Adjusting for Unmeasured Spatial Confounding with Distance Adjusted Propensity Scores

Georgia Papadogeorgou

Harvard Chan School of Public Health

1.12.2018

Air Quality Regulations

- Regulations
  - Clean Air Interstate Rule (CAIR)
  - Cross-State Air Pollution Rule (CSAPR)
  - 1990 Clear Air Act
  - Acid Rain Program
- Power plants followed various compliance strategies
- Comparative effectiveness of NO_x emission control technologies on ambient ozone levels

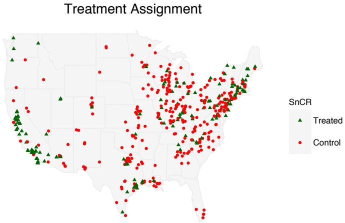
NO_x: Nitric oxide and nitrogen dioxides

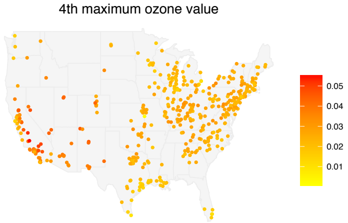
Motivation

- Ozone is a secondary pollutant (Allen, 2002)
  - Created from chemical reactions in the atmosphere
  - Sunlight, Higher temperature
- Selective Catalytic Reduction (SCR) and Selective Non-Catalytic Reduction (SNCR) are the most effective in reducing NO_x
- Reductions in NO_x emissions $\rightarrow$ reduction in ozone concentrations
- Effect of SCR/SNCR on ambient ozone

Data

- Coal and natural gas power plants during June-August 2004
- $A = 1$ if at least half of facility heat input is used by units with installed SCR/SNCR technologies, $A = 0$ otherwise
- 152 treated facilities, 321 controls
- Y: NO_x emissions / 4^{th} maximum ambient ozone concentration





Notation

- For unit i
  - Treatment $A_i \in \{0, 1\}$
  - Potential outcomes $Y_i(1), Y_i(0)$ (SUTVA)
  - Covariates $C_i = (C_{i1}, C_{i2}, \ldots, C_{ip})$

Notation

- For unit i
  - Treatment $A_i \in \{0, 1\}$
  - Potential outcomes $Y_i(1), Y_i(0)$ (SUTVA)
  - Covariates $C_i = (C_{i1}, C_{i2}, \ldots, C_{ip})$
- Average Treatment Effect on the Treated:

$$ATT = E[Y(1) - Y(0)|A = 1]$$

Notation

- For unit i
  - Treatment $A_i \in \{0, 1\}$
  - Potential outcomes $Y_i(1), Y_i(0)$ (SUTVA)
  - Covariates $C_i = (C_{i1}, C_{i2}, \ldots, C_{ip})$
- Average Treatment Effect on the Treated:

$$ATT = E[Y(1) - Y(0)|A = 1]$$

- $P(A = 1|\mathbf{C}) \in (0, 1)$
- $Y(1), Y(0) \amalg A|\mathbf{C}$

Notation

- For unit i
  - Treatment $A_i \in \{0, 1\}$
  - Potential outcomes $Y_i(1), Y_i(0)$ (SUTVA)
  - Covariates $C_i = (C_{i1}, C_{i2}, \ldots, C_{ip})$
- Average Treatment Effect on the Treated:

$$ATT = E[Y(1) - Y(0)|A = 1]$$

- $P(A = 1|\mathbf{C}) \in (0, 1)$
- $Y(1), Y(0) \amalg A|\mathbf{C}$
- Propensity score matching
  - PS model $P(A = 1|\mathbf{C})$
  - Match treated units to controls with similar PS estimates

Unmeasured spatial confounding

- Confounders $\mathbf{C} = (\mathbf{X}, \mathbf{U})$
  - $\mathbf{X}$ are observed, $\mathbf{U}$ are unobserved
- If $\mathbf{U}$ varies spatially, can we adjust for it?

Unmeasured spatial confounding

- Confounders $\mathbf{C} = (\mathbf{X}, \mathbf{U})$
  - $\mathbf{X}$ are observed, $\mathbf{U}$ are unobserved
- If $\mathbf{U}$ varies spatially, can we adjust for it?

- Temperature, and weather conditions may confound the relationship of NO_x control strategies and ambient ozone.
  - Temperature, barometric pressure, humidity
- Weather and atmospheric covariate information varies spatially

Unmeasured Spatial Confounding

- Observed variables $\mathbf{X}$:
  - Use the propensity score to adjust for the observed confounders
  - $P(A_i = 1|X_i) = f(X_i) = \text{expit}\left(X_i^T \beta\right)$
- Unmeasured spatial confounders $\mathbf{U}$
  - The correlation of $\mathbf{U}$ is high for small enough distances
  - If a matched pair is sufficiently close, the treated and control units will have similar values of $\mathbf{U}$

Rosenbaum and Rubin (1983)

Distance Adjusted Propensity Score Matching

- For a treated unit i and a control unit j define

$$DAPS_{ij} = w|PS_i - PS_j| + (1-w) * Dist_{ij},\ w \in [0,1]$$

  where PS propensity score estimates, and $Dist$ spatial proximity.

