Quick Primer on Doing Empirical Research

Reed Walker
UC Berkeley

Spring 2022

Doing Research: 2nd year and beyond

Graduate school is well-structured to teach you:

1. **Economics**, i.e. What are the interesting and important questions?
2. **Technical skills**, i.e. How to answer them

How to come up with a question / find data?

How to get a job?

- Job market candidates: backwards induction...

Step 1: Coming up with Ideas

All (good) research starts with a question that is interesting

- Can you explain to others why it's interesting and exciting?
- Your peers and professors
- Non economists
- Your family
- Non-economist friends (if any)

Are you interested in this question?

- If you are not interested and excited in your project, how can you possibly expect anyone else to be?
- And you will certainly not enjoying working on it for many years!

Some Ways to Come up with Ideas

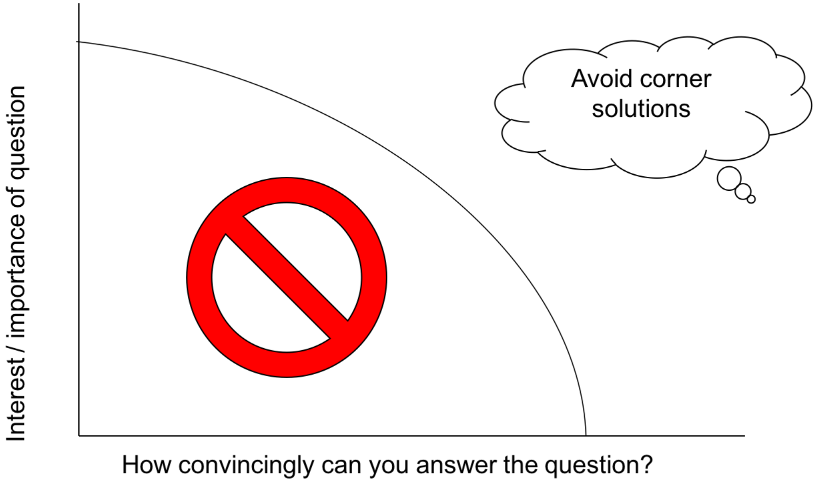
Sources of ideas:

- Classes – what are the important unanswered questions?
- Seminars – what does the seminar make me think about?
- In general do not go to the literature for ideas
  - Broad survey articles can stimulate ideas
  - JEL, JEP, Handbook Chapters
- Read the newspaper with an eye towards economic questions
  - Look at the real world, not just the economics literature
- Read non-economics non-fiction (e.g. Biography, history)
- Talk to people – economists and non economists

Ideas come at random times

- Be sure to write them down whenever you have them
- Keep at it

Always Be on the Frontier



Fast Forward to the End

If you think you have a good idea:

- Imagine you came up with a way to answer your question convincingly
- I know it's hard but "fast forward" to the end where you've produced a really convincing answer

Now ask yourself:

- So what? Why is this interesting?
- What would make it more interesting?
- May help you modify / fine tune your question...

Bottom Line Revisited

Research is not a solo process

- Form weekly working groups with your friends to talk about your latest thoughts
- Force everyone to talk about at least one idea, no matter how lame they think it is
- Think / talk / discuss your ideas or project constantly
- As project progresses, will get different kinds of feedback

Keep a file on your thoughts and people's comments

- You'd be amazed how quickly you can forget...

Proceed in a systematic fashion:

- If X is essential for the project's success, look into X now, not later!

Paper/Idea Evaluation Taxonomy

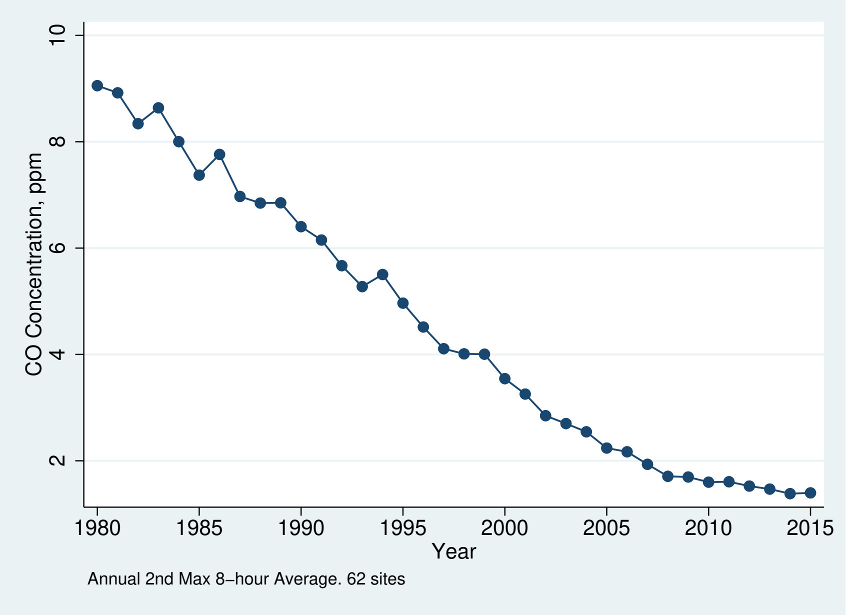
1. What is the question, and why is it interesting?
2. Why is the existing literature crappy, non-existent, and/or unresolved?
3. What is this paper going to do to solve it?

Setting the Stage: An Overview of Where Environmental Policy is Today

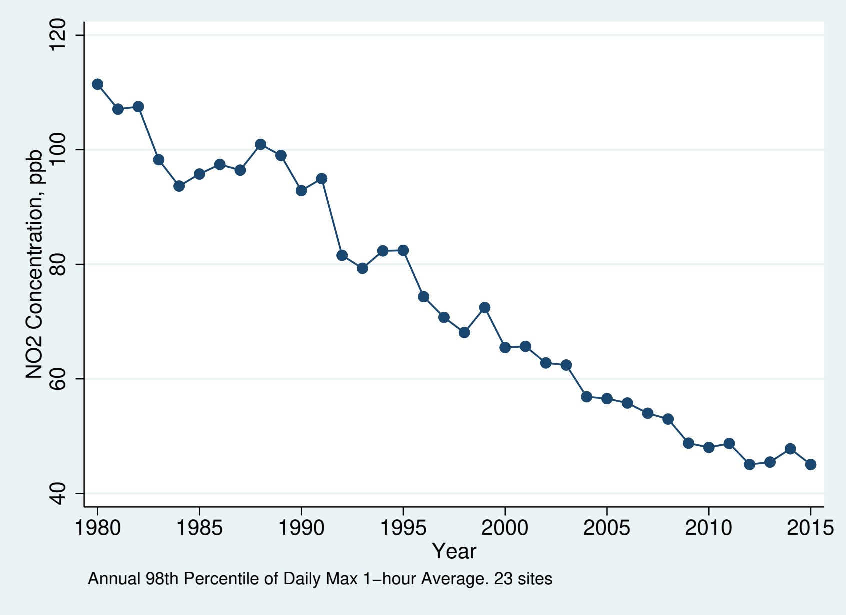
Reed Walker
UC Berkeley

Spring 2022

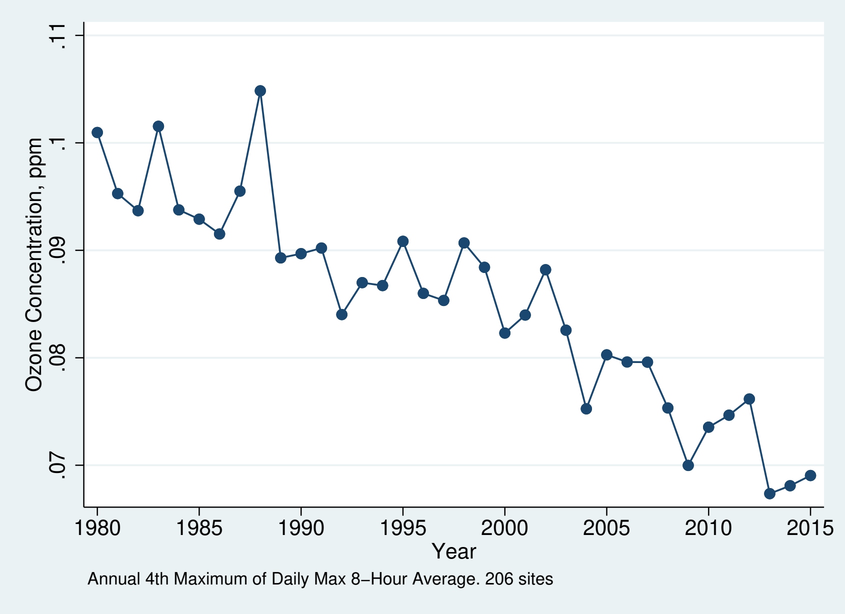
Carbon Monoxide Levels Decreased by 85% Since 1980



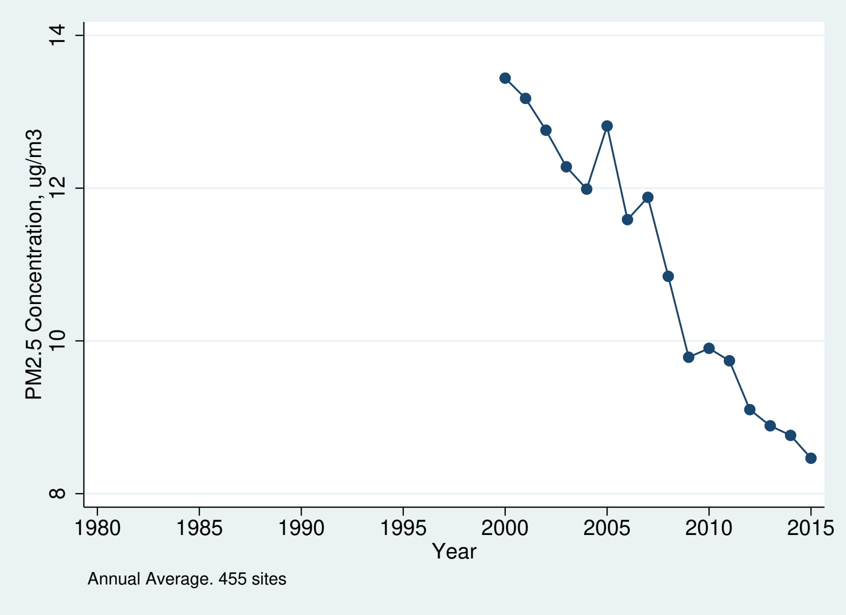
Nitrogen Dioxide Levels Decreased by 61% Since 1980



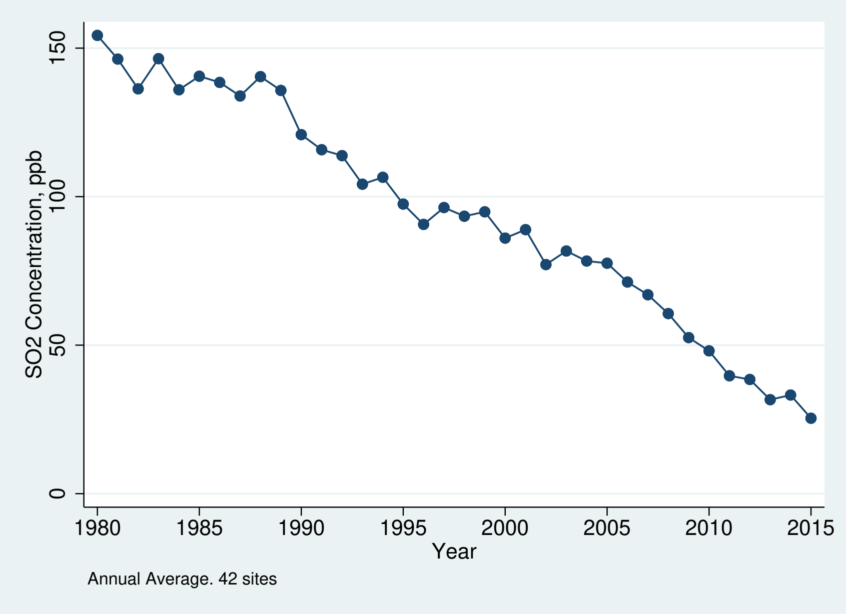
Ozone Levels Decreased by 31% Since 1980

