Mitigating Unobserved Spatial Confounding Bias with Mixed Models

Georgia Papadogeorgou

ACIC, May 23, 2019

Causal inference and unmeasured structured confounding

- Causal inference formalizes the notion of an *effect*, and provides identifiability assumptions
- One often invoked assumption is the no unmeasured confounding assumption (+ positivity = ignorability)
- No unmeasured confounding cannot be tested but sensitivity of results to violations of this assumption can be evaluated [Rosenbaum, 2002]

Causal inference and unmeasured structured confounding

- Causal inference formalizes the notion of an *effect*, and provides identifiability assumptions
- One often invoked assumption is the no unmeasured confounding assumption (+ positivity = ignorability)
- No unmeasured confounding cannot be tested but sensitivity of results to violations of this assumption can be evaluated [Rosenbaum, 2002]

- Can we use unmeasured confounders' *structure* to adjust for them?
  - Spatial structure: spatial variables vary continuously over space

Unmeasured spatial confounding in the literature

- Causal literature [Keele et al., 2015, Papadogeorgou et al., 2018]
  - Spatial information in treatment assignment
- Spatial literature [Hodges and Reich, 2010, Paciorek, 2010]
  - Inspired by spatial structure in regression residuals

Unmeasured spatial confounding in the literature

- Causal literature [Keele et al., 2015, Papadogeorgou et al., 2018]
  - Spatial information in treatment assignment

  → *not immediately compatible with spatial models*

- Spatial literature [Hodges and Reich, 2010, Paciorek, 2010]
  - Inspired by spatial structure in regression residuals

  → *Confusion about what these spatial models are capable of accounting for*

  → *Spatial random effects do not eliminate bias*

Unmeasured spatial confounding in the literature

- Causal literature [Keele et al., 2015, Papadogeorgou et al., 2018]
  - Spatial information in treatment assignment

  → *not immediately compatible with spatial models*
- Spatial literature [Hodges and Reich, 2010, Paciorek, 2010]
  - Inspired by spatial structure in regression residuals

  → *Confusion about what these spatial models are capable of accounting for*

  → *Spatial random effects do not eliminate bias*
- Spatial and causal inference literatures remain largely separated
- Bridging the two strands of literature by
  - Unmeasured confounding within the causal inference framework
  - Estimation using models and tools common among spatial statisticians

Patrick Schnell

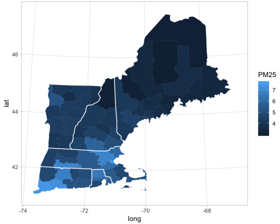
Spatial data and causal inference in air pollution research

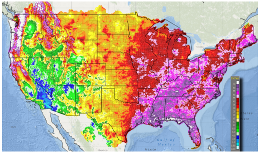
- The scientific questions are causal
  - Do emissions cause pollution?
  - What effect does an intervention on polluting sources have on air pollution concentrations?

https://kcstormfront.wordpress.com/2015/01/11/2014-in-review/

Spatial data and causal inference in air pollution research

- The scientific questions are causal
  - Do emissions cause pollution?
  - What effect does an intervention on polluting sources have on air pollution concentrations?
- The data are spatial
  - Spatially-indexed
  - Exposure, outcome, and covariates are spatially structured
  - Unmeasured confounders are spatial!





https://kcstormfront.wordpress.com/2015/01/11/2014-in-review/

Notation

- For unit i
  - Treatment or exposure $Z_i \in \mathcal{Z}$
  - Potential outcomes $\{Y_i(z), z \in \mathcal{Z}\}$
  - Observed outcome $Y_i = Y_i(Z_i)$
  - Covariates $\boldsymbol{W}_i = (W_{i1}, W_{i2}, \ldots, W_{ip})$
- Average potential outcome: $\overline{Y}(z) = E[Y(z)]$

Notation

- For unit i
  - Treatment or exposure $Z_i \in \mathcal{Z}$
  - Potential outcomes $\{Y_i(z), z \in \mathcal{Z}\}$
  - Observed outcome $Y_i = Y_i(Z_i)$
  - Covariates $\boldsymbol{W}_i = (W_{i1}, W_{i2}, \ldots, W_{ip})$
- Average potential outcome: $\overline{Y}(z) = E[Y(z)]$
- Common identifiability assumptions
  - Positivity: $p(Z = z|\boldsymbol{W}) > 0, z \in \mathcal{Z}$
  - No unmeasured confounding: $Y(z) \perp\!\!\!\perp Z|\boldsymbol{W}$
- Estimate the average potential outcome via propensity score methods, outcome regression, or combinations

