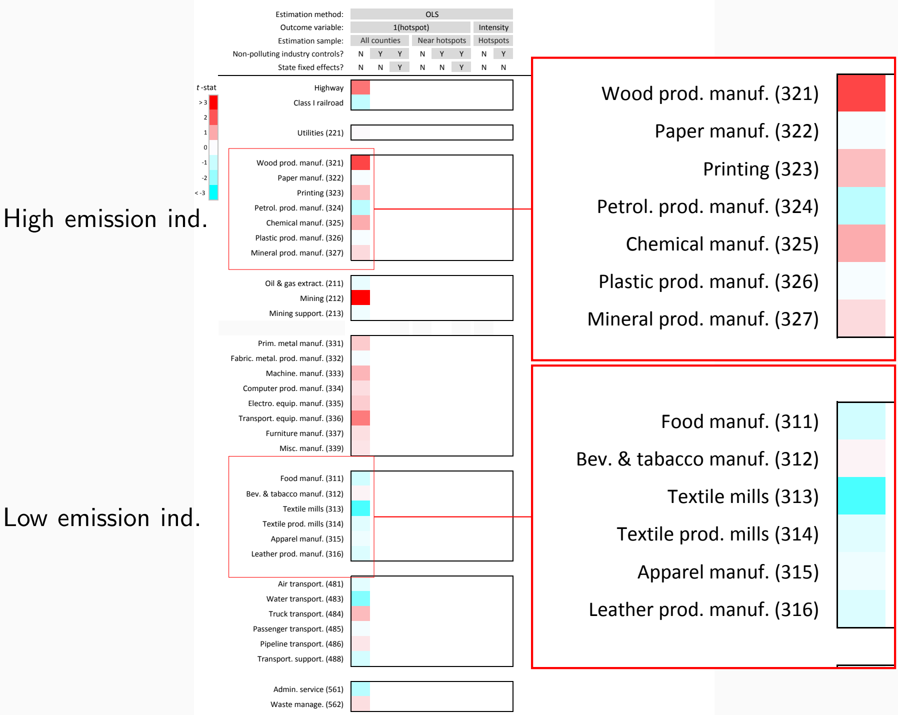
Industry correlates of pollution gap hot spots



Vary sources of variation:

- Full cross section
- State fixed effects

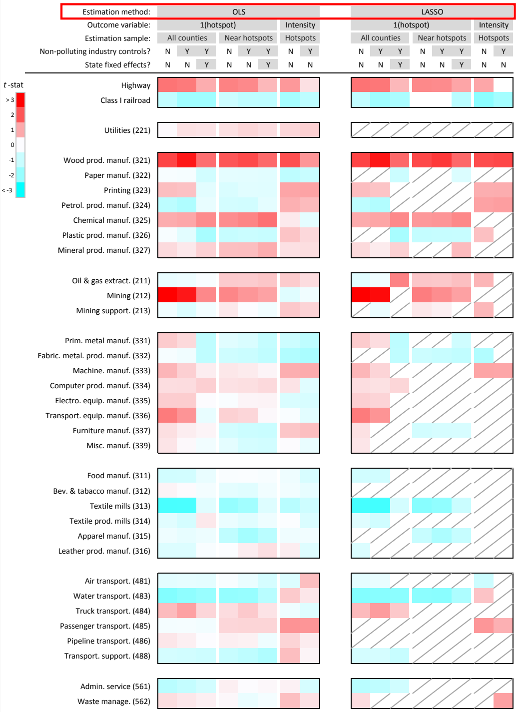

Estimation method:
OLS
Outcome variable:
1(hotspot)
Intensity
Estimation sample:
All counties
Near hotspots
Hotspots
Non-polluting industry controls?
N Y Y N Y Y N Y
State fixed effects?
N N Y N N Y N N
t -stat
> 3
2
1
0
-1
-2
< -3
Highway
Class I railroad
Utilities (221)
Wood prod. manuf. (321)
Paper manuf. (322)
Printing (323)
Petrol. prod. manuf. (324)
Chemical manuf. (325)
Plastic prod. manuf. (326)
Mineral prod. manuf. (327)
Oil & gas extract. (211)
Mining (212)
Mining support. (213)
Prim. metal manuf. (331)
Fabric. metal. prod. manuf. (332)
Machine. manuf. (333)
Computer prod. manuf. (334)
Electro. equip. manuf. (335)
Transport. equip. manuf. (336)
Furniture manuf. (337)
Misc. manuf. (339)
Food manuf. (311)
Bev. & tabacco manuf. (312)
Textile mills (313)
Textile prod. mills (314)
Apparel manuf. (315)
Leather prod. manuf. (316)
Air transport. (481)
Water transport. (483)
Truck transport. (484)
Passenger transport. (485)
Pipeline transport. (486)
Transport. support. (488)
Admin. service (561)
Waste manage. (562)