- w expresses our belief of the relative importance of the observed and unobserved confounders
- $Dist$ is the measure the expresses our belief of similarity of U as a function of distance

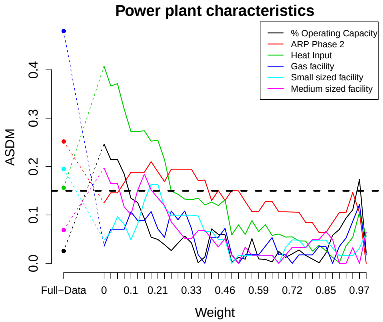
Choosing w

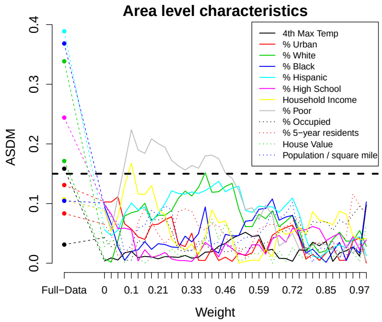
$$DAPS_{ij} = w|PS_i - PS_j| + (1-w) * Dist_{ij}, \; w \in [0,1]$$

- Interplay between distance of observed covariates and distance of matched pairs
- w can be specified using subject-matter knowledge on an unmeasured spatial confounder
- Automated procedure
  - Re-calculates DAPS and performs matching for many values of w
  - Balance of the observed covariates is assessed
  - The smallest value that acheives covariate balance is chosen

Checking covariate balance

- Absolute standardized difference of means as a function of w

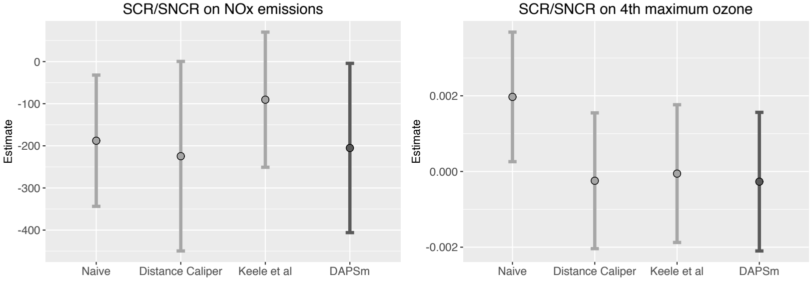




Matches

- Average distance of matched pairs
  - Naïve: 1066 miles
  - DAPSm: 141 miles

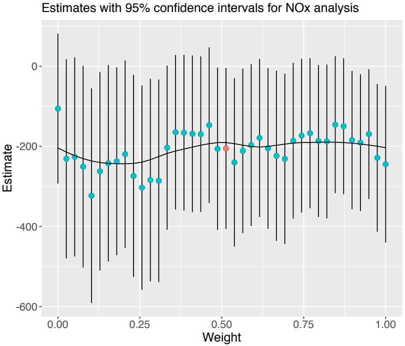
Results

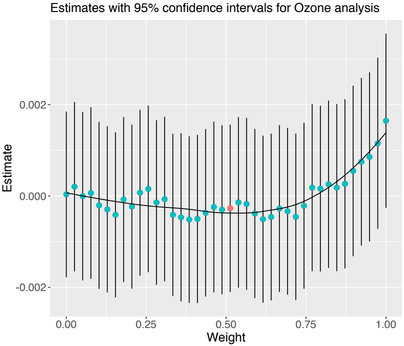


- 205 NO_x tons (95% CI: 4 – 406)
- -0.27 parts per billion (95% CI: -2.1 – 1.56)

The national ozone air quality standard of 70 parts per billion.
Keele et al. (2015)

DAPSm results as a function of w





Conclusions

- SCR/SNCR control technologies seem to be associated with reduced NO_x emissions
- Their effect on ozone is not significant
- Unobserved confounding can lead to severe bias of estimates

References

Allen, J. (2002). Chemistry in the Sunlight. *Earth Observatory NASA* .

Keele, L., Titiunik, R., and Zubizarreta, J. (2015). Enhancing a Geographic Regression Discontinuity Design Through Matching to Estimate the Effect of Ballot Initiatives on Voter Turnout. *Journal of Royal Statistical Society A* **178,** 223–239.

Rosenbaum, P. R. and Rubin, D. B. (1983). The Central Role of the Propensity Score in Observational Studies for Causal Effects. *Biometrika* **70,** 41–55.