Unmeasured spatial confounding

Georgia Papadogeorgou

University of Florida

Learning from Interventions, Simons Institute, February 17 2022

Joint work with Cory Zigler (UT Austin),
Christine Choirat (Swiss Data Science Center),
and Patrick Schnell (Ohio State)

Spatial confounding: What does it mean?

- Spatial statistics and causal inference have different views
- Common thread: Confounding hides what we want to learn
- Difference: The goal, what it is that we want to learn

Spatial confounding: What does it mean?

- Spatial statistics and causal inference have different views
- Common thread: Confounding hides what we want to learn
- Difference: The goal, what it is that we want to learn
- Notation $\rightsquigarrow$ $\begin{cases} \text{Treatment or exposure} & Z_i \\ \text{Potential outcomes} & Y_i(z) \\ \text{Outcome} & Y_i = Y_i(Z_i) \\ \text{Measured covariates} & \boldsymbol{C}_i \end{cases}$

Spatial confounding in spatial statistics

The goal: Learn the relationship $Y \mid Z, \boldsymbol{C}$

- Spatial Model 1: $Y \sim Z + \boldsymbol{C}$
  $\rightsquigarrow$ Residuals are spatially correlated!

The goal: Learn the relationship $Y \mid Z, \boldsymbol{C}$

- Spatial Model 1: $Y \sim Z + \boldsymbol{C}$
  $\rightsquigarrow$ Residuals are spatially correlated!
- Spatial Model 2: $Y \sim Z + \boldsymbol{C} + U$
  $\rightsquigarrow$ U spatial random effect
  $\rightsquigarrow$ Reason for U: Inference, capture residual spatial dependence
  $\rightsquigarrow$ U is given a correlation structure (Exponential, Matérn, etc)

Spatial confounding in spatial statistics

The goal: Learn the relationship $Y \mid Z, \boldsymbol{C}$

- Spatial Model 1: $Y \sim Z + \boldsymbol{C}$
  $\rightsquigarrow$ Residuals are spatially correlated!

- Spatial Model 2: $Y \sim Z + \boldsymbol{C} + U$
  $\rightsquigarrow$ U spatial random effect
  $\rightsquigarrow$ Reason for U: Inference, capture residual spatial dependence
  $\rightsquigarrow$ U is given a correlation structure (Exponential, Matérn, etc)

- U is collinear with the exposure Z
  $\rightsquigarrow$ "Steals" from the exposure
  $\rightsquigarrow$ Confounding by the spatial U

Hodges and Reich (2010); Paciorek (2010); Hughes and Haran (2013); Hanks et al. (2015); Vicente et al. (2020); Azevedo et al. (2020); Reich et al. (2020)

Unmeasured spatial confounding in causal inference

The goal: Learn the relationship $Y \mid Z, \boldsymbol{L}$

where $\boldsymbol{L}$ such that $Z \perp\!\!\!\perp Y(z) \mid \boldsymbol{L}$

Here: $\boldsymbol{L} = (\boldsymbol{C}, U)$, for U spatial

The goal: Learn the relationship $Y \mid Z, \boldsymbol{L}$

where $\boldsymbol{L}$ such that $Z \perp\!\!\!\perp Y(z) \mid \boldsymbol{L}$

Here: $\boldsymbol{L} = (\boldsymbol{C}, U)$, for U spatial

- Spatial Model 1: $Y \sim Z + \boldsymbol{C}$
  $\rightsquigarrow$ does not include all confounders
  $\rightsquigarrow$ cannot be used to learn causal effects

Unmeasured spatial confounding in causal inference

The goal: Learn the relationship $Y \mid Z, \boldsymbol{L}$

where $\boldsymbol{L}$ such that $Z \perp\!\!\!\perp Y(z) \mid \boldsymbol{L}$

Here: $\boldsymbol{L} = (\boldsymbol{C}, U)$, for U spatial

- Spatial Model 1: $Y \sim Z + \boldsymbol{C}$
  $\rightsquigarrow$ does not include all confounders
  $\rightsquigarrow$ cannot be used to learn causal effects

- Spatial Model 2: $Y \sim Z + \boldsymbol{C} + U$
  $\rightsquigarrow$ U spatial random effect
  $\rightsquigarrow$ U cannot "learn" the unmeasured variable (Paciorek, 2010; Schnell and Papadogeorgou, 2020)
  $\rightsquigarrow$ cannot be used to learn causal effects

Spatial confounding: What does it mean?