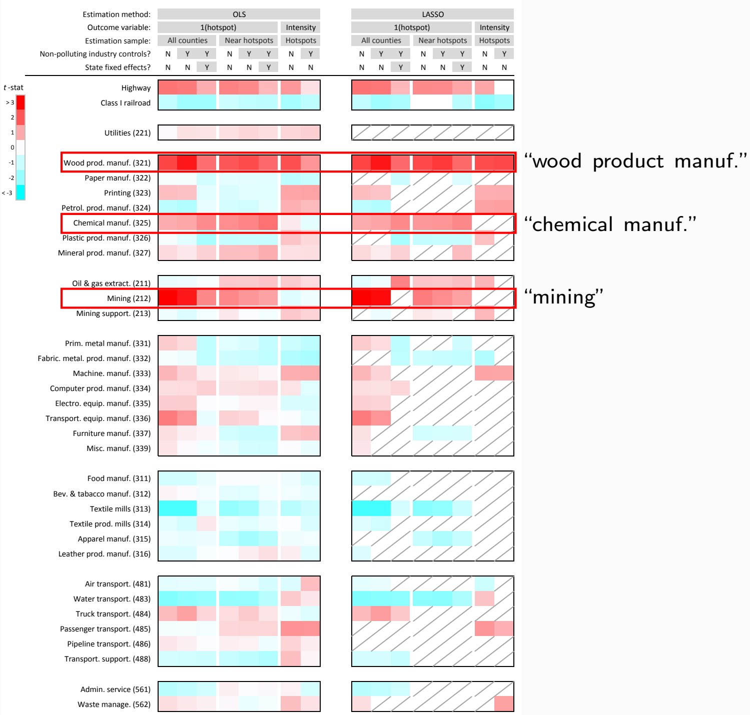


Estimation method: OLS
Outcome variable: 1(hotspot) Intensity
Estimation sample: All counties Near hotspots Hotspots
Non-polluting industry controls? N Y Y N Y Y N Y
State fixed effects? N N Y N N Y N N
t-stat
> 3
2
1
0
-1
-2
< -3
Highway
Class I railroad
Utilities (221)
High emission ind.
Wood prod. manuf. (321)
Paper manuf. (322)
Printing (323)
Petrol. prod. manuf. (324)
Chemical manuf. (325)
Plastic prod. manuf. (326)
Mineral prod. manuf. (327)
Oil & gas extract. (211)
Mining (212)
Mining support. (213)
Prim. metal manuf. (331)
Fabric. metal. prod. manuf. (332)
Machine. manuf. (333)
Computer prod. manuf. (334)
Electro. equip. manuf. (335)
Transport. equip. manuf. (336)
Furniture manuf. (337)
Misc. manuf. (339)
Low emission ind.
Food manuf. (311)
Bev. & tabacco manuf. (312)
Textile mills (313)
Textile prod. mills (314)
Apparel manuf. (315)
Leather prod. manuf. (316)
Air transport. (481)
Water transport. (483)
Truck transport. (484)
Passenger transport. (485)
Pipeline transport. (486)
Transport. support. (488)
Admin. service (561)
Waste manage. (562)


Estimation method:
OLS
Outcome variable:
1(hotspot)
Intensity
Estimation sample:
All counties
Near hotspots
Hotspots
Non-polluting industry controls?
N Y Y N Y Y N Y
State fixed effects?
N N Y N N Y N N
t -stat
> 3
2
1
0
-1
-2
< -3
Highway
Class I railroad
Utilities (221)
Wood prod. manuf. (321)
Paper manuf. (322)
Printing (323)
Petrol. prod. manuf. (324)
Chemical manuf. (325)
Plastic prod. manuf. (326)
Mineral prod. manuf. (327)
Oil & gas extract. (211)
Mining (212)
Mining support. (213)
Prim. metal manuf. (331)
Fabric. metal. prod. manuf. (332)
Machine. manuf. (333)
Computer prod. manuf. (334)
Electro. equip. manuf. (335)
Transport. equip. manuf. (336)
Furniture manuf. (337)
Misc. manuf. (339)
Food manuf. (311)
Bev. & tabacco manuf. (312)
Textile mills (313)
Textile prod. mills (314)
Apparel manuf. (315)
Leather prod. manuf. (316)
Air transport. (481)
Water transport. (483)
Truck transport. (484)
Passenger transport. (485)
Pipeline transport. (486)
Transport. support. (488)
Admin. service (561)
Waste manage. (562)

Sparse models (Lasso)




Estimation method:
OLS
LASSO
Outcome variable:
1(hotspot)
Intensity
Estimation sample:
All counties
Near hotspots
Hotspots
Non-polluting industry controls?
N Y Y N Y Y N Y
State fixed effects?
N N Y N N Y N N
t -stat
> 3
2
1
0
-1
-2
< -3
Highway
Class I railroad
Utilities (221)
Wood prod. manuf. (321)
Paper manuf. (322)
Printing (323)
Petrol. prod. manuf. (324)
Chemical manuf. (325)
Plastic prod. manuf. (326)
Mineral prod. manuf. (327)
Oil & gas extract. (211)
Mining (212)
Mining support. (213)
Prim. metal manuf. (331)
Fabric. metal. prod. manuf. (332)
Machine. manuf. (333)
Computer prod. manuf. (334)
Electro. equip. manuf. (335)
Transport. equip. manuf. (336)
Furniture manuf. (337)
Misc. manuf. (339)
Food manuf. (311)
Bev. & tabacco manuf. (312)
Textile mills (313)
Textile prod. mills (314)
Apparel manuf. (315)
Leather prod. manuf. (316)
Air transport. (481)
Water transport. (483)
Truck transport. (484)
Passenger transport. (485)
Pipeline transport. (486)
Transport. support. (488)
Admin. service (561)
Waste manage. (562)
"wood product manuf."
"chemical manuf."
"mining"