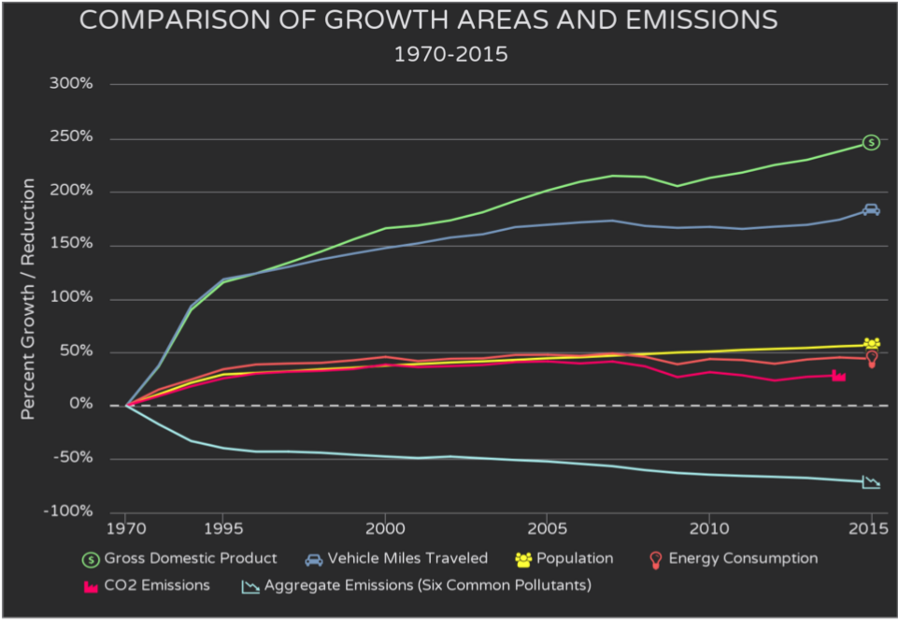


Particulate Matter Levels Decreased by 42% Since 2000



Sulfur Dioxide Levels Decreased by 87% Since 1980




COMPARISON OF GROWTH AREAS AND EMISSIONS
1970-2015
Percent Growth / Reduction
300%
250%
200%
150%
100%
50%
0%
-50%
-100%
1970
1995
2000
2005
2010
2015
Gross Domestic Product
Vehicle Miles Traveled
Population
Energy Consumption
CO2 Emissions
Aggregate Emissions (Six Common Pollutants)

Clean Air Act

Impressive improvements in air quality, but not without costs!

- What do we know about the costs and benefits of the CAA?

Answer: There is a lot we do not know

Insights/tools from other fields that might be useful going forward:

- International Trade - General equilibrium (increasingly parsimonious)
- Industrial Organization - Ground up institutional details
- Public Finance - Welfare analysis w/ emphasis on credible empirics

Outline

1. Very Brief Primer on the Clean Air Act
2. What do we know about the costs and benefits?
3. What don't we know (that we would like to know!)?

Outline

1. **Very Brief Primer on the Clean Air Act**
2. What do we know about the costs and benefits?
3. What don't we know (that we would like to know!)?

Clean Air Act (CAA) Primer

The largest environmental program in the United States

- First enacted in 1963 with major revisions in 1970, 1977, and 1990.
- National Ambient Air Quality Standards (NAAQS) for certain criteria air pollutants (CO, O_3, NO_2, SO_2, PM)
- Administered at the County$\times$Pollutant$\times$Year level
- Areas that exceed the EPA pollution threshold in a given year are designated as "Nonattainment"

CAA Nonattainment Designation: Regulating Polluters

Polluting firms in nonattainment areas are subject to stringent emissions regulations

- Limits on how much pollution may be allowed
- Mandated abatement technology
- Require new sources of pollution to offset their emissions by reducing it from somewhere else within the county.

Environmental regulations are costly to firms:

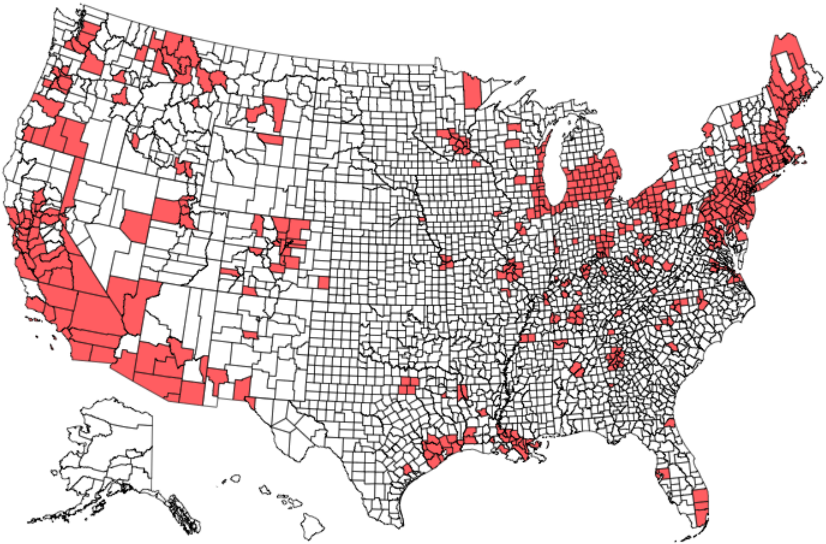
- $20 billion in annual, aggregate operating costs (U.S. Census 2008)
- $6 billion in capital expenditures for pollution abatement in 2005 (U.S. Census 2008).

Clean Air Act as the Basis for a Research Design

Polluting firms in **nonattainment counties** are regulated

1. Only some counties are in nonattainment in a given year
   - Nationwide or industry wide controls

Nonattainment Counties in 1991



NOTE: Figure shows county nonattainment designation for any pollutant. Source EPA.

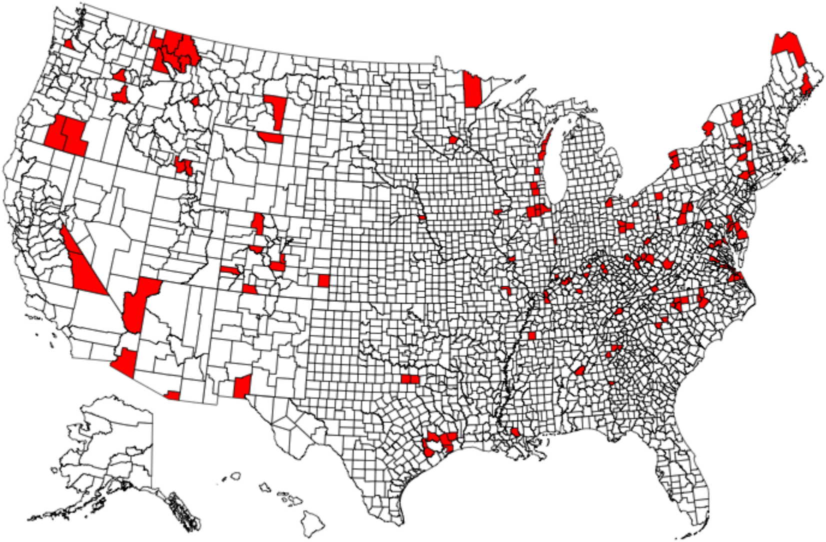
Clean Air Act as the Basis for a Research Design

Polluting firms in **nonattainment counties** are regulated

1. Only some counties are in nonattainment in a given year
   - Nationwide or industry wide controls
2. Counties enter/exit nonattainment based on ambient pollution levels
   - Pre-post comparisons within counties/firms/workers

Clean Air Act Amendments of 1990: New Nonattainment Counties

Clean Air Act Amendments created new and stronger pollution standards



Clean Air Act as the Basis for a Research Design

Polluting firms in **nonattainment counties** are regulated

1. Only some counties are in nonattainment in a given year
   - Nationwide or industry wide controls
2. Counties enter/exit nonattainment each year based on pollution levels
   - Pre-post comparisons within counties/firms/workers
3. Only polluting firms are regulated in a county
   - Facilitates within county comparisons of regulated/unregulated firms
   - Control for unobserved local economic conditions

Outline

1. Very Brief Primer on the Clean Air Act
2. **What do we know about the costs and benefits?**
3. What don't we know (that we would like to know!)?

Big Picture: Social Cost of Air Pollution Externalities

In presence of externality, optimal policy typically entail a tax on emissions equal to marginal damages.

- Pigouvian tax set equal to marginal damages (at the optimal quantity) moves market back to efficient level

The key input in designing optimal environmental policy is the marginal damage associated with a given externality.

- Research community / policy makers have only a cursory understanding as to the social costs of air pollution externalities

There are **five main** challenges that researchers face in this arena

Five Main Challenges: Social Costs of Pollution

1. Causal Inference
   - correlation $\neq$ causation
2. Pollutants / Emissions are Correlated
   - Instruments for pollution often affect multiple pollutants
3. Short-Run vs. Long-Run Exposure
   - Short-run more "exogenous"; long-run more policy relevant
4. Adaptation / Heterogeneity in Dose-Response Functions
   - Avoidance behavior / defensive investments costly $\Rightarrow$ also affect dose-response estimates
5. Monetization
   - Hospitalization costs $\neq$ welfare

Deryugina et al. (2016)

Research Question:

- Effect of acute air pollution exposure on mortality, life-years lost, and health care utilization among the US elderly?

"Frontier" Paper: Addresses many challenges listed above

- Also ignores others; no paper is perfect.

Key Innovations:

- Full Medicare Claims Data - Comprehensive health outcomes/expenditures for 65+
- Predict remaining life expectancy for each individual (LASSO) and use as LHS... reduces life years lost calculations by 50+%
- Daily variation in wind direction to instrument pollution

Deryugina, et al. (2016).

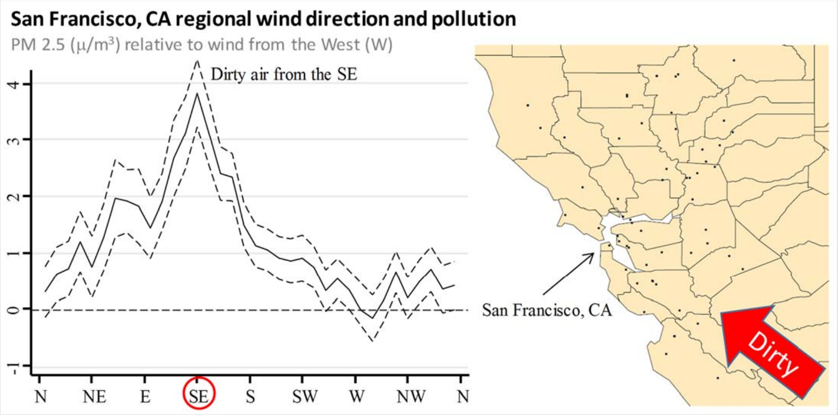


Figure 2. Relationship between daily average wind direction and PM 2.5 concentrations for counties in and around the Bay Area, CA. The left panel shows regression estimates of equation (A1) from the appendix, where the dependent variable is the county average daily PM 2.5 concentration and the key independent variables are a set of indicators for the daily wind direction falling into a particular 10-degree angle bin. Controls include county, month-by-year, and state-by-month fixed effects, as well as a flexible function of maximum and minimum temperatures, precipitation, wind speed, and the interactions between them. The dashed lines represent 95 percent confidence intervals based on robust standard errors. The right panel shows the location of the PM 2.5 pollution monitors (black dots) in the Bay Area that provided the pollution measures for this regression.

Deryugina, et al. (2016).