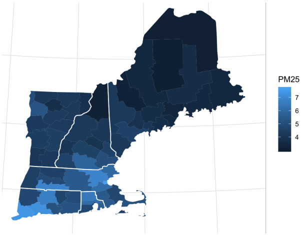
Spatial statistics

- Want to learn the relationship between outcome and exposure given the measured variables
- Spatial adjustment is for inference
- Collinearity of exposure and random effect "blurs" the results

Causal inference

- Want to learn the relationship between outcome and exposure conditioning on all confounders
- Some confounders are spatial and unmeasured
- Spatial models cannot adjust for unmeasured spatial confounders
- Can we use unmeasured confounders' *structure* to adjust for them?

- Regulations enforce stricter rules on emissions to reduce air pollution
  $\rightsquigarrow$ Power plants follow various strategies to comply to these regulations
- Install SCR/SNCR systems for reducing NO_x emissions
  $\rightsquigarrow$ NO_x reacts with VOCs and carbon monoxide in the presence of sunlight to create ozone
- The scientific questions are causal
  $\rightsquigarrow$ SCR/SNCR systems VS alternatives on ambient air pollution
- The data are spatial
  $\rightsquigarrow$ Exposure, outcome, measured and unmeasured covariates are spatially structured
  $\rightsquigarrow$ VOCs, sunlight spatial & unmeasured



Question 2: Supermarket access and cardiovascular health

- Access to supermarkets $\rightsquigarrow$ Healthy habits $\rightsquigarrow$ Cardiovascular health
- Potential confounders: Income, demographics, regional culture, personal vehicles, diet, local-level support for individuals with disabilities
- Data $\begin{cases} \text{Exposure} & \text{Continuous} \in (0, 100) \\ \text{Level} & \text{Measured at counties} \\ \text{Confounders} & \text{Unmeasured, hard to define} \end{cases}$

Two approaches to unmeasured spatial confounding

1. "Causal" approach
   Incorporating spatial information in propensity score methods
   with Christine Choirat, Cory Zigler

2. "Spatial" approach
   Bias correction in outcome regression
   with Patrick Schnell

Distance adjusted propensity score matching for binary treatments

Distance Adjusted Propensity Score Matching

- Estimate the ATT = $E[Y(1) - Y(0)|Z = 1]$
- Propensity score model using measured variables $\boldsymbol{C}$:

$$P(Z_i = 1|\boldsymbol{C}_i) = \text{expit}\left(\boldsymbol{C}_i^T \beta\right)$$

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

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

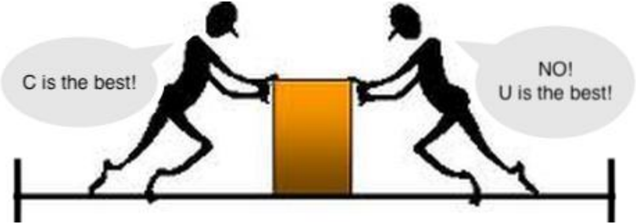
  where PS propensity score estimates, and Dist spatial proximity.

- Small value $DAPS_{ij}$ means:
  - Similar propensity scores
  - Points in close geographical distance (similar values of $\boldsymbol{U}$!)