Industry sources of pollution gap: Additional evidence

- Using plant location data, directly estimate pollution gap around plants
- Further estimate pollution gap *gradient* by plant's distance to intermittent PM monitors
  - Wood / chemical plants: Toxic Release Inventory
  - Coal mining sites: MHSA databases
  - For coal mines, suggestive evidence of higher injury rate during off-days

Estimation method:
OLS
LASSO
Outcome variable:
1(hotspot)
Intensity
Estimation sample:
All counties
Near hotspots
Hotspots
Non-polluting industry controls?
N Y Y N Y Y N Y
State fixed effects?
N N Y N N Y N N
t-stat
> 3
2
1
0
-1
-2
< -3
Highway
Class I railroad
Utilities (221)
Wood prod. manuf. (321)
Paper manuf. (322)
Printing (323)
Petrol. prod. manuf. (324)
Chemical manuf. (325)
Plastic prod. manuf. (326)
Mineral prod. manuf. (327)
Oil & gas extract. (211)
Mining (212)
Mining support. (213)
Prim. metal manuf. (331)
Fabric. metal. prod. manuf. (332)
Machine. manuf. (333)
Computer prod. manuf. (334)
Electro. equip. manuf. (335)
Transport. equip. manuf. (336)
Furniture manuf. (337)
Misc. manuf. (339)
Food manuf. (311)
Bev. & tabacco manuf. (312)
Textile mills (313)
Textile prod. mills (314)
Apparel manuf. (315)
Leather prod. manuf. (316)
Air transport. (481)
Water transport. (483)
Truck transport. (484)
Passenger transport. (485)
Pipeline transport. (486)
Transport. support. (488)
Admin. service (561)
Waste manage. (562)
“has highway”

Pollution gap estimates for 1-in-6 day monitors

Dep. var. = Aerosol (log)				
	(1) Sample: sites w. **any** 1in6d monitor	(2) Sample: sites w. **any** 1in6d monitor	(3) Sample: sites w. **only** 1in6d monitor	(4) Sample: counties w. **only** 1in6d monitor
1(*off-days*)	0.016*** (0.004)	0.016*** (0.004)	0.018*** (0.004)	0.018*** (0.006)
Ctrls		✓	✓	✓
N	685,060	685,060	427,846	176,225
N (site)	1,193	1,193	899	489

Notes: Controls include FEs (site, year, month-of-year, day-of-week), daily temperature bins, precipitation, and wind speed bins. SEs clustered at the county lvl. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Pollution gap around "placebo" monitors.

Dep. var. = Aerosol (log)			
	(1) Sample: retired 1in6d PM monitors	(2) Sample: 1in1d PM monitors	(3) Sample: 1in6d toxic. monitors
1(*off-days*)	-0.0020	0.0023	0.0023
	(0.0046)	(0.0080)	(0.0044)
$\text{Power}_{(1.5\% \text{ effect, } 5\% \text{ sig.})}$	0.940	0.803	0.910
N	372,989	231,532	370,020
N (site)	490	556	792

Notes: Controls include FEs (site, year, month-of-year, day-of-week), daily temperature bins, precipitation, and wind speed bins. SEs clustered at the county lvl. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Intuition: Pollution gap shifts

No.

Yes.

