

Differences-in-Differences with Spatial Spillover

Kyle Butts

December 4, 2020

Spatial Spillovers

Researchers aim to estimate the average treatment effect on the treated:

$$\tau \equiv \mathbb{E} [Y_{i1}(1) - Y_{i1}(0) \mid D_i = 1]$$

Spillover effects are when effect of treatment extend over boundaries such as states, counties, etc.

- e.g. large employer opening/closing in a county have positive or negative employment effects on nearby counties

Bias from Spatial Spillovers

The canonical difference-in-differences estimate is:

$$\hat{\tau} = \underbrace{\hat{\mathbb{E}} [Y_{i1} - Y_{i0} \mid D_i = 1]}_{\text{Counterfactual Trend} + \tau} - \underbrace{\hat{\mathbb{E}} [Y_{i1} - Y_{i0} \mid D_i = 0]}_{\text{Counterfactual Trend}}$$

Two problems in presence of spillover effects:

- **Spillover onto Control Units:**

Nearby “control” units fail to estimate counterfactual trends because they are affected by treatment

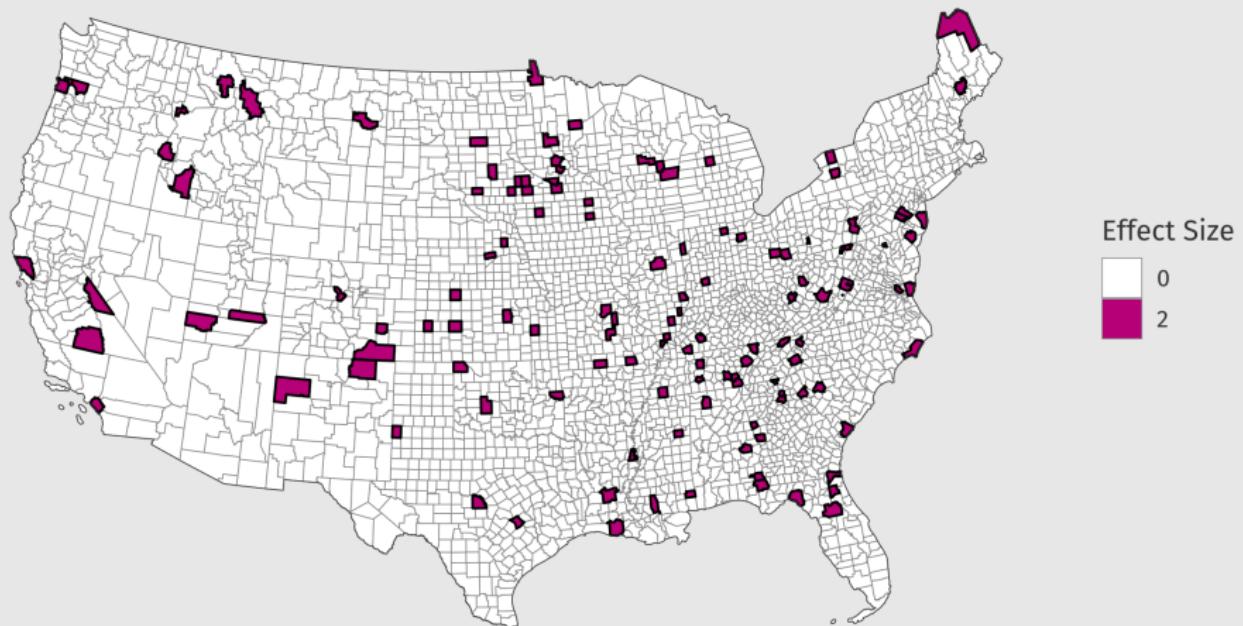
- **Spillover onto other Treated Units:**

Treated units are also affected by nearby units and therefore combines “direct” effects with spillover effects