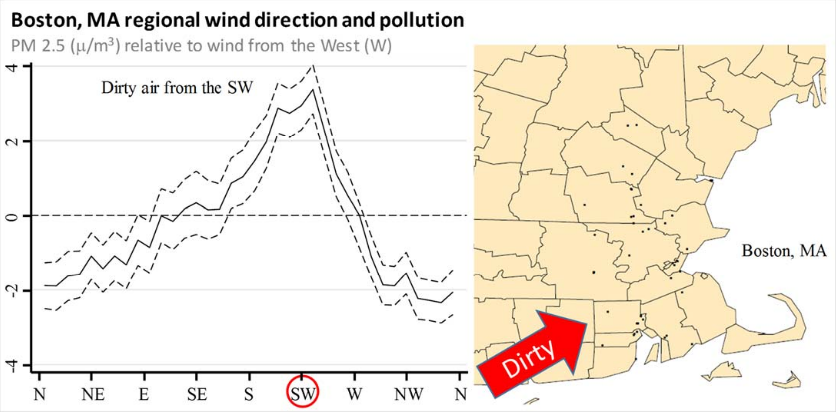


Figure 3. Relationship between daily average wind direction and PM 2.5 concentrations for counties in and around the Boston Area, MA. The left panel shows regression estimates of equation (A1) from the appendix, where the dependent variable is the county average daily PM 2.5 concentration and the key independent variables are a set of indicators for the daily wind direction falling into a particular 10-degree angle bin. Controls include county, month-by-year, and state-by-month fixed effects, as well as a flexible function of maximum and minimum temperatures, precipitation, wind speed, and the interactions between them. The dashed lines represent 95 percent confidence intervals based on robust standard errors. The right panel shows the location of the PM 2.5 pollution monitors (black dots) in the Boston Area that provided the pollution measures for this regression.

Wind direction as an IV for daily variation in pollution

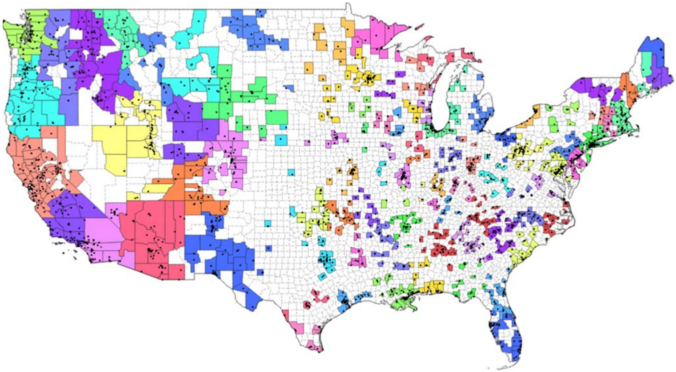


Figure 4. Counties assigned to each monitor group. Different colors correspond to different monitor groups. White corresponds to counties not assigned to any monitor group due to lack of monitors. Black dots represent PM 2.5 pollution monitors.

Implications

Between 1999-2011, national reduction in PM 2.5 levels of 3.65 $\mu g/m^3$

- Estimates imply reduction in elderly deaths of 55,000 per year and reduction in life-years lost by 150,000 per year
- Assuming value of $100,000 per statistical life-year implies a corresponding benefit of $15 billion per year
- Estimating life-years lost using average life expectancy increases estimate to $47 billion

How Does This Paper Stack Up?

1. **Plausibly causal**: ✓
2. **Multiple pollutants** ✓
3. **Representative**: ✓ for age 65+
4. **Non-market valuations**: ✓ VSL
5. **Flexible / Non-linearity**: not really - perhaps underpowered
6. **Reflect adaptation and it's costs**: no
7. **Long-run dose-response**: paper = short-run

Other Costs of Air Pollution Not Captured by Health Based Dose-Response Estimates

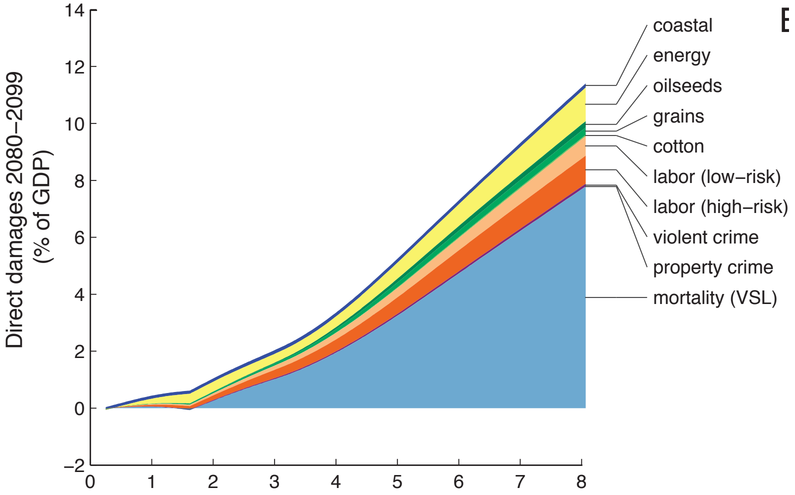
Some examples (there are others):

- Amenity values associated with cleaner air (Chay & Greenstone 2003)
- Cognition and/or labor productivity effects (Neidell & Graff-Ziven)
- School and workplace absenteeism (Currie et al.),
- Defensive investments (Deschenes, Greenstone, & Shapiro).

In principal, each of these components should factor into damage function

- Each subject to same challenges associated with estimating social costs: (1) causal inference, (2) multiple pollutants, (3) short-run vs. long-run, (4) non-linear damage functions, and (5) monetization

A Roadmap: Insights from SCC / CO2 Damage Functions



Source: Hsiang et al. American Climate Prospectus (2014)

The Clean Air Act: Benefits

Damage Function $\neq$ Clean Air Act Benefits

- But it is *the* key input

Need to Know:

1. Causal effect of CAA on air quality [doable]
2. Use damage function to integrate out benefits

Back of the envelope: Deryugina, et al. (2016)

- $15B annually for 65+ ignoring non-health benefits
- Ignoring benefits from other criteria pollutant reductions + non-health benefits

What Do We Know Empirically About CAA Costs?

Taking Stock of Possible Regulatory Effects

1. The average price level / structure of prices / quantities
   - General equilibrium: costs in one market may spill over to other markets
2. The static costs of production:
   1. input distortions
   2. direct regulatory costs
   3. changes in input prices
3. Dynamic efficiency: rate/direction of innovation, productivity, barriers to entry
4. Product quality and variety.
5. Transition Costs: capital and/or labor
6. Distributional concerns
   - less a concern for overall welfare but important

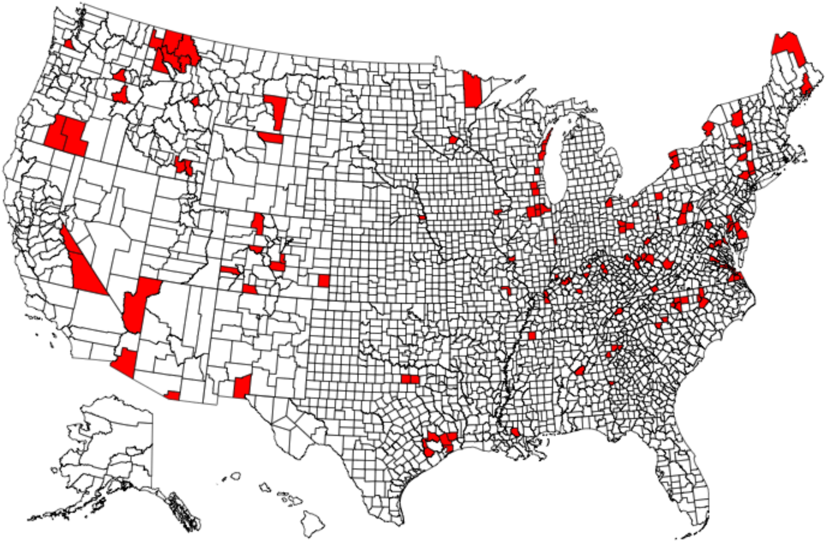
What Do We Know Empirically About CAA Costs?

How to Estimate?: Three Main Approaches

1. Program Evaluation Methods (i.e. ex post analysis)
2. CGE Models + Counterfactuals
3. "Ground-up" IO Models of a Single Sector + Counterfactuals

How Well Do Different Methods Capture Entirety of Compliance Costs?

Clean Air Act Amendments created new and stronger pollution standards



Clean Air Act: Lessons from the Past

NAAQS Nonattainment Status affects...

- Plant entry / exit (Henderson, List et al.,)
- Factor inputs (Greenstone 2002)
- Productivity (Greenstone, List, and Syverson)
- Factor inputs $\Rightarrow$ reallocation costs (Walker 2012)
- Foreign Direct Investment (Hanna 2010)

Program Evaluation Methods / Ex-Post Analysis

Challenges for this literature:

- Largely ignores economy-wide equilibrium implications
- Largely ignoring effects that are more directly connected to welfare (i.e. prices, marginal costs, fixed costs)
- Nonattainment is a blunt instrument for a time-varying and heterogenous treatment

Low hanging fruit:

- Prices, markups, MC (e.g. Ganapati, Shapiro, and Walker)
- Effects on product quality / variety (see trade literature)
- Can we use reduced form estimates to say something about welfare/incidence?
  - See e.g. recent methods in trade / PF

Computable General Equilibrium Models

Used extensively by the US Environmental Protection Agency and other organizations

Jorgenson and Wilcoxen (1990): IGEM Model

- Features over 2000 equations that jointly fit together to define an equilibrium in each period.
- Very black-box with thousands of parameters

Headline Findings: Mandated abatement costs reduce real GNP growth rate by 0.2 percentage points per year (between 1974-1985)

- Mostly through an increase in the cost of capital

Economy-Wide Quantitative Models

Recent innovations in quantitative modeling: International trade / economic geography

- Melitz (2003), Dekle, Eaton and Kortum (2008), Caliendo and Parro (2015), Redding (2016)
- Somewhat more parsimonious
- Not without strong assumptions

Shapiro and Walker (2018)

- Borrow insights from this literature
- Develop quantitative model / counterfactuals to understand role of environmental regulation (accounting for equilibrium feedback, etc...)
- No cost estimate - tractability comes from "relative" statements

Quantitative Models

More to be done here

- Overall net benefits of CAA accounting for externalities, regulatory costs, reallocation, etc... ?
- Input distortions / misallocation / etc...
  - See e.g. Fajgelbaum, Morales, Suarez-Serrato, Zidar (2016)

“Ground-up” IO Models of a single sector

“New Empirical Industrial Organization” (Bresnahan 1989)

- Motivation: individual industries sufficiently distinct (+ details sufficiently important), that cross-industry variation isn’t so useful
- Focus on a single industry + attention to institutional details, measurement of key variables, and econometric identification issues.
- Hope to learn generalizable insights from narrow focus

Some Drawbacks:

- Not clear how results generalize to other industries/markets
- Partial equilibrium
- Models often rely on equilibrium assumptions / non-transparent empirics
    - Can obscure link between estimates and underlying data/variation

Ryan (2012): "The Costs of Environmental Regulation in a Concentrated Industry"

- CAA could affect firm's entry + investment decisions (which might have welfare effects).
  - e.g. increased sunk costs, fixed costs, and/or investment costs
- Estimates dynamic game of entry/exit + investment in US cement industry.
- Estimated model used to evaluate welfare effects of 1990 CAAA

Headline Findings

- Amendments doubled sunk entry costs $\Rightarrow$ reduced net entry and # of plants $\Rightarrow$ increasing market power.
- Amendments led to higher investment by incumbents, but lower aggregate market capacity.
- Consumer welfare decreased 25% due to lower entry and increased market power (approx. $1.2B).