Choosing w

w: relative importance of the observed and unobserved confounders

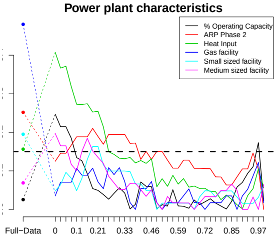
- High values of w - priority to observed covariates
- Low values of w - priority to spatial proximity

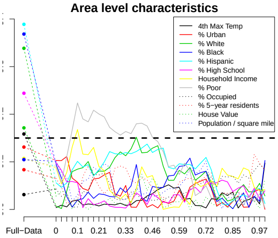


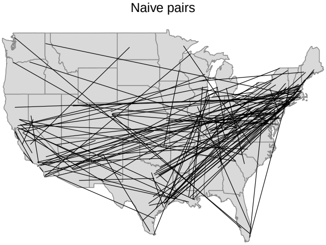
Navigate the tradeoff between:

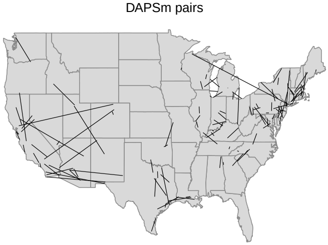
1. Making matches as similar as possible with respect to $\boldsymbol{C}$
2. Small distance of matched pairs to capture similarity in $\boldsymbol{U}$

Balance and distance

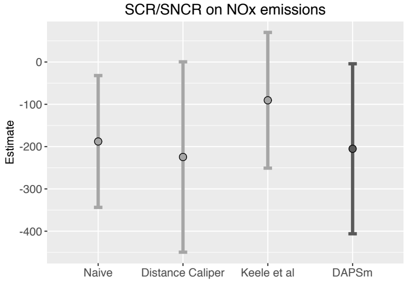


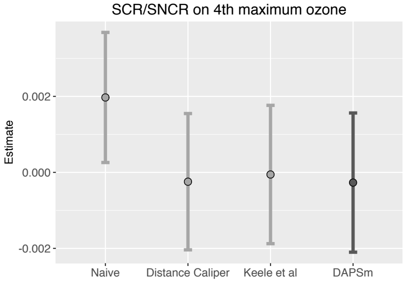






Results





Keele et al. (2015)

Outcome regression with bias correction to mitigate bias from unmeasured spatial variables

Potential outcomes

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

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

  - $\boldsymbol{W}^u$ are unmeasured variables
  - Additive model, $\mathbf{W}^u$ do not interact with Z and $\boldsymbol{C}$

Potential outcomes

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

$$Y_i(z) = \eta(z, \boldsymbol{C}) + U + \varepsilon_i(z)$$

  - $\boldsymbol{W}^u$ are unmeasured variables
  - Additive model, $\mathbf{W}^u$ do not interact with Z and $\boldsymbol{C}$
  - Denote $U = g(\boldsymbol{W}^u)$

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

$$Y_i(z) = \beta_0 + \beta_1 z + U + \varepsilon_i(z)$$

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

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.

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

    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.

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

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

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

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

A Gaussian Markov random field construction of the joint distribution

1. $(\boldsymbol{U}, \boldsymbol{Z})$ is mean 0 normal

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

2. **Cross-Markov property:** $p(Z_i | \boldsymbol{Z}_{-i}, \boldsymbol{U}) = p(Z_i | \boldsymbol{Z}_{-i}, U_i)$
   $\rightsquigarrow Q$ is diagonal

3. **Constant conditional correlation:** $\mathrm{Cor}(U_i, Z_i | \boldsymbol{U}_{-i}, \boldsymbol{Z}_{-i}) = \rho$
   $\rightsquigarrow q_{ii} = -\rho\sqrt{g_{ii} h_{ii}}$