Illustration: Pollution gap shifts

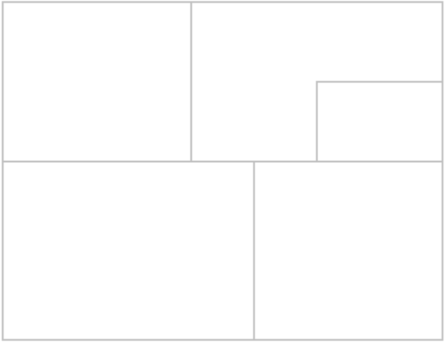


Illustration: Pollution gap shifts

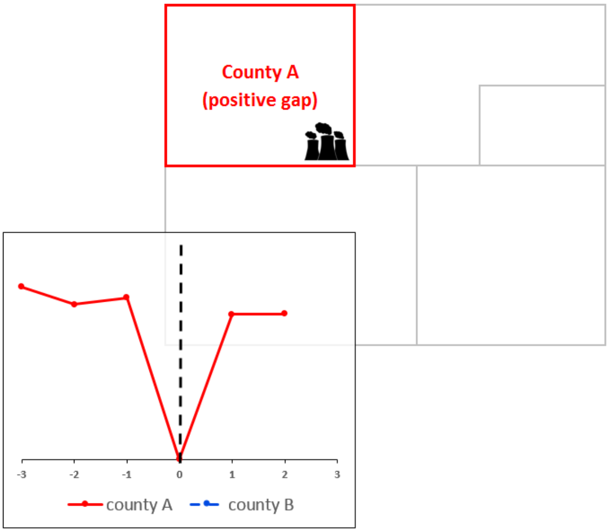
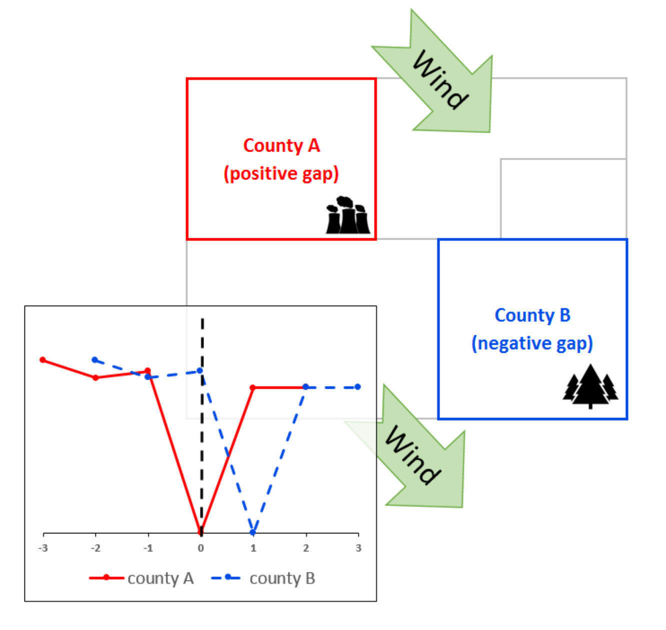
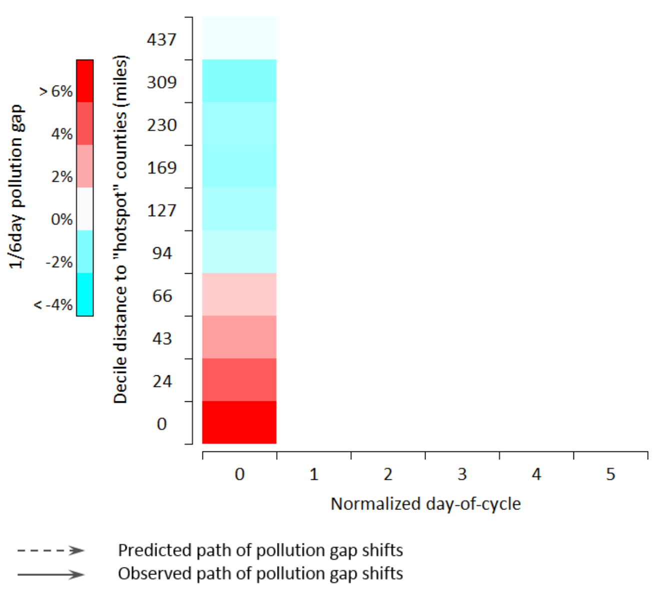


Illustration: Pollution gap shifts

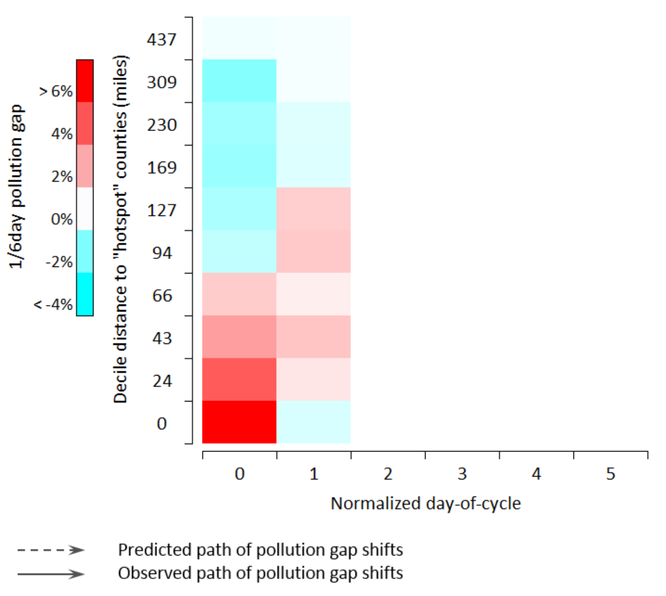


Test: Pollution gap shifts



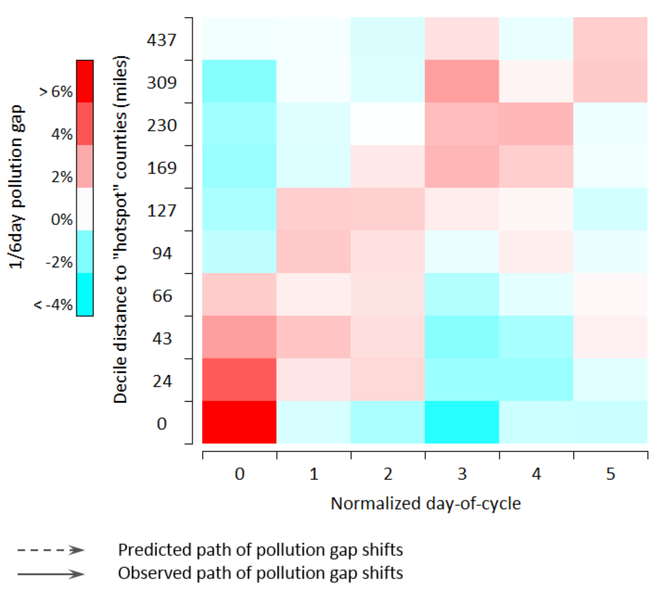
Notes: Figure shows 1/6day pollution paths estimated separately by distance decile groups to the nearest hotspot counties.