Taking Stock of CAA Regulatory Effects

1. The average price level and the structure of prices **[nope]**
2. The static costs of production:
    1. input distortions **[nope]**
    2. direct regulatory costs **[nope]**
    3. changes in input prices **[nope]**
3. Dynamic efficiency: rate/direction of innovation/productivity
    - Some work: Popp, Greenstone, List, and Syverson
4. Product quality and variety **[nope]**
5. Transition Costs: capital and/or labor
    - Labor transition costs but nothing on capital

Taking Stock of CAA Regulatory Effects

Distributional concerns: less a concern for overall welfare but important

1. profitability of regulated firms **[not really]**
2. rent-sharing with factors of production **[nope]**
3. other relative surplus measures (e.g. producers vs. consumers) **[nope]**

Benefits versus Costs

General sentiment that benefits $>$ costs

- But, when do diminishing returns to abatement start to bind?
- $> 80\%$ reductions in ambient air pollution is no joke!

Marginal Abatement Cost Curve

- What does the short-run abatement cost curve look like?
- What does the long-run abatement cost curve look like?
  - Is the long-run abatement cost curve even upward sloping?
  - Learning by doing / endogenous innovation / etc...

Wrapping Up:

Costs and Benefits of Environmental Policy: Many key questions remain without answers:

- Marginal Damages: need more coordinated efforts to synthesize research into a usable damage function (see e.g. SCC)
- Regulatory Costs: Limited understanding of potential costs to date

New data providing unique opportunities to answer questions that have been central to environmental economics for decades

- Also provide a means to rethink the way we design and implement policy going forward

Non-Market Valuation: Environmental Externalities, Part I

Reed Walker
UC Berkeley

Spring 2022

Externality: Review

Quick Review: Externality correcting taxes (i.e. Pigou)

An externality arises whenever the utility or production possibility of an agent depends **directly** on the actions of another agent.

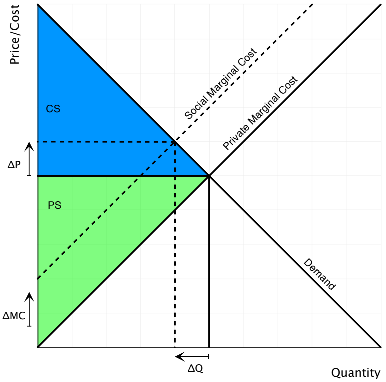
Externalities: What is the best way to correct externalities and move closer to the social optimum?

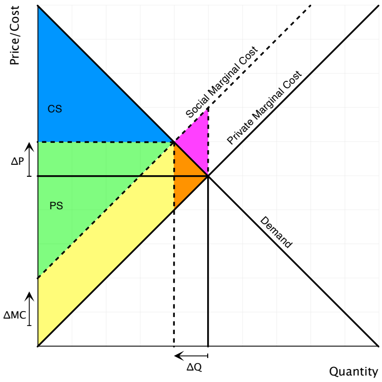
Pigouvian Taxation

Punchline: Pigouvian tax set equal to marginal damages **at the optimal quantity** moves market back to efficient level

- Impose tax $t = MD(Q^*)$
- Restores Pareto efficiency and maximizes social welfare
- Practical limitations:
  - Must know marginal damage function to set t
  - Difficult to measure the marginal damage in practice

Fix Ideas: Welfare Analysis

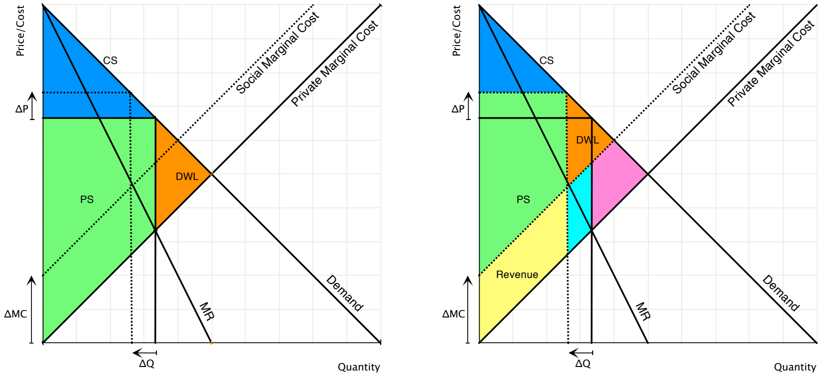




- **Welfare, Net Social Benefit**: Total Change in Welfare Due to Tax

$$\Delta W = dCS + dPS + dG + dE$$

Fix Ideas: Welfare Analysis Differs with Market Structure



- **Welfare, Net Social Benefit**: Total Change in Welfare Due to Tax

$$\Delta W = dCS + dPS + dG + dE$$

Empirics: Welfare, Incidence, and Market Structure

In general, to estimate Δwelfare/incidence, need to know:

- supply curve / marginal costs
- demand curve / preferences
- market structure
- marginal damage function / externality

How to make empirical progress?

- That's one goal of this class
- With enough assumptions – easy
- Hopefully veracity of results not strongly dependent upon assumption
  - Level/shape of damage function
  - Preference heterogeneity
  - etc...

Valuing Environmental Externalities: Direct Measurement - Air Pollution

Reed Walker
UC Berkeley

Spring 2022

Air Pollution Exposure Varies Widely

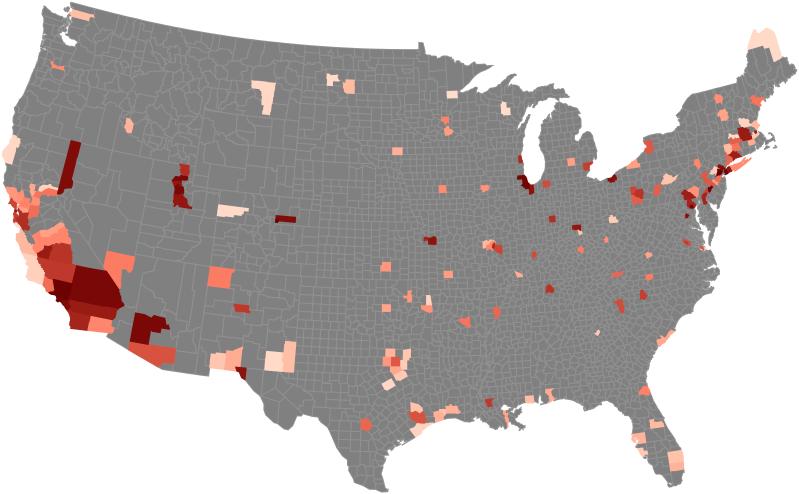
Improving data and associated body of evidence suggests that some forms of air pollution are very spatially concentrated

- Large differences even within very local areas $<$1km

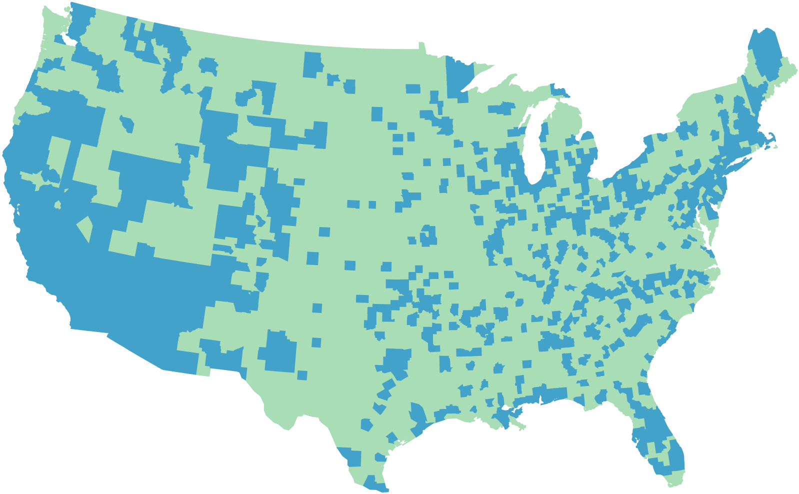
These differences have important implications for the efficacy and efficiency of existing air quality regulations

- Better measurement $=$ better targeting $=$ better policy

NO2 Levels as Measured By EPA Monitors

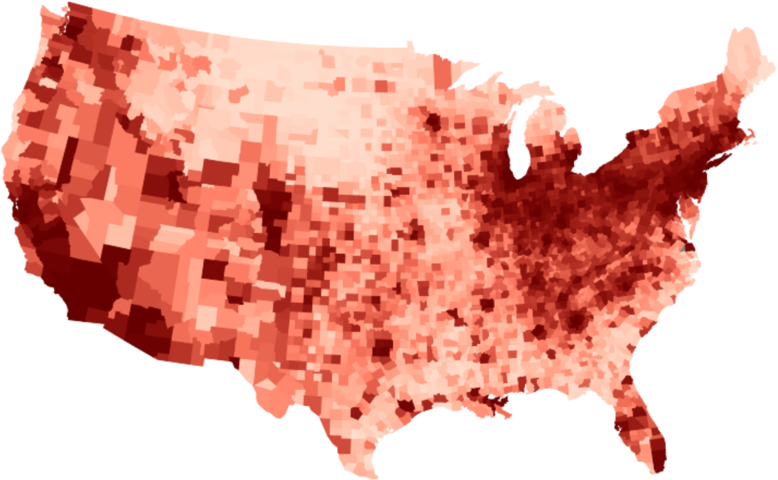


Counties with and without EPA monitors

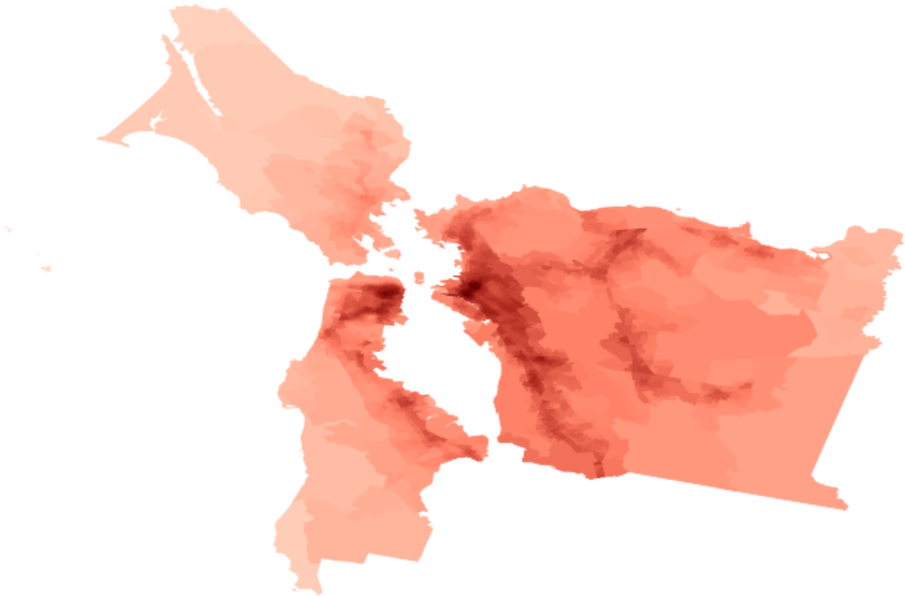


Satellite Data and Remote Sensing Technology Open Up Possibilities

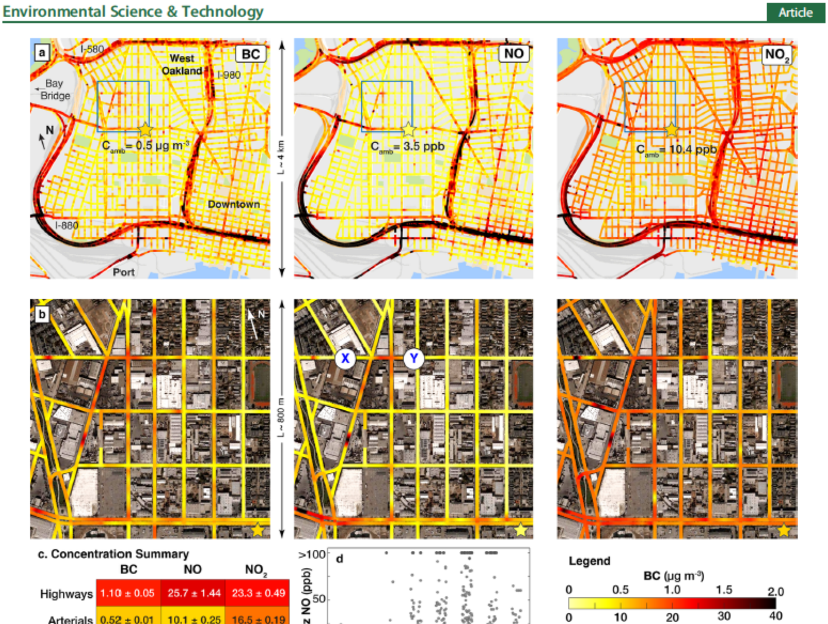
NO_2 Pollution Heterogeneity Across United States



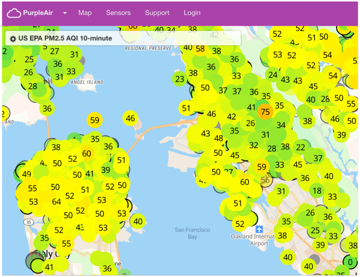
NO_2 Pollution Heterogeneity Within Bay Area



Google Street Car Measurement: Oakland



New Technologies Showing up Everyday



aclima

Spatial Data on Air Pollution: Data Repository

Not Secure | spatialmodel.com/conc...