Notation

- For unit i
  - Treatment or exposure $Z_i \in \mathcal{Z}$
  - Potential outcomes $\{Y_i(z), z \in \mathcal{Z}\}$
  - Observed outcome $Y_i = Y_i(Z_i)$
  - Covariates $\boldsymbol{W}_i = (W_{i1}, W_{i2}, \dots, W_{ip})$
- Average potential outcome: $\overline{Y}(z) = E[Y(z)]$
- Common identifiability assumptions
  - Positivity: $p(Z = z|\boldsymbol{W}) > 0, z \in \mathcal{Z}$
  - No unmeasured confounding: $Y(z) \perp\!\!\!\perp Z|\boldsymbol{W}$
- Estimate the average potential outcome via propensity score methods, outcome regression, or combinations
- Confounders $\boldsymbol{W} = (\boldsymbol{W}^m, \boldsymbol{W}^u)$, $\boldsymbol{W}^m$ are observed, $\boldsymbol{W}^u$ are unobserved
- If $\boldsymbol{W}^u$ vary spatially, can we adjust for it?

Potential outcomes

- We assume the following *true* model for the potential outcomes:

$$Y_i(z) = \eta(z, \boldsymbol{W}^m) + g(\boldsymbol{W}^u) + \varepsilon_i$$

  - $\boldsymbol{W}^u$ are unmeasured variables, denote $U = g(\boldsymbol{W}^u)$
  - g is such that $E[g(\boldsymbol{W}^u)] = E[U] = 0$
  - Additive model, $\boldsymbol{W}^u$ do not interact with Z and $\boldsymbol{W}^m$
- For ease of presentation, assume $\boldsymbol{W}^m$ empty, $\eta(z) = \beta_0 + \beta_1 z$
- Focus on $\beta_1 = \overline{Y}(z+1) - \overline{Y}(z)$

Bias of common estimators, and the affine estimator

If we could fit model $Y \sim Z + U$, we could estimate β_1 without bias.

Bias of common estimators, and the affine estimator

If we could fit model $Y \sim Z + U$, we could estimate β_1 without bias.

- $Y \sim Z \to \widehat{\boldsymbol{\beta}}$

- $Y \sim Z + \text{Spatial RE} \to \widetilde{\boldsymbol{\beta}}$

Bias of common estimators, and the affine estimator

If we could fit model $Y \sim Z + U$, we could estimate β_1 without bias.

- $Y \sim Z \to \widehat{\boldsymbol{\beta}}$
  - $\mathrm{E}\big(\widehat{\boldsymbol{\beta}}|\boldsymbol{Z}\big) = \boldsymbol{\beta} + \left(\mathbf{X}^\intercal\mathbf{X}\right)^{-1}\mathbf{X}^\intercal\mathrm{E}(\boldsymbol{U}|\boldsymbol{Z})$
- $Y \sim Z +$ Spatial RE $\to \widetilde{\boldsymbol{\beta}}$
  - $\mathrm{E}\big(\widetilde{\boldsymbol{\beta}}|\boldsymbol{Z}\big) = \boldsymbol{\beta} + \{\mathbf{X}^\intercal(\mathrm{Var}[\boldsymbol{Y}|\boldsymbol{Z}])^{-1}\mathbf{X}\}^{-1}\mathbf{X}^\intercal(\mathrm{Var}[\boldsymbol{Y}|\boldsymbol{Z}])^{-1}\mathrm{E}[\boldsymbol{U}|\boldsymbol{Z}]$

    where $\mathbf{X} = (\mathbf{1}, \boldsymbol{Z})$

Bias of common estimators, and the affine estimator

If we could fit model $Y \sim Z + U$, we could estimate β_1 without bias.

- $Y \sim Z \to \widehat{\boldsymbol{\beta}}$
  - $\mathrm{E}\big(\widehat{\boldsymbol{\beta}}|\boldsymbol{Z}\big) = \boldsymbol{\beta} + (\mathbf{X}^\intercal\mathbf{X})^{-1}\,\mathbf{X}^\intercal\mathrm{E}(\boldsymbol{U}|\boldsymbol{Z})$
- $Y \sim Z +$ Spatial RE $\to \widetilde{\boldsymbol{\beta}}$
  - $\mathrm{E}\big(\widetilde{\boldsymbol{\beta}}|\boldsymbol{Z}\big) = \boldsymbol{\beta} + \{\mathbf{X}^\intercal(\mathrm{Var}[\boldsymbol{Y}|\boldsymbol{Z}])^{-1}\mathbf{X}\}^{-1}\mathbf{X}^\intercal(\mathrm{Var}[\boldsymbol{Y}|\boldsymbol{Z}])^{-1}\mathrm{E}[\boldsymbol{U}|\boldsymbol{Z}]$

    where $\mathbf{X} = (\mathbf{1}, \boldsymbol{Z})$
- Identify the bias term, and subtract it

$$\bar{\boldsymbol{\beta}} = \{\mathbf{X}^\intercal(\mathrm{Var}[\boldsymbol{Y}|\boldsymbol{Z}])^{-1}\mathbf{X}\}^{-1}\mathbf{X}^\intercal(\mathrm{Var}[\boldsymbol{Y}|\boldsymbol{Z}])^{-1}\{\boldsymbol{Y} - \mathrm{E}[\boldsymbol{U}|\boldsymbol{Z}]\}$$

- Find a way to identify $E[\boldsymbol{U}|\boldsymbol{Z}]$!