Direct Effect of Treatment

$Pr(\text{treatment}) = 5\%;$ Bias = 0

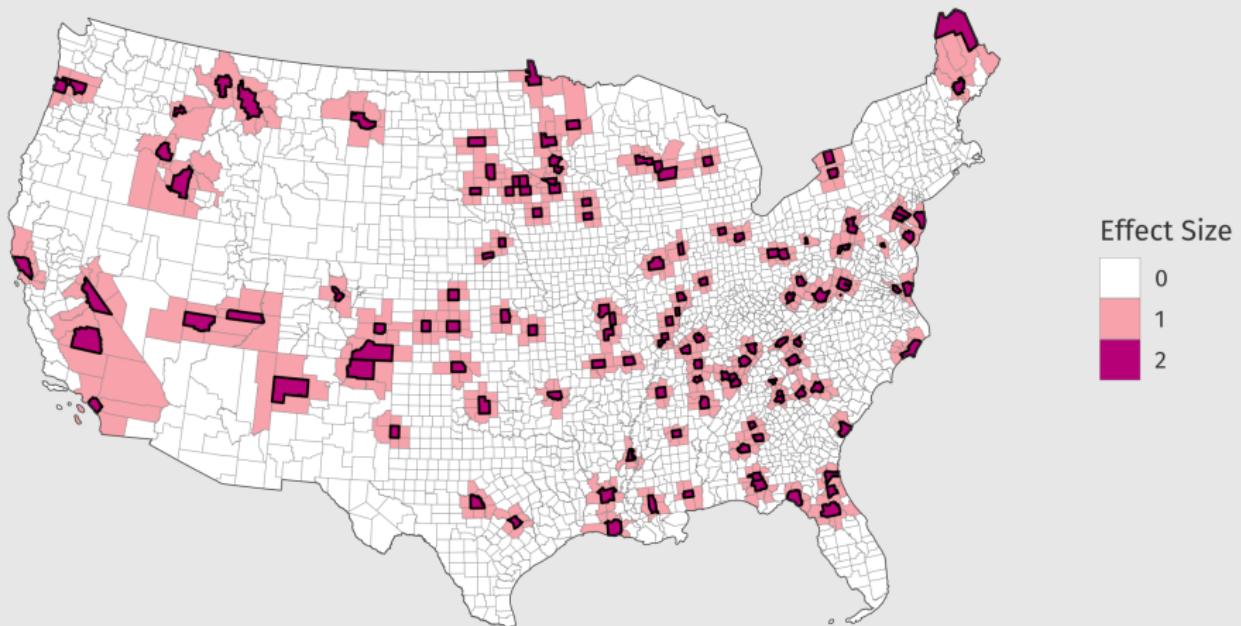
Treated units outlined in black



Direct Effect + Spillover on Control

$Pr(\text{treatment}) = 5\%$; $\text{Bias} = -0.13$

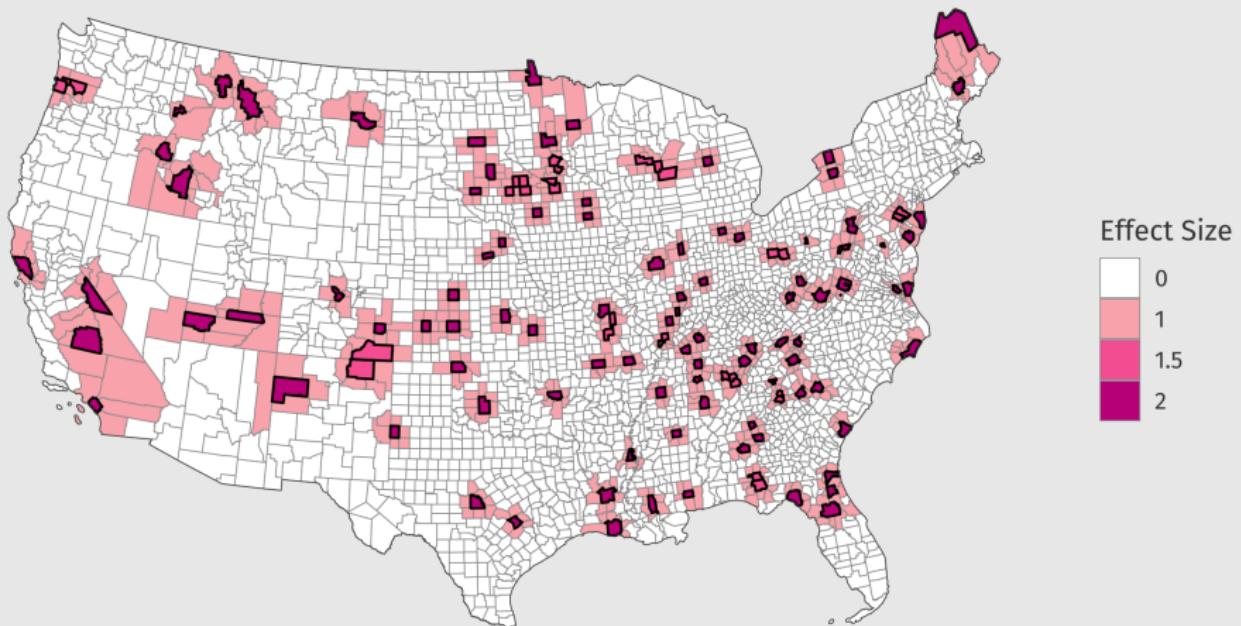
Treated units outlined in black



Direct Effect + Spillover on Control and Treated

$Pr(\text{treatment}) = 0.05$; Bias = $-0.09 - 0.13 = -0.22$

Treated units outlined in black



Outline

- 1– Using a potential outcomes framework, I derive the bias in the presence of spillovers
- 2– Using Monte Carlo simulations, show when biases are large and analyze effectiveness of current adjustments done by applied researchers
- 3– Revisit Kline and Moretti (2014) analysis of the Tennessee Valley Authority

Preview of Results

There are two sources of bias:

1. Treatment effects spillover onto control units
2. Treatment effects spillover onto also treated units

Solutions:

- Fixing bias 1 is relatively easy and works better when you have relatively few treated units
- Bias 2 worsens as treatment locations bunch more (spatial correlation of treatment increases)
- However, spatial correlation of treatment decreases bias 1

Roadmap of Talk

Theory

Simulations

Application: Tennessee Valley Authority

Conclusion

Potential Outcomes Framework

For exposition, I will label units as counties. Assume all treatment occurs at the same time (2-periods or pre-post averages).¹

- Let $Y_{it}(D_i, h_i(\vec{D}))$ be the potential outcome of county $i \in \{1, \dots, N\}$ at time t with treatment status $D_i \in \{0, 1\}$.
- $\vec{D} \in \{0, 1\}^N$ is the vector of all units treatments.
- The function $h_i(\vec{D})$ maps the entire treatment vector into a scalar that completely determines the spillover.
- In the world with no treatment, $\vec{0}$, we say $h_i(\vec{0}) = 0$.