Calculating the affine estimator

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

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

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

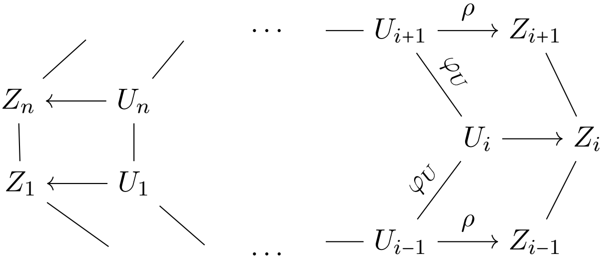
- The restricted likelihood is

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

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

- We can calculate $\bar{\boldsymbol{\beta}}$ using the RL maximizers / Bayesian

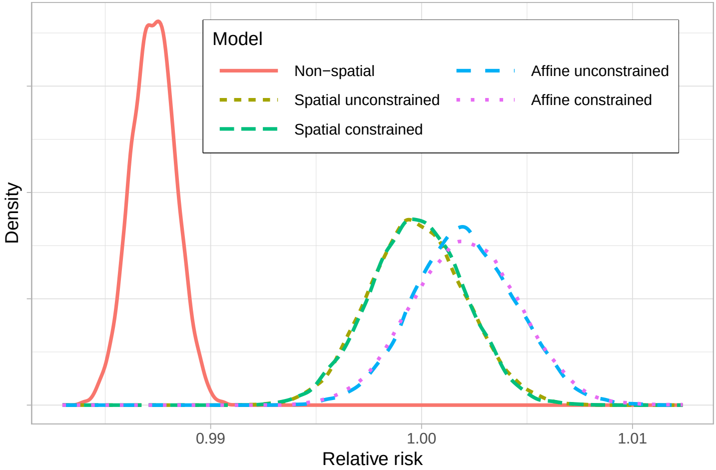
(1) the unmeasured U is (a) spatial and (b) a confounder
(2) the dependencies can be depicted on a ring
$\rightsquigarrow$ everything is identifiable based on $(\boldsymbol{Z}, \boldsymbol{Y})$



Conjecture: Similar results hold for graphs with "enough pairs" of locations at varying lags

If U not spatial: Effect is not identifiable

Effect of poor supermarket access on CVD deaths



Conclusions

- Spatial statistics VS causality: different goals, different tools
- In causal inference with structured data it is possible to use the structure information to alleviate unmeasured confounding bias
- Identifiability is hinged on structure (spatial correlation)
- Interplay between the scale of variation in the unmeasured variable and the causal positivity assumption
  $\rightsquigarrow$ scale restriction in estimation

Paper 1 — Data, R package, PDF
`gpapadogeorgou.netlify.app/publication/dapsm/`

Paper 2 — Data, Code, PDF
`gpapadogeorgou.netlify.app/publication/spatial_confounding2/`

D. R. M. Azevedo, M. O. Prates, and D. Bandyopadhyay. Alleviating spatial confounding in spatial frailty models. *arXiv:2008.06911*, 2020.

E. M. Hanks, E. M. Schliep, M. B. Hooten, and J. A. Hoeting. Restricted spatial regression in practice: geostatistical models, confounding, and robustness under model misspecification. *Environmetrics*, 26(4):243–254, 2015.

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

J. Hughes and M. Haran. Dimension reduction and alleviation of confounding for spatial generalized linear mixed models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(1):139–159, 2013.

L. Keele, R. Titiunik, and J. Zubizarreta. 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, 2015.

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

G. Papadogeorgou, C. Choirat, and C. M. Zigler. Adjusting for unmeasured spatial confounding with distance adjusted propensity score matching. *Biostatistics*, 20(2):256–272, 2019.

B. J. Reich, S. Yang, Y. Guan, A. B. Giffin, M. J. Miller, and A. G. Rappold. A review of spatial causal inference methods for environmental and epidemiological applications. *arXiv:2007.02714*, 2020.

P. Schnell and G. Papadogeorgou. Mitigating unobserved spatial confounding when estimating the effect of supermarket access on cardiovascular disease deaths. *Annals of Applied Statistics*, 14(4):2069–2095, 2020.

G. Vicente, T. Goicoa, P. Fernandez-Rasines, and M. Ugarte. Crime against women in india: unveiling spatial patterns and temporal trends of dowry deaths in the districts of uttar pradesh. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 183(2): 655–679, 2020.