A Gaussian Markov random field construction of the joint distribution

- Assumptions on the joint distribution of $(\boldsymbol{U}, \boldsymbol{Z})$ to identify $E[\boldsymbol{U}|\boldsymbol{Z}]$
- $(\boldsymbol{U}, \boldsymbol{Z})$ is mean 0 normal

1. **Cross-Markov property:** $p(Z_i|\boldsymbol{Z}_{-i}, \boldsymbol{U}) = p(Z_i|\boldsymbol{Z}_{-i}, U_i)$,
2. **Constant conditional correlation:** $\mathrm{Cor}(U_i, Z_i|\boldsymbol{U}_{-i}, \boldsymbol{Z}_{-i}) = \rho$.

A Gaussian Markov random field construction of the joint distribution

- Assumptions on the joint distribution of $(\boldsymbol{U}, \boldsymbol{Z})$ to identify $E[\boldsymbol{U}|\boldsymbol{Z}]$
- $(\boldsymbol{U}, \boldsymbol{Z})$ is mean 0 normal

1. **Cross-Markov property:** $p(Z_i|\boldsymbol{Z}_{-i}, \boldsymbol{U}) = p(Z_i|\boldsymbol{Z}_{-i}, U_i)$,
2. **Constant conditional correlation:** $\mathrm{Cor}(U_i, Z_i|\boldsymbol{U}_{-i}, \boldsymbol{Z}_{-i}) = \rho$.

$$\begin{pmatrix} \boldsymbol{U} \\ \boldsymbol{Z} \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mathbf{0} \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{G} & \mathbf{Q} \\ \mathbf{Q}^\intercal & \mathbf{H} \end{pmatrix}^{-1}\right],$$

- $\mathbf{Q}$ is diagonal, and $q_{ii} = -\rho\sqrt{g_{ii}h_{ii}}$
- For areal data, we specify $\mathbf{G}, \mathbf{H}$ as CAR

Calculating the affine estimator

- Integrating $\mathbf{U}|\boldsymbol{Z}$ out

$$\boldsymbol{Y}|\boldsymbol{Z} \sim \mathcal{N}[\mathbf{X}\boldsymbol{\beta} - \mathbf{G}^{-1}\mathbf{Q}\boldsymbol{Z}, \mathbf{G}^{-1} + \mathbf{R}^{-1}],$$
$$\boldsymbol{Z} \sim \mathcal{N}[\mathbf{0}, (\mathbf{H} - \mathbf{Q}^\intercal\mathbf{G}^{-1}\mathbf{Q})^{-1}]$$

  where $\mathbf{R}^{-1} = \mathrm{Cov}(\boldsymbol{\varepsilon})$

- Parameters are estimated based on the restricted likelihood

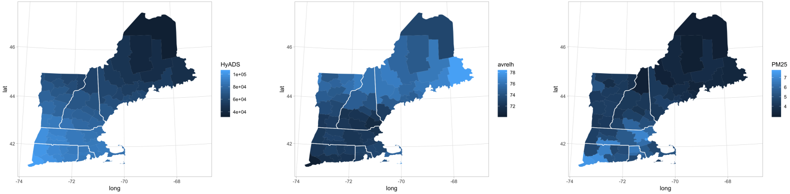
$$RL \propto C_1 \exp\left[-\frac{1}{2}\left\{(\boldsymbol{Y} - \mathbf{B}\boldsymbol{Z})^\intercal C_2(\boldsymbol{Y} - \mathbf{B}\boldsymbol{Z}) + \boldsymbol{Z}^\intercal\mathbf{A}^{-1}\boldsymbol{Z}\right\}\right]$$

  where $\mathbf{A} = (\mathbf{H} - \mathbf{Q}^\intercal\mathbf{G}^{-1}\mathbf{Q})^{-1}$, and $\mathbf{B} = -\mathbf{G}^{-1}\mathbf{Q}$