Click to go forward, hold to see history

Update

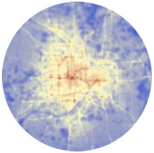
Empirical Model Database

Empirical Model Database (BETA version)

A Repository of Air Quality Estimates From Empirical Models

What is the Empirical Model Database?

Air pollution researchers in the U.S. and worldwide have developed national and continental-scale models for estimating concentrations of several air pollutants. The purpose of the modeling is to estimate concentrations at locations without measurements, and to do so with excellent spatial precision.



Pollution Granularity - Questions and Implications

Newly available data and associated findings raise a number of interesting and important questions:

1. What does this spatial granularity imply for current air quality regulations (e.g. the Clean Air Act is enforced at the county level)?
2. What does this spatial granularity imply for our understanding of environmental inequality or environmental justice?
3. How could policy makers use this data to improve existing policy? Advocacy?

Pollution Granularity - Questions and Implications

Newly available data and associated findings raise a number of interesting and important questions:

1. **What does this spatial granularity imply for current air quality regulations (e.g. the Clean Air Act is enforced at the county level)?**

2. What does this spatial granularity imply for our understanding of environmental inequality or environmental justice?

3. How could policy makers use this data to improve existing policy? Advocacy?

Fowlie, Rubin, and Walker (2019): Bringing Satellite-Based Air Quality Estimates Down to Earth

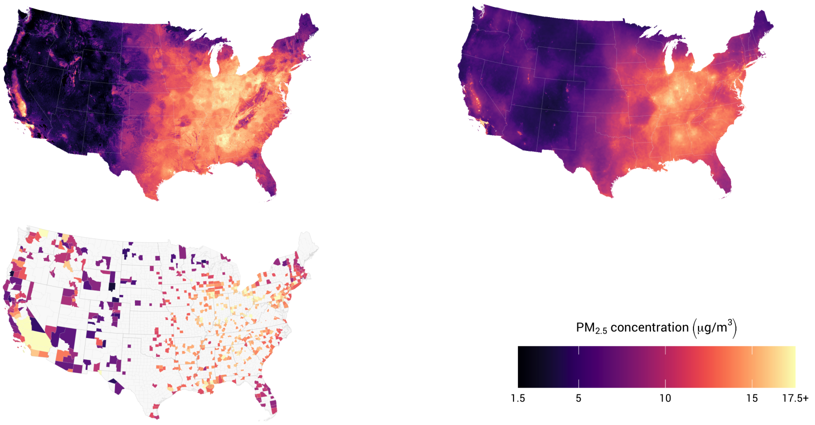
EPA pollution monitor grid is sparse

- Particulate matter pollution less than 2.5 micrometers (PM2.5) poses serious health risks.
- Regulatory limits on ambient PM2.5 concentrations are enforced using measurements collected by a sparse network of regulatory-grade air pollution monitors (US EPA).
- The spatial coarseness of these monitor measurements has potentially significant implications for the design and implementation of existing air quality regulations.

The sensor revolution takes on air pollution!

- Advances in satellite remote sensing support estimation of PM2.5 concentrations at fine spatial resolution.
- We obtained two state-of-the-art, satellite-based data products:
    - Di et al. (2016) predict daily PM2.5 concentrations at nationwide 1 km$\times$1 km grid cells over the period 2000 to 2015.
    - Van Donkelaar et al. (2019) similarly estimate 1 km$\times$1 km resolution for the years 1998-2016.
- We spatially intersect these estimates of annual average concentrations with 2000 census block group boundary files.

(a) Di et al. 2016 | (b) vanDonkelaar et al. 2019



(c) EPA Monitors

Figure: Satellite-Based PM2.5 Measurements and EPA AQS Monitoring Network, 2005

Research Objectives

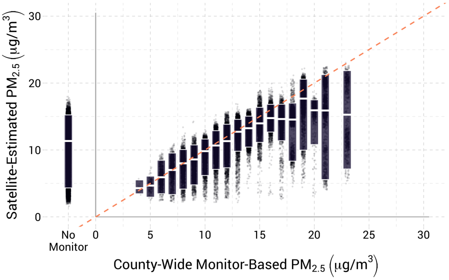
- Use state-of-the-art, satellite-based PM2.5 concentration estimates to assess the extent to which limited network of EPA monitors has led to over and/or under detection of violations of PM2.5 standards.
- Characterize the health and distributional implications of any policy implementation errors, taking satellite-based estimates as truth.
- Characterize imprecision of satellite-based estimates and assess the implications for policy calibration.

Policy Context: National Ambient Air Quality Standards (NAAQS) for PM2.5

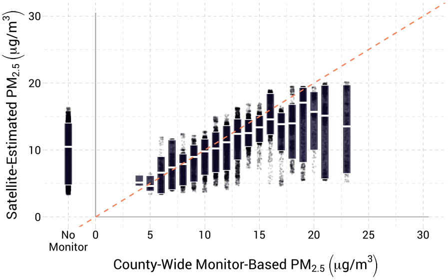
- EPA monitors used to construct two design values (DVs): a 3-year annual average concentration and the 3-year average 98th percentile of 24 hour concentrations.
- If either DV exceeds the PM2.5 NAAQS standard, the monitor's jurisdiction (usu. county) is classified into "non-attainment."
- We use satellite-based estimates to construct annual average design values and compare these to the de jure, monitor-based DVs.
- Focus on 1997 PM2.5 NAAQS (effective 2005) because violations are due to the annual standard of 15 $\mu g/m^3$.

Annual average design values: Monitor-based measures vs. satellite-based estimates

(a) Di et al.



(b) van Donkelaar et al.



Notes: These figures plot the relationship between satellite-based design values and monitor-based design values in 2005. An observation is a census block group. Note the variation in satellite-based design values within each level of EPA monitor-based design values

Conceptual framework: Policy designation errors

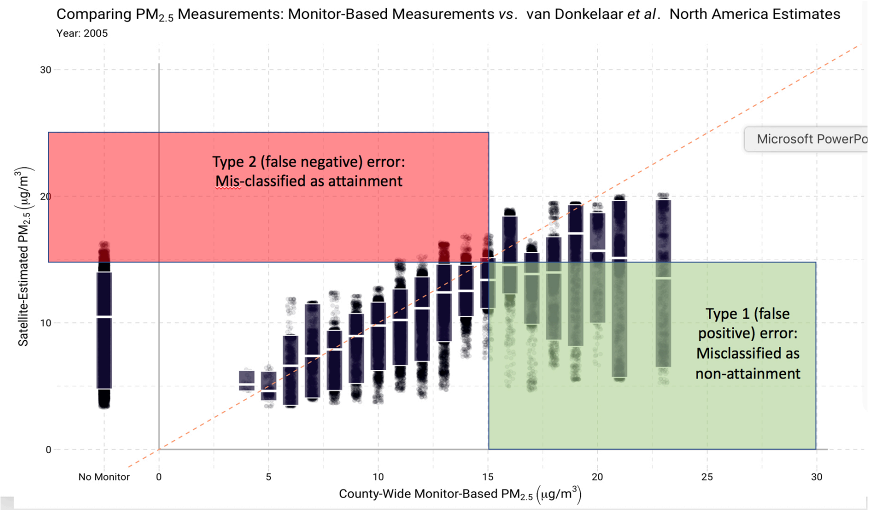


Table: Comparing NAAQS Designations: Monitors and Satellite-Based Estimates

	Satellite Attain		Satellite Nonattain	
	AQS Attain (1)	AQS Nonattain (2) **Type 1 error**	AQS Attain (3) **Type 2 error**	AQS Nonattain (4)
	Panel A: Population Shares (Di et al.)			
Population (millions)	236.8	26.7	7.1	30.2
Pop. Share	78.7%	8.9%	2.4%	10.1%
	Panel B: Population Shares (van Donkelaar et al.)			
Population (millions)	242.6	35.7	1.3	21.3
Pop. Share	80.6%	11.9%	0.4%	7.1%

Notes: These estimates come from comparing satellite-based estimates to EPA AQS monitor data. The satellite-based pollution data have been spatially intersected with census block groups. The column classifications are based on the 2005 3-year Annual Design Values, calculated at the county-level using EPA monitors and at the census block group level using the Di et al. measurements.

Health implications?

- Damages associated with PM2.5 exposure are primarily mortality related: cardiovascular diseases, ischemic heart disease, cerebrovascular disease, respiratory complications.
- Concentration-Response functions relate variation in PM2.5 exposure to variation in mortality rates.
- Hazard ratios typically estimated using log-linear, random-effects Cox proportional-hazard models.
- We estimate impacts of a 1 $\mu g/m^3$ change in PM2.5 concentrations for each census block group based Krewski et al. 2009 and Lepeule et al. 2012.

Table: Mortality Impacts of a 1 $\mu g/m^3$ reduction in PM2.5 (2005)

	Satellite Attain		Satellite Nonattain	
	AQS Attain (1)	AQS Nonattain (2) **Type 1 error**	AQS Attain (3) **Type 2 error**	AQS Nonattain (4)
	Panel A: Di et al.			
Avoided deaths Lower estimate (Krewski et al.)	4,670	589	141	627
Avoided deaths Higher estimate (**?**)Lepeule et al.)	12,762	1,558	387	1,654
	Panel B: van Donkelaar et al.			
Avoided deaths Lower estimate (Krewski et al.)	4,787	781	24	435
Avoided deaths Higher estimate (Lepeule et al.)	13,084	2,085	66	1,127

Results Summary

- Type 1 errors appear more prevalent than Type 2 errors.
- Sub-populations mis-classified as non-attainment (i.e. Type 1) are more urban and have larger minority populations as compared to Type 2 CBGs.
- Mortality implications of Type 1 errors (i.e. avoided deaths) could be much more consequential than the foregone mortality benefits associated with Type 2 errors.