Test: Pollution gap shifts



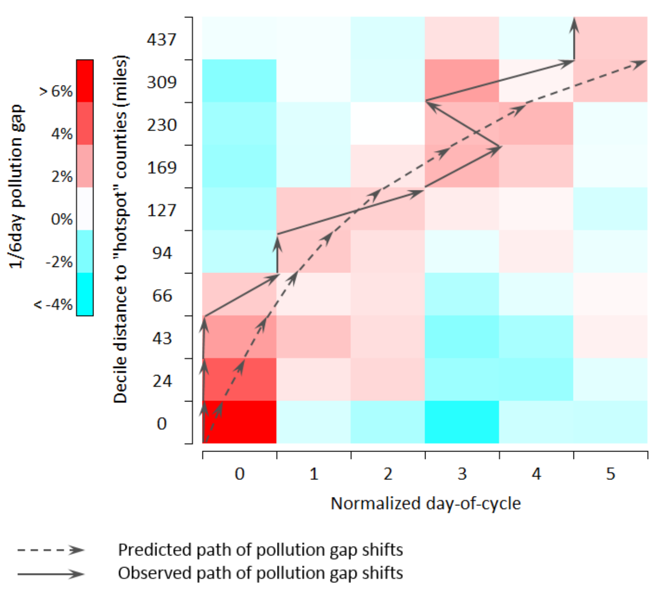
Notes: Figure shows 1/6day pollution paths estimated separately by distance decile groups to the nearest hotspot counties.

Test: Pollution gap shifts



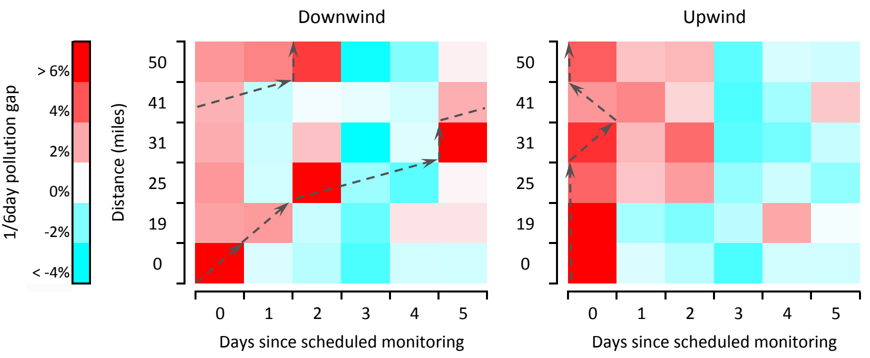
Notes: Figure shows 1/6day pollution paths estimated separately by distance decile groups to the nearest hotspot counties.

Test: Pollution gap shifts



Notes: Figure shows 1/6day pollution paths estimated separately by distance decile groups to the nearest hotspot counties.

Test: Pollution gap shifts by wind direction.



Notes: Each row represents a group of counties by distance to the hotspot counties. Each column represents a day in a 1-in-6 day monitoring cycle, with 0 being the on-day according to the EPA's monitoring schedule. Each cell shows the (log) pollution difference between that day and the other five days on the same row (so, for column day-0, this is just the off-day vs. on-day pollution gap). Regression models are estimated separately for each row. Sample restricts to counties within 50 miles to hotspots, cut by quintile distance, and then further grouped into ones that are downwind and upwind the hotspot counties. A county is downwind if its centroid falls within a 30-degree cone relative to the prevailing wind direction at the nearest hotspot county. Upwind counties are defined symmetrically. Prevailing wind direction is measured by 13 year average daily wind direction from 2001 to 2013 at the hotspot county centroid.

Pollution gap estimates by "Action Day" declaration

Dep. var. = Aerosol (log)		
	(1)	(2)
1(off-days) × 1(warning)	0.069***	0.051***
	(0.014)	(0.013)
1(*off-days*) × 1(no warning)	0.011**	0.013***
	(0.005)	(0.005)
1(*off-days*) × 1(no "Action Day" program)	0.011*	0.016***
	(0.011)	(0.006)
Ctrls.		✓
N	685,060	685,060

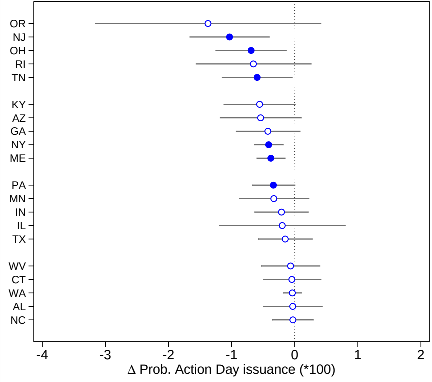
Notes: Each column represents a separate regression. Controls include FEs (site, year, month-of-year, day-of-week), daily temperature bins, precipitation, and wind speed bins. SEs clustered at the county lvl. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

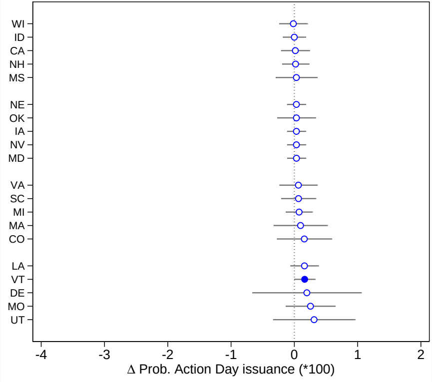
Strategic "Action Day" declaration by non-attainment status