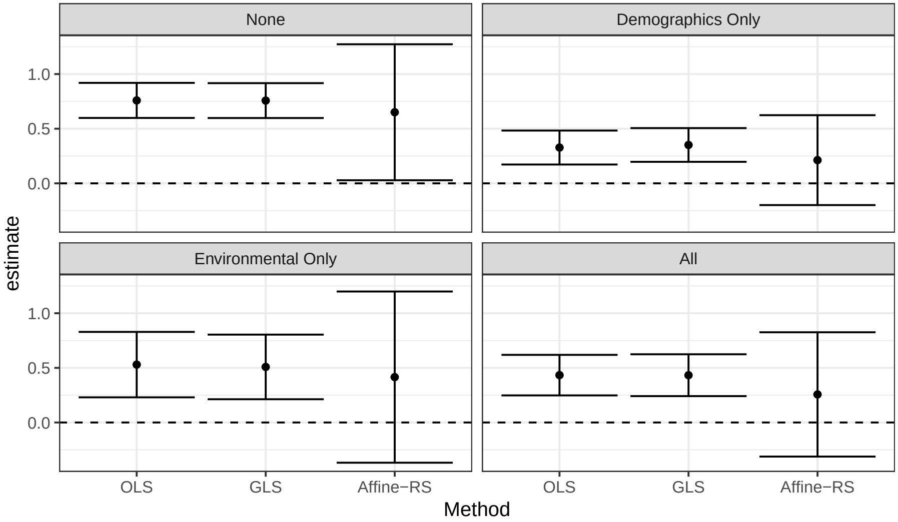
- *Spatial scale restriction* [Paciorek, 2010]
- We calculate $\bar{\boldsymbol{\beta}}$ using the RL maximizers

Data

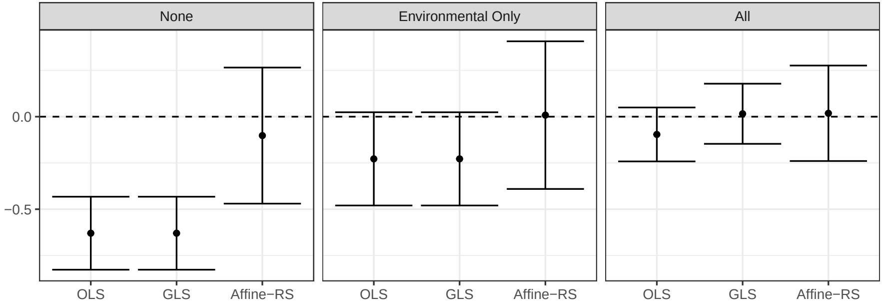
- Counties in New England in 2012
- Z: Emissions from coal power plants in the county [Henneman et al., 2019]
- Y: Average annual $\text{PM}_{2.5}$ concetration
- Covariates: Power plant characteristics, demographics, weather



Effect of coal emissions on ambient $PM_{2.5}$



Effect of relative humidity on ambient $PM_{2.5}$



Conclusions

- Unmeasured spatial confounding is not identified in the analysis of the effect of coal emissions on ambient $PM_{2.5}$ concentrations
- The affine estimator appears to mitigate unmeasured spatial bias in the analysis of the effect of relative humidity on $PM_{2.5}$

Conclusions

- Unmeasured spatial confounding is not identified in the analysis of the effect of coal emissions on ambient $PM_{2.5}$ concentrations
- The affine estimator appears to mitigate unmeasured spatial bias in the analysis of the effect of relative humidity on $PM_{2.5}$

- Unmeasured confounding is one of the main criticisms of air pollution epidemiology
- We can address the sensitivity of results through
  - Sensitivitiy analysis
  - Analysis mitigating bias by unmeasured structured confounders

References

Lucas R.F. Henneman, Christine Choirat, Cesunica Ivey, Kevin Cummiskey, and Corwin M. Zigler. Characterizing population exposure to coal emissions sources in the United States using the HyADS model. *Atmospheric Environment*, 203:271–280, apr 2019.

James S Hodges and Brian J Reich. Adding Spatially-Correlated Errors Can Mess Up the Fixed Effect You Love. *The American Statistician*, 64(4):325–334, 2010.

Luke Keele, Rocio Titiunik, and José R Zubizarreta. Enhancing a geographic regression discontinuity design through matching to estimate the effect of ballot initiatives on voter turnout. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(1): 223–239, 2015.

Christopher J Paciorek. The importance of scale for spatial-confounding bias and precision of spatial regression estimators. *Statistical Science*, 25(1):107, 2010.

Georgia Papadogeorgou, Christine Choirat, and Corwin M Zigler. Adjusting for unmeasured spatial confounding with distance adjusted propensity score matching. *Biostatistics*, 20(2): 256–272, 2018.

Paul R. Rosenbaum. *Sensitivity to Hidden Bias*, pages 105–170. Springer New York, New York, NY, 2002. ISBN 978-1-4757-3692-2.