Satellite-based estimates are *estimates*

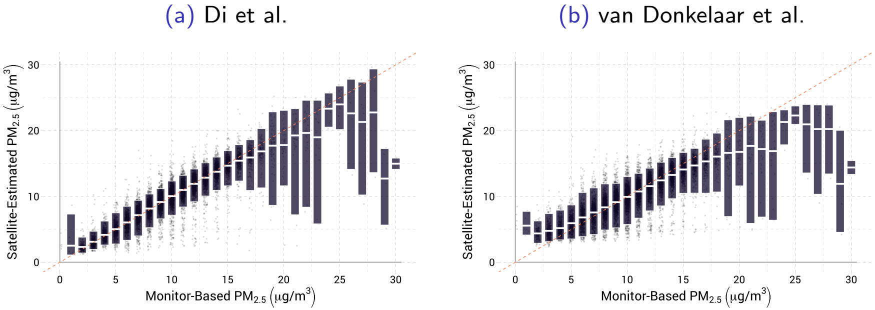


Figure: Comparing PM2.5: Monitors' Measurements vs. Satellite-Based Estimates (monitored census block groups only)

NOTES: The boxes depict the range of estimates (2.5^{th}–97.5^{th} percentiles) from the satellite-based datasets (y axis) for the given PM2.5 level measured at the corresponding (within CBG) AQS monitor (x axis).

Conclusions

- Newly available, spatially resolved pollution data present a host of new opportunities, both for research and policy.
- Using more spatially disaggregated estimates of PM2.5 concentrations to determine NAAQS attainment need not be welfare improving, relative to the current status-quo.
- Our work highlights both data limitations *and* policy design limitations: how should regulatory frameworks evolve to leverage more spatially resolved information about pollution concentrations?
- Work that rigorously explores the precision, bias, limits of these data will also be important in determining the appropriate use of these data.

Pollution Granularity - Questions and Implications

Newly available data and associated findings raise a number of interesting and important questions:

1. What does this spatial granularity imply for current air quality regulations (e.g. the Clean Air Act is enforced at the county level)?

2. **What does this spatial granularity imply for our understanding of environmental inequality or environmental justice?**

3. How could policy makers use this data to improve existing policy? Advocacy?

Currie, Voorheis, and Walker (2021): Environmental Inequality

Salient narrative that low income minorities disproportionately live in areas that are characterized by environmental degradation/ elevated pollution levels etc

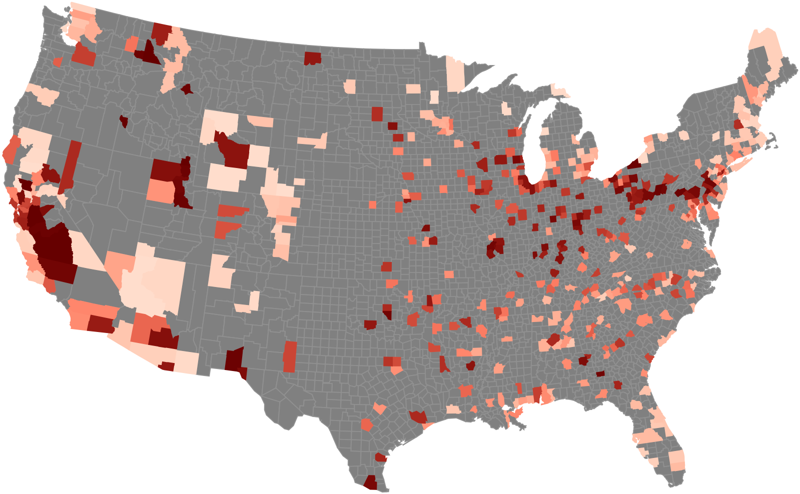
Given rise to the modern day "Environmental Justice" movement

Existing evidence is somewhat piecemeal and indirect

- Proxies for environmental exposure (i.e. proximity to toxic plant)
- Data is scarce (i.e. 775 counties in US with EPA pollution monitors)

Very little evidence on population-wide patterns in racial pollution disparities and/or underlying drivers of these patterns

PM2.5 Levels as Measured By EPA Monitors



Environmental Inequality

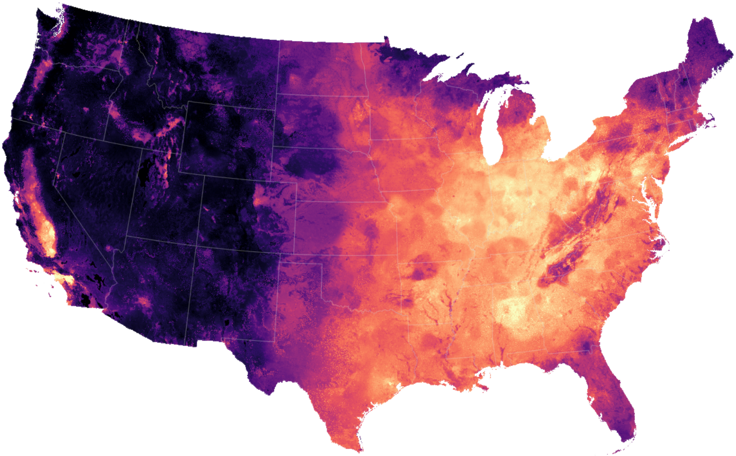
Emerging Opportunities: Increased availability of spatial data on environmental exposure

- Satellite/Remote Sensing, Assimilation/Reanlysis, Land Use Regression, and combinations thereof
- "Solves" both issue of proxies but also representativeness

Newly available data and associated findings raise a number of interesting and important questions:

- To what extent does spatial granularity of exposure alter our understanding of questions pertaining to environmental inequality?

PM2.5 Satellite / Remote Sensing Pollution Measurements



Di et al. (2018), 1km resolution, 2000-2015

This Project

Combine spatially continuous pollution data with large-scale demographic data (Census + ACS) $\Rightarrow$ provide new facts on environmental disparities

- Part 1: Documenting Black/White differences and trends over time
- Part 2: Decomposing Black/White Differences in Exposure
- Part 3: Exploring underlying [causal] mechanisms/explanations.

Data

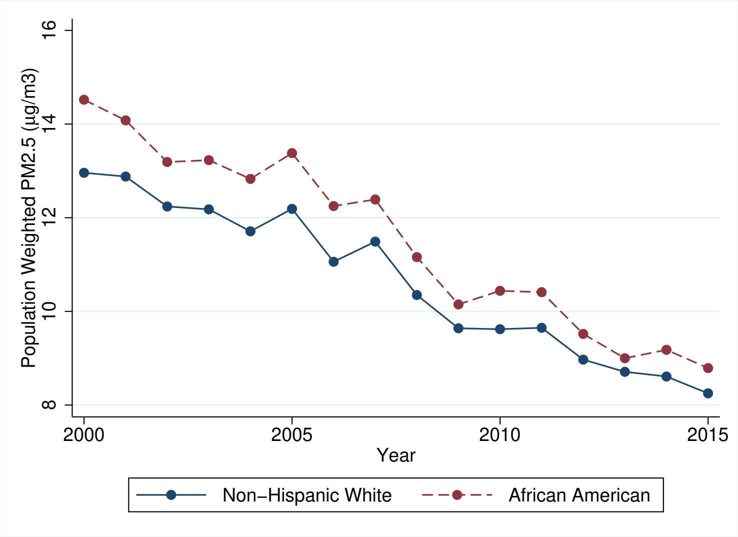
Decennial Census and American Community Survey: 2000-2015

- Administrative microdata with Census-<u>block</u> geo-identifiers
- Survey responses on race, income, education, family size, home-owner, etc...
- $\approx$ 32 million individual responses/observations, sample weights

PM2.5 Pollution: 2000-2015

- 0.01×0.01 grid/raster from Di et al. 2018
- Machine learning / neural net using satellite imagery + monitor ground truth + other inputs / bias corrections
- <u>This project</u>: intersected raster with Census block shapefiles, 2000-2015

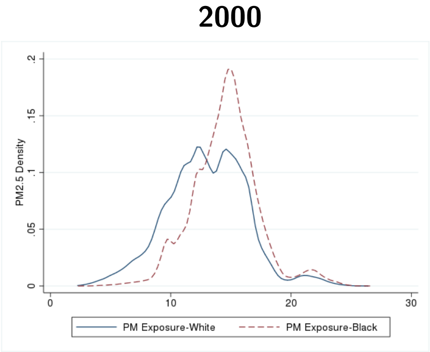
Black White Pollution Trends: PM2.5

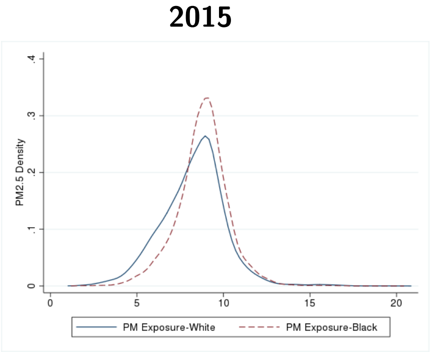


Source: Di et al. (2018)

Black White Pollution Distributions

Changes in other parts of the PM2.5 distribution





Note: Densities censored at 5th and 95th percentiles due to Census disclosure avoidance

What are the Underlying Correlates/Drivers of B-W Pollution Gaps?

Conditional vs. Unconditional Differences in Pollution Exposure

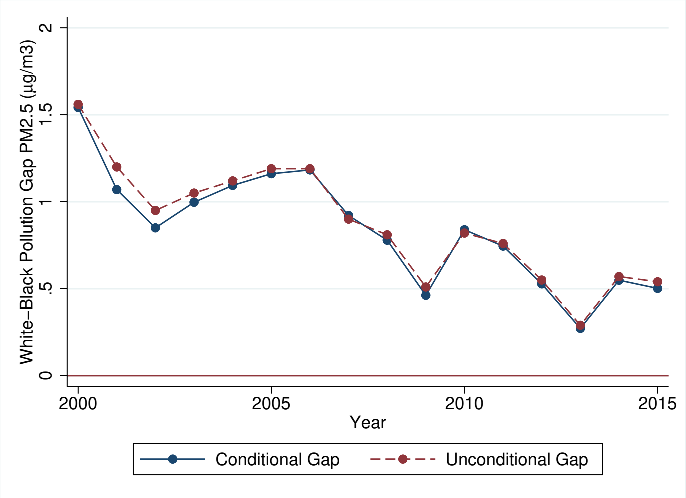
- What fraction of gaps are explained by observable individual characteristics?
  - Is environmental inequality related to economic inequality?
  - Do income differences explain the observed pollution gap?

- A range of individual characteristics (X_{it}) that we can control for:
  - Income, Age, Schooling; Kids, Gender, Homeownership

$$PM2.5_{it} = \beta_t Black_i + X_{it} + \epsilon_{it}$$

- Conditional pollution gap: $\widehat{\beta}_t$

Residual Black-White Pollution Gap: PM2.5



15 separate regressions of pollution regressed on black indicator + controls. Coefficient on black dummy + 95% confidence intervals plotted. Source: Di et al. (2018)

Summary

Individual Characteristics Explain Little B-W Pollution Gap
(or 2000-2015 change in gap)

- Differences in income, family structure, education, etc.. can account for 3% of mean differences, leaving 97% unexplained
- Similar story when looking at quantile differences

Oaxaca-Blinder | Decomposing Long Differences | Decomposing Quantiles | Dinardo, Fortin, Lemieux

- What is the role of relative mobility and/or neighborhood changes in explaining convergence?

Can Differential Migration Explain Observed Convergence?

One simple test is to fix individuals in their 2000 locations and assign them the 2015 pollution of their 2000 census block

- Holding populations fixed, what would be the 2015 B-W gap?

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

Can Differential Migration Explain Observed Convergence?

One simple test is to fix individuals in their 2000 locations and assign them the 2015 pollution of their 2000 census block

- Holding populations fixed, what would be the 2015 B-W gap?

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

Message: 2000-2015 shifts in population contributed to racial convergence, but population shifts are not a central explanatory factor ($\approx$ 13%)

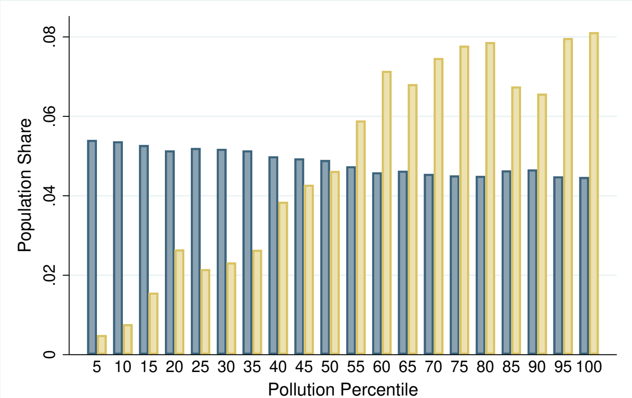
- Instead, black neighborhoods are seeing larger improvements than

Why are Black Neighborhoods Getting (Relatively) Cleaner?