Dep. var. = Issuance of Action Day Advisories (coeff. $\times$ 100)		
	(1)	(2)
$1(\textit{off-days}) \times 1(\text{Attainment})$	-0.087**	-0.086**
	(0.035)	(0.035)
$\mathbf{1(off\text{-}days) \times 1(Non\text{-}attainment)}$	-0.604**	-0.599**
	(0.263)	(0.263)
Equality $p-$value	0.055	0.057
Ctrls.		✓
Mean dep. var. ($\times$ 100)	0.998	0.998
N	624,663	624,663

Notes: Each column represents a separate regression. First issuance is counted in cases of consecutive issuances. Controls include fixed effects (CBSA, year, month-of-year, and day-of-week). Standard errors are clustered at the CBSA level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

State heterogeneity





Notes: Outcome variable is core based statistical area (CBSA) $\times$ daily dummy for weather any action day is issued. Sample spans 2004-2013 and includes 14,945 issuances across 171 CBSAs. In cases of issuances that span a consecutive number of days, only the first day of issuance is counted. Panel B shows off-days vs. on-days issuance probability differential estimated separately for each state.

Mortality gap estimates

Dep. var.: Mortality rate (per million Medicare beneficiaries)				
	(1)	(2)	(3)	(4)
Estimation sample:	**1/6day counties**		**1/1day counties (placebo)**	
Mortality gap	0.903**	0.865**	-0.069	-0.091
	(0.418)	(0.418)	(0.420)	(0.425)
Ctrls.		✓		✓
Dep. var. mean	130.29	130.29	128.10	128.10
N	432,825	432,825	162,660	162,660
N (counties)	321	321	152	152

Notes: "Mortality gap" is the mortality difference between mortality non-decline-day and decline-day in the 6-day monitoring cycle. Controls include FEs (site, year, month-of-year, day-of-week), daily temperature bins, precipitation, and wind speed bins. SEs clustered at the county lvl. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Mortality cost of intermittent monitoring

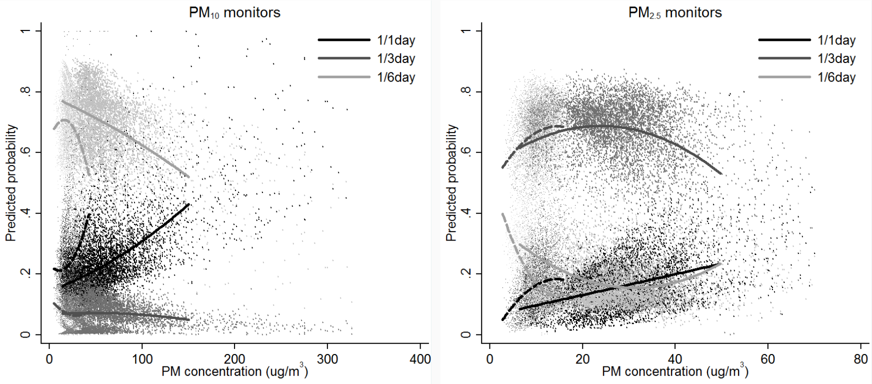
- We calculate loss in life values relative to the counterfactual in which mortality rates do not deviate from the mortality-decline-day
  - Use **lower bound on 95% CI** of effect estimate
  - Use co-morbidity adjusted life years lost of **5 years per death** (Deryugina et al., 2016)
  - Assume the policy only affects the **2 million Medicare beneficiaries** living in 1/6day monitoring counties
  - Assume a conventional **VSL of $100,000 per life year**
- Annual loss in life value ≈ **$20 million/year**
- Cost-savings from intermittent monitoring ≈ **$12 million/year**

Monitoring compliance.

	(1) Samples required	(2) Samples taken	(3) Fraction taking $\geq$90% required samples	(4) Fraction taking 100% required samples
1/6day monitors	60 or 61	58.4 [2.2]	96.74%	19.21%
1/3day monitors	121 or 122	115.6 [4.4]	94.72%	5.42%
1/1day monitors	365 or 366	349.1 [13.0]	92.54%	6.33%

Notes: Statistics are computed from monitor-year observations. Sample includes all monitors eligible for NAAQS comparison. Standard deviation in brackets.

Frequency selection.



Notes: Graph reports predicted probability of monitoring schedule assignment for PM_{10} (left panel) and $PM_{2.5}$ (right panel) by annual PM concentration. Predictions are obtained from a multinomial logistic model that predicts selection into monitoring schedule by annual average and 99^{th} percentile PM value fully interacted with Census region dummies, 5 year lags in annual average as well as 99^{th} percentile value, and calendar year dummies. Each dot on the graph represent a monitor-pollutant metric. Lines show quadratic fits of predicted probability over annual average concentration (dashed) and annual 99^{th} percentile concentration (solid).

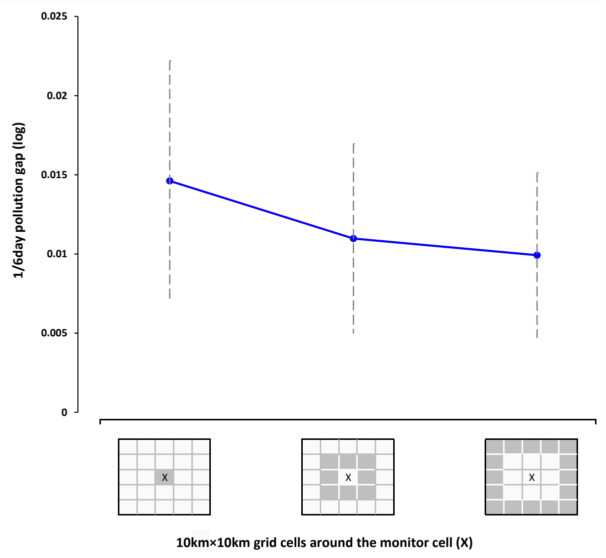
Action day estimates.