One Possibility: Environmental policy disproportionately improves air quality in areas where African Americans are overrepresented

The Clean Air Act targets / cleans up only the most polluted areas

- Explore how this spatial targeting has affected the black/white pollution gap.



Clean Air Act's National Ambient Air Quality Standards (NAAQS)

EPA sets maximum allowable concentrations of air quality for certain "criteria" pollutants

Enforced at the "County"$\times$Pollutant$\times$Year level

Areas that exceed the EPA pollution threshold in a given year are designated as "Nonattainment"

- Limits on how much pollution may be allowed
- Mandated abatement technology
- Require new sources of pollution to offset their emissions by reducing it from somewhere else within the county.

CAA Particulate Matter (PM2.5) Nonattainment Standards

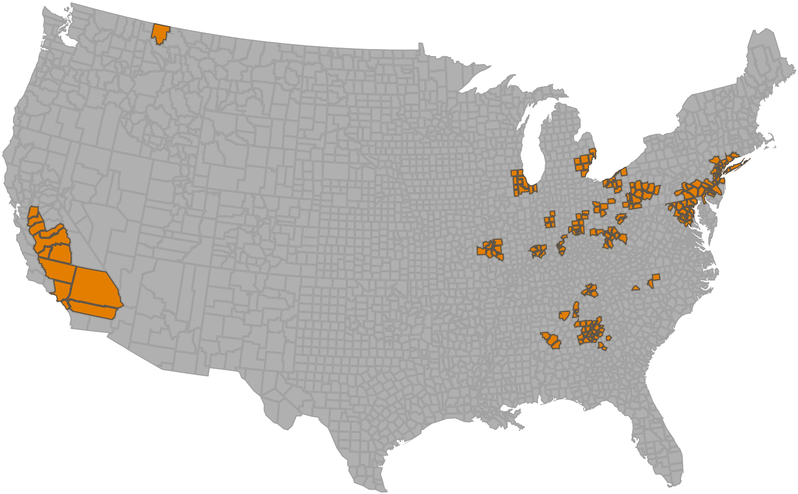
EPA introduced new PM2.5 Standards in 1997

- Nonattainment designations first promulgated in 2005
- $\approx$ 200 counties newly designated as nonattainment for PM2.5

Explore how this county$\times$year variation affects satellite based measures of pollution via difference-in-differences

- i.e. the regulation affects newly designated nonattainment counties in the years after the policy change

Nonattainment Counties in 2005 (New PM2.5 Standard)

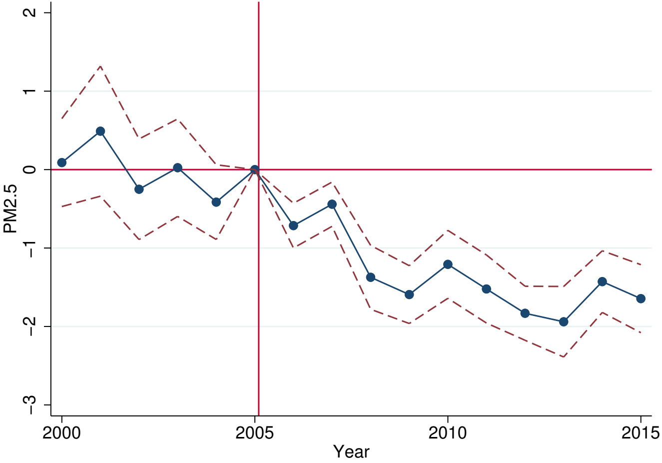


Event Study Representation

$$P_{ict} = \sum_{t=2000}^{2015} \beta(1[Nonattain_c] \times 1[year = t]) + \gamma_c + \rho_t + X'\eta + \epsilon_{ict}$$

Pollution P_{ict} for person i residing in county c in year t with county fixed effects γ_c and state-year fixed effects ρ_t

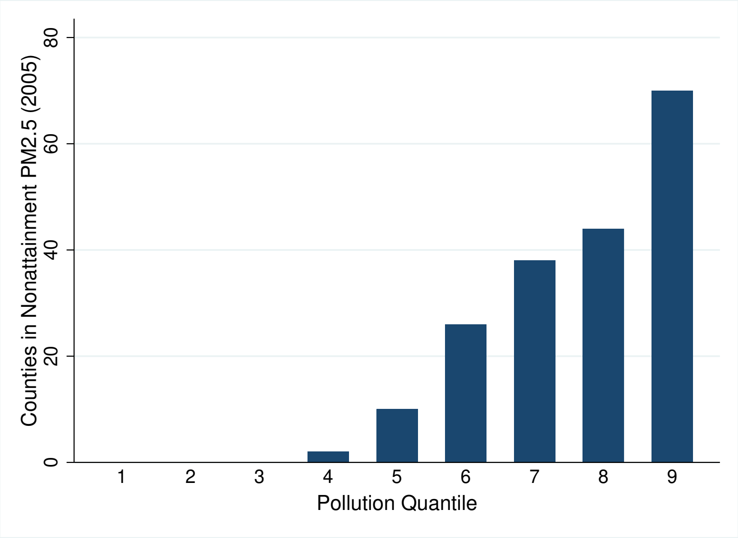
Results: PM2.5 Event Study



On average about a 8% drop in PM2.5 in treated counties in years after NAAQS rollout State-Year

County PM2.5 Nonattainment Designations by Pollution Quantile

CAA Only Affects Part of the Pollution Distribution



Effect of CAA on Pollution Distribution

CAA reduced pollution levels on average, but what about different parts of pollution distribution?

Unconditional Quantile Regression (Firpo et al. (2009))

- Can use Recentered Influence Functions (RIF) to explore causal effect of CAA on different parts of (unconditional) pollution distribution

In practice:

- Apply RIF transformations to PM2.5 for each quantile of interest
- Regression of RIF on covariates $\approx$ effect of covariates on distributional statistic of interest (applied to unconditional distribution).

RIF Overview/Transformations

Unconditional Quantile Regressions (RIF)

For quantile τ denoted Q_τ, the quantile RIF is given by

$$RIF(p, Q_\tau) = Q_\tau + \frac{\tau - \mathbf{1}(p_i < Q_\tau)}{f_p(Q_\tau)}$$

Example: 90th percentile of pollution distribution $\approx 15\mu g/m^3$:

$$RIF(p, 15) = 15 + \frac{90 - \mathbf{1}(p_i < 15)}{f_p(15)}$$

Unconditional Quantile Regressions (RIF)

For quantile τ denoted Q_τ, the quantile RIF is given by

$$RIF(p, Q_\tau) = Q_\tau + \frac{\tau - \mathbf{1}(p_i < Q_\tau)}{f_p(Q_\tau)}$$

Example: 90th percentile of pollution distribution $\approx 15\mu g/m^3$:

$$RIF(p, 15) = 15 + \frac{90 - \mathbf{1}(p_i < 15)}{f_p(15)}$$

Key Property: Expectation of RIF = distributional statistic of interest.

Taking expectation of RIF variable recovers quantile of interest

$$E[RIF(p, Q_\tau)] = Q_\tau$$

Can then use OLS / law of iterated expectations to estimate CAA's effect on the unconditional pollution quantile

Effect of CAA on Pollution Distribution: RIF in Practice

Methodologically:

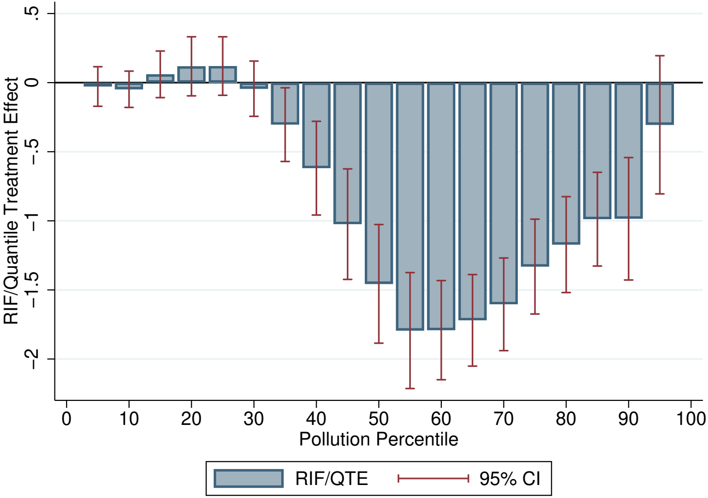
1. Create 20 different RIF variables for each pollution vigintile
2. Replace dependent variable from before with RIF_p

$$RIF^{p}_{ict} = \beta(1[Nonattain_c] \times 1[Post_t]) + \gamma_c + \rho_t + X'\eta + \epsilon_{ict}$$

where $p \in [5, 10, 15, 20, ..., 95]$

- Estimate 20 separate regressions

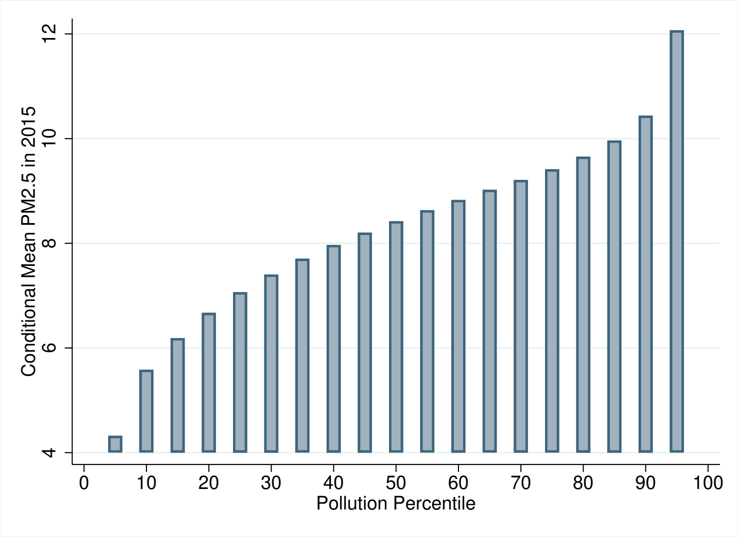
Quantile Treatment Effects of CAA on PM2.5 Distribution



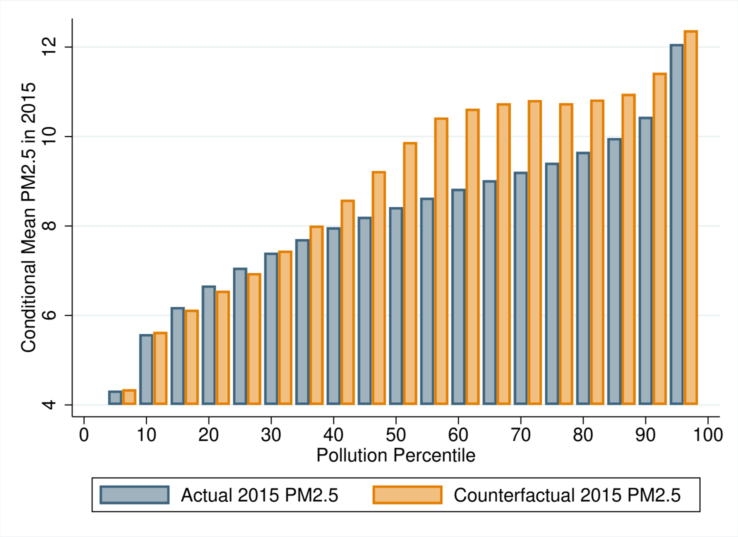
Estimates include county and State-Year fixed effects. Standard errors clustered by CZ

No State-Year FE

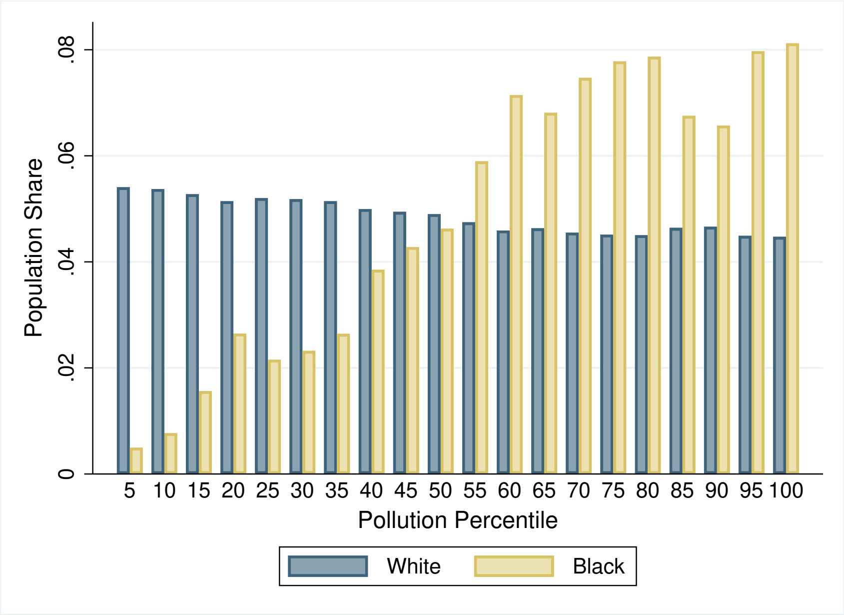
2015 PM2.5 Pollution, Conditional Mean by Quantile



Counterfactual 2015 PM2.5 Pollution Implied by RIF Treatment Effects



Population Share by Pollution Decile and Race



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

(1) PM2.5 Quantile Bin	(2) Actual PM2.5 in 2005	(3) Actual PM2.5 in 2015	(4) Counterfactual PM2.5 in 2015 Without CAA
5	5.32	4.34	4.37
10	7.87	5.63	5.69
15	8.91	6.25	6.18
20	9.65	6.72	6.62
25	10.33	7.11	7.03
30	10.90	7.45	7.56
35	11.42	7.75	8.12
40	11.90	8.01	8.67
45	12.34	8.24	9.28
50	12.73	8.44	9.89
55	13.09	8.65	10.39
60	13.44	8.84	10.57
65	13.80	9.03	10.68
70	14.15	9.22	10.75
75	14.51	9.42	10.71
80	14.91	9.67	10.80
85	15.27	9.98	10.93
90	15.72	10.49	11.41
95	17.01	12.21	12.46

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

Main Counterfactual	
2015 Counterfactual Black-White Gap:	0.97
Counterfactual *Change* in Black-White Gap:	-0.23
Actual *Change* in Black-White Gap:	-0.59
% of Actual Gap Attributable to CAA:	**61.2%**

Wrapping Up

A 60% improvement in the Black-White PM2.5 gap from 2000-2015

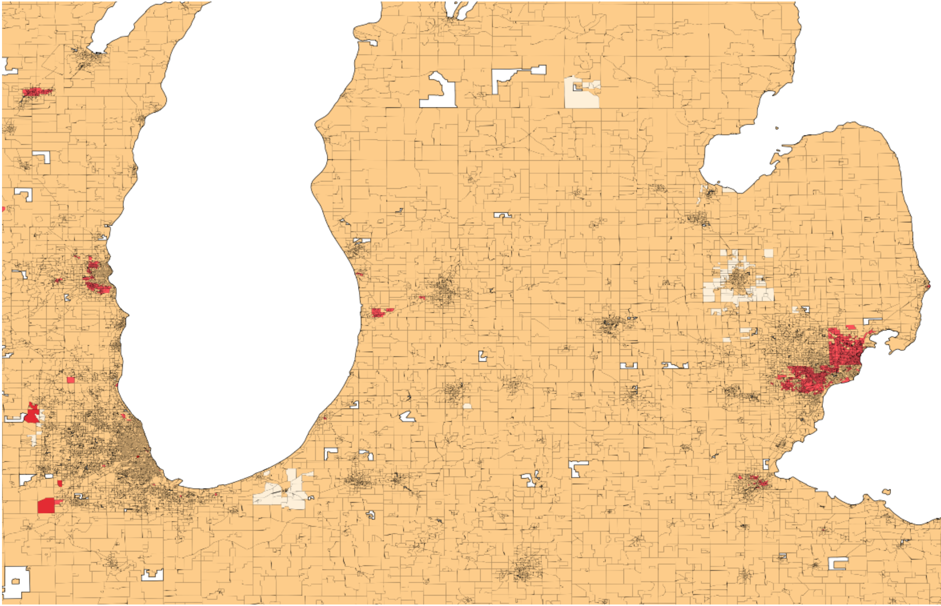
- Existing differences in exposure and reductions in disparities *not* explained by individual characteristics or differential mobility
- Minority communities are seeing greater improvements in air quality in large part due to the targeted nature of the CAA
- The CAA has compressed the pollution distribution from the top, disproportionately benefitting African Americans

Pollution Granularity - Questions and Implications

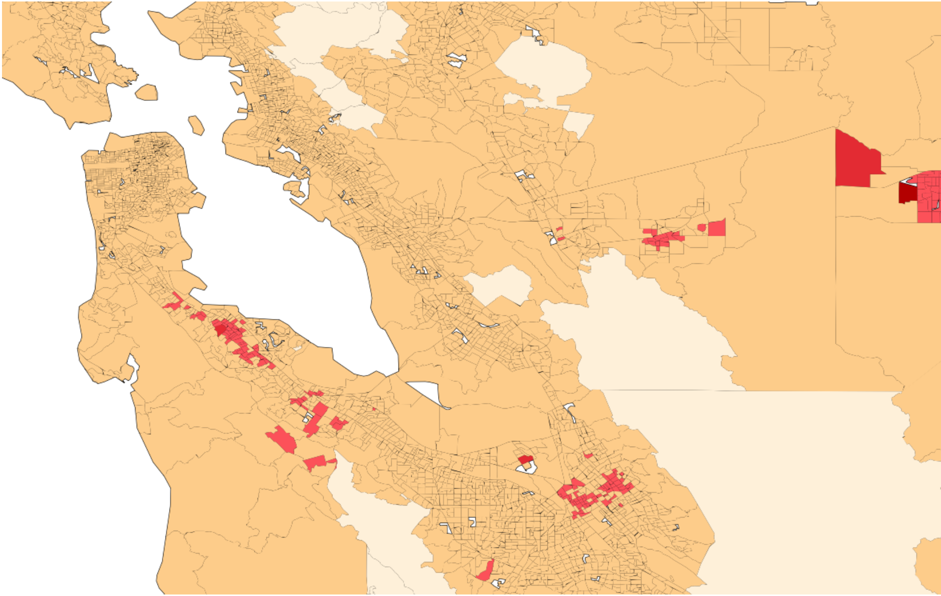
Newly available data and associated findings raise a number of interesting and important questions:

1. What does this spatial granularity imply for current air quality regulations (e.g. the Clean Air Act is enforced at the county level)?
2. What does this spatial granularity imply for our understanding of environmental inequality or environmental justice?
3. **How could policy makers use this data to improve existing policy? Advocacy?**

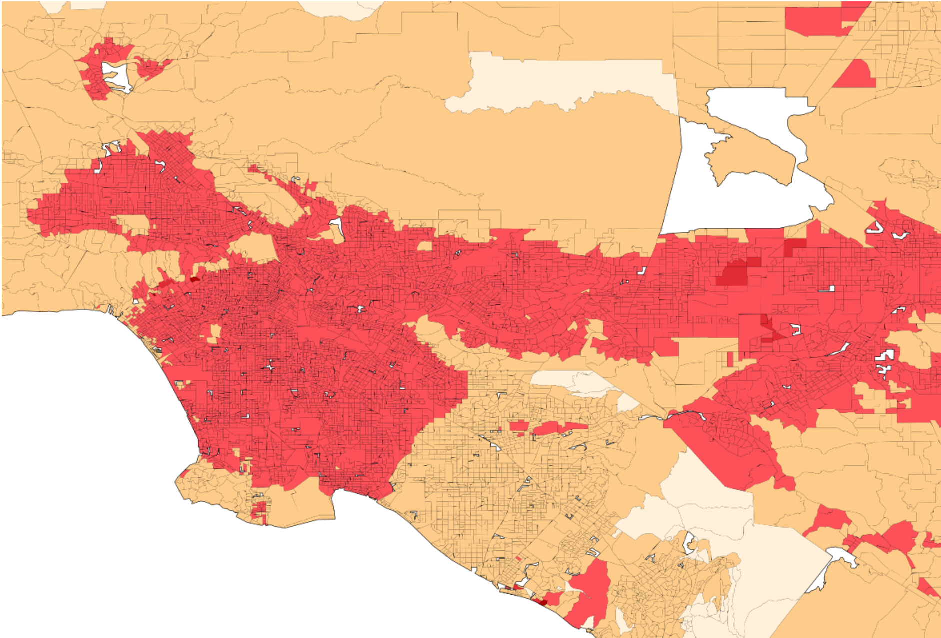
A More Spatially Differentiated Policy?



A More Spatially Differentiated Policy?



A More Spatially Differentiated Policy?



Other Ways Better Data Can Improve Regulatory Policy

Emissions are a classic externality

- Plant owner does not internalize his/her effect on surrounding population health
- Solution: internalize the externality through a tax equal to marginal damages
    - Clean Air Act has almost no marginal incentives!
- Question: what are marginal damages of a facility? Depends on location, population density, wind/weather

This is an empirical question

Other Ways Better Data Can Improve Regulatory Policy

Data Driven Answers:

- Where does 1 ton of CO emitted from a facility end up on average?
- Monitoring technology is becoming cheap and portable
- Instead of relying on engineering / dispersion models, measure directly (or at least validate)
- Piloting / Permitting: sparse grid of monitors to measure emissions dispersion

Convert dispersion estimates to damages via dose-response / valuation

- Pollution "Permit" - charge producer the local marginal damage for every unit of pollutant emitted
- Marginal incentives to abate local emissions

Average Plume/Emissions Dispersion from CA Airports

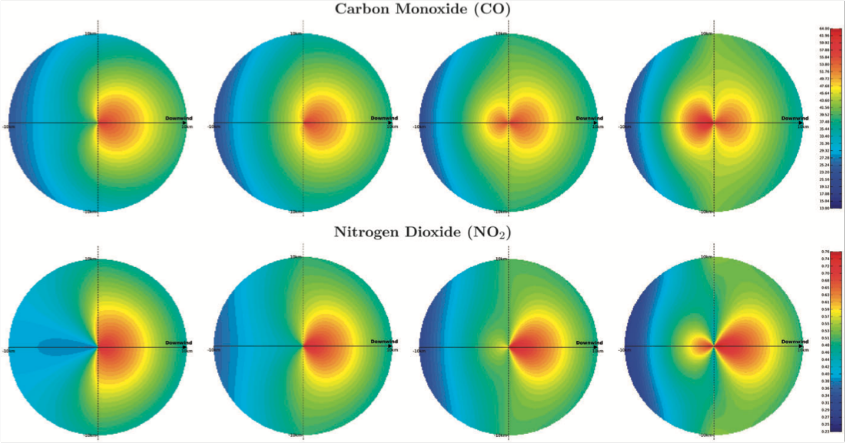


FIGURE 3
Contour maps: marginal impact of taxi time on pollution levels.

Source: Schlenker and Walker (2015)