1-in-6 day vs. 1-in-1 day heterogeneity

Dep. var. : Issuance of Action Day Advisories (coeff. $\times$ 100)				
	(1)	(2)	(3)	(4)
	Sample: All issuance		Sample: Non-consecutive issuance	
Panel A: CBSAs with 1/6day monitoring				
off-days	-0.169***	-0.167***	-0.121***	-0.120***
	(0.049)	(0.049)	(0.041)	(0.041)
Mean dep. var. ($\times$ 100)	2.57	2.57	1.05	1.05
Ctrls.		✓		✓
N	467,221	467,221	467,221	467,221
Panel B: CBSAs with 1/1day monitoring				
off-days	0.042	0.037	-0.009	-0.010
	(0.131)	(0.133)	(0.082)	(0.082)
Mean dep. var. ($\times$ 100)	1.19	1.19	0.571	0.571
Ctrls.		✓		✓
N	25,938	25,938	25,938	25,938

Notes: Each panel $\times$ column represents a separate regression. "Non-consecutive issuance" is the day of Action Day in cases of consecutive issuances. Controls include fixed effects (CBSA, year, month-of-year, and day-of-week). Standard errors are clustered at the CBSA level. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

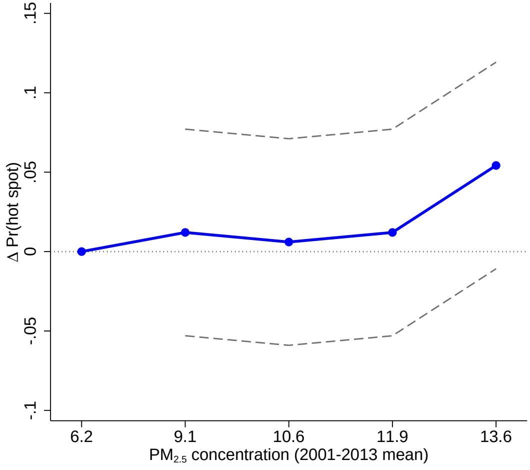
Pollution gap in neighboring grids



Notes: Graph plots 1-in-6 day pollution gap estimates for the 10km × 10km grid that contains the monitor (left), first-order neighboring grids (middle), and second-order neighboring grids (right). Gray dashed bars show 95% confidence intervals constructed using standard errors clustered at the county level.

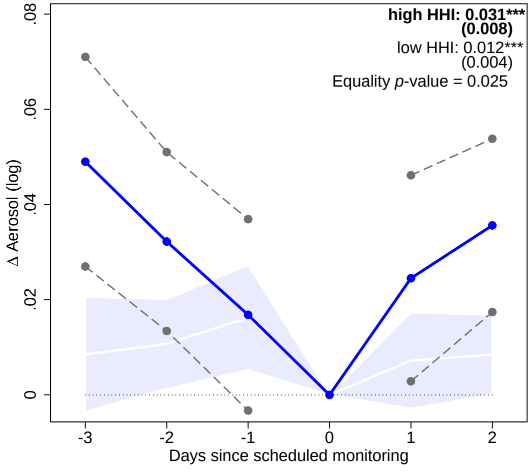
NAAQS interaction.

Cross-section estimation: Hot spots vs. 13-year average $PM_{2.5}$



Notes: Graph plots probability of a county being a pollution gap hot spot by quintiles of 2001-2013 average $PM_{2.5}$ concentrations. The regression restricts to counties that ever had $PM_{2.5}$ monitors from 2001-2013. x-axis indicates mean $PM_{2.5}$ within each concentration quintile. Coefficient for the lowest concentration bin is normalized to zero.

HHI interaction: panel estimation.



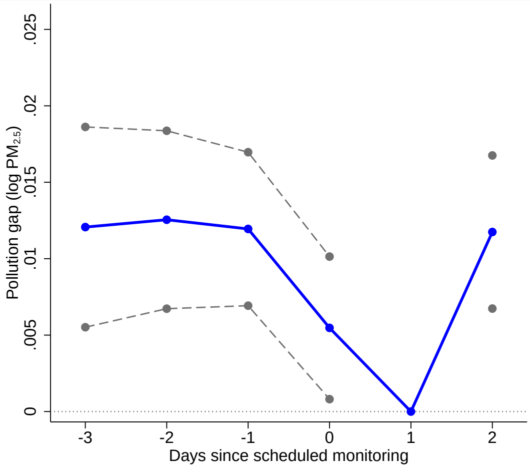
Notes: Figure displays 1-in-6 day pollution pattern separately for high Herfindahl index (≥ 0.9) vs. low Herfindahl index (< 0.9) counties. Estimates are obtained from a single regression. Foreground graph objects represent estimates for the high Herfindahl index counties while the background graph objects show estimates for the rest of the samples. Dashed lines and the shades represent 95% confidence interval constructed from standard errors clustered at the county level. Point estimates shown on the upper-right corner shows average pollution gap. Equality *p*-value corresponds to the null hypothesis that there is no difference in the off-days effect for the two groups. The regression includes fixed effects dummies (site, year, month-of-year, and day-of-week) and weather controls.

Example: Regulation & industrial structure correlates

Dep. var. = 1(hot-spot counties). Mean = 0.10		
	(1)	(2)
1(Non-attainment)	0.042**	0.046**
	(0.019)	(0.019)
1(Has 1-in-6 day monitors)	0.029*	0.001
	(0.015)	(0.017)
1(Emission Herfindahl $\geq$ 0.9)	0.007	-0.007
	(0.011)	(0.012)
1(Emission Herfindahl $\geq$ 0.9)		0.109***
$\times$ 1(Has 1-in-6 day monitors)		(0.033)
N	3,199	3,199

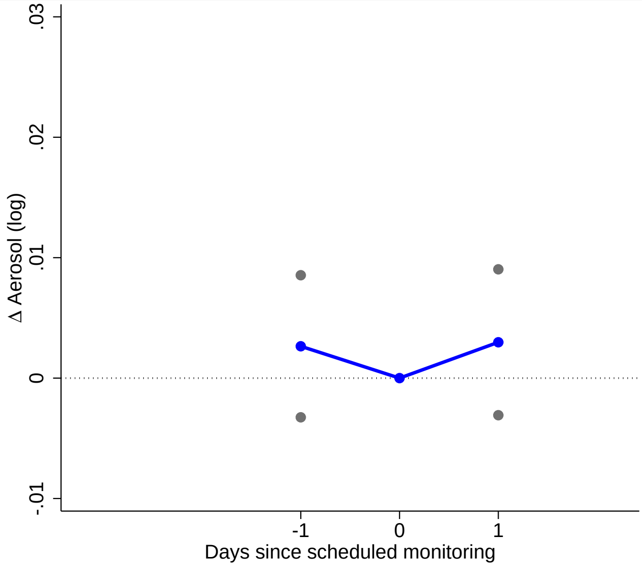
Notes: Emission Herfindahl = county-level avg TRI emission Herfindahl Index 2001-2013. High index indicates total emissions concentrated in the hands of few polluters. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Continuous $PM_{2.5}$ data.



Notes: Graph plots 1-in-6 day pollution gap as detected by the EPA's continuous $PM_{2.5}$ monitors. These monitors use indirect methods (e.g. beta-ray attenuation and microbalance) to infer $PM_{2.5}$ concentration, and are mainly used toward public air quality disclosure and forecast purposes, rather than comparison to NAAQS. See Appendix A for more details. Estimation sample restrict to monitor$\times$months with at least 28 daily $PM_{2.5}$ observations available. The day next to the monitoring day is normalized to 0. Regression includes fixed effects dummies (site, year, month-of-year, and day-of-week) and weather controls. Dashed lines represent 95% confidence intervals constructed using standard errors clustered at the county level.

Pollution gap at 1-in-3 day sites.



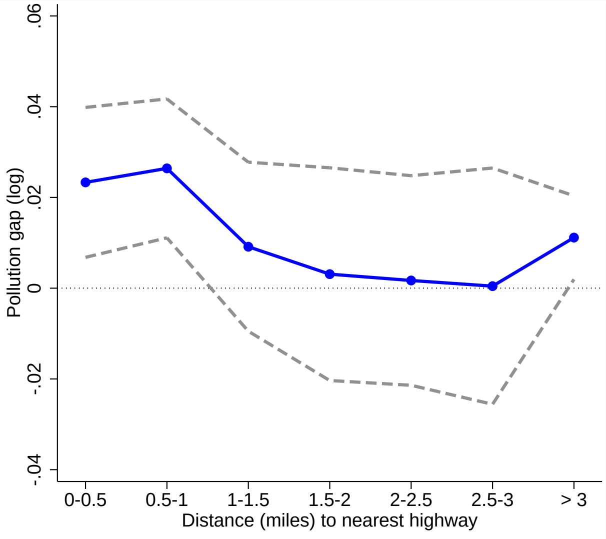
Notes: Sample spans 2001-2013. Dep var = aerosol optical depth within 10km grid cell containing a 1/3day monitoring site. Results similar with or without controls. Dashed lines show 95% CI using SEs clustered at the county level.

Pollution gap at 1-in-3 day sites..

Dep. var. = Aerosol concentration (log)				
	(1) Sample: sites w. any 1/3d monitor	(2) Sample: sites w. any 1/3d monitor	(3) Sample: sites w. only 1/3d monitor	(4) Sample: counties w. only 1/3d monitor
1(off-days)	0.0028	0.0029	0.0024	0.0054*
	(0.0026)	(0.0020)	(0.0025)	(0.0030)
Ctrls		✓	✓	✓
N	598,859	598,859	386,854	244,071
N (site)	1,064	1,064	849	562

Notes: Controls include FEs (site, year, month-of-year, day-of-week), daily temperature bins, precipitation, and wind speed bins. SEs clustered at the county lvl. *: $p < 0.10$; **: $p < 0.05$; ***: $p < 0.01$.

Near road pollution gap



Notes: Figure plots interaction of pollution gap with the 1-in-6 day PM monitor's distance (bins) to the nearest highway. The group ">3" pools all monitors that fall more than 3 miles from the nearest highway. Dashed lines represent 95% confidence intervals constructed using standard errors clustered at the county level.