¹This form adapts the framework of Vazquez-Bare (2019) for the diff-in-diff setting.

Examples of $h_i(\vec{D})$

Examples of $h_i(\vec{D})$:

- **Treatment within x miles:**

$h_i(\vec{D})$ equals 1 if there is a treated unit within x -miles of unit i and equals 0 otherwise

- **Average treatment among k -nearest neighbors:**

$h_i(\vec{D}) = \sum_{j=1}^k D_{n(i,j)} / k$ where $n(i, j)$ represents the j -th closest neighbor to unit i .

Estimand of Interest

Estimand of Interest:

$$\tau_{\text{direct}} \equiv \mathbb{E} [Y_{i,1}(1, 0) - Y_{i,1}(0, 0) \mid D_i = 1]$$

This is the direct effect in the absence of spillovers.

Parallel Trends

I assume a modified version of the parallel trends assumption:

$$\begin{aligned} & \mathbb{E} \left[Y_{i,1}(0, 0) - Y_{i,0}(0, 0) \mid D_i = 1, \vec{D} = 0 \right] \\ &= \mathbb{E} \left[Y_{i,1}(0, 0) - Y_{i,0}(0, 0) \mid D_i = 0, \vec{D} = 0 \right], \end{aligned}$$

In the complete absence of treatment (not just the absence of individual i 's treatment), the change in potential outcomes from period 0 to 1 would not depend on treatment status

What does Diff-in-Diff identify?

With the parallel trends assumption and random assignment of D_i , I decompose the diff-in-diff estimate as follows:

$$\begin{aligned} & \mathbb{E}[Y_{i1} - Y_{i0} \mid D_i = 1] - \mathbb{E}[Y_{i1} - Y_{i0} \mid D_i = 0] \\ &= \mathbb{E}[Y_{i1}(1, 0) - Y_{i1}(0, 0) \mid D_i = 1] \\ &\quad + \mathbb{E}\left[Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(1, 0) \mid D_i = 1\right] \\ &\quad - \mathbb{E}\left[Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0, 0) \mid D_i = 0\right] \\ &\equiv \tau_{\text{direct}} + \tau_{\text{spillover, treated}} - \tau_{\text{spillover, control}} \end{aligned}$$

Signing the Bias

$$\tau_{\text{direct}} + \tau_{\text{spillover, treated}} - \tau_{\text{spillover, control}}$$

Table 1: Bias from Spillovers

	+ Spillover	- Spillover
onto Treated	+ Bias	- Bias
onto Control	- Bias	+ Bias

Example of Bias

Treated Spillover:

$Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(1, 0) = -0.5$ if county i shares a border with a treated county and 0 otherwise.

Control Spillover:

$Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0, 0) \mid D_i = 1$ if county i shares a border with a treated county and 0 otherwise.

Then, the bias term is

- $0.5 * \%$ of treated counties sharing a border with a treated county
- $1 * \%$ of treated counties sharing a border with a treated county

Partial Identification

Researchers can use the above method to create bounds for the biases resulting from spillovers. The choices should follow from the empirical context at hand.

To formalize this, I follow the literature on partial identification for treatment effects:

[Armstrong and Kolesár (2020), Rambachan and Roth (2020), and Manski and Pepper (2018)]

Partially-identified Set

Following Rambachan and Roth (2020) and Manski and Pepper (2018), researchers can provide bounds to the maximum spillovers.

Let Δ denote the interval of potential biases:

$$\Delta \equiv [\underline{\Delta}, \bar{\Delta}]$$

$$= \left[\min \tau_{\text{spill,treated}} - \tau_{\text{spill,control}}, \max \tau_{\text{spill,treated}} - \tau_{\text{spill,control}} \right]$$

The partially identified set is formed by

$$\left[\hat{\tau} - \underline{\Delta}, \hat{\tau} + \bar{\Delta} \right]$$

Bias-adjusted Inference

Inference can be done following Rambachan and Roth (2020) by creating a Fixed-Length Confidence Interval:

$$\left[\hat{\tau} - \underline{\Delta} - z_{1-\alpha/2} \text{SE}_{\hat{\tau}}, \hat{\tau} + \bar{\Delta} + z_{1-\alpha/2} \text{SE}_{\hat{\tau}} \right]$$

These standard errors should account for the spatial structure of outcomes following Conley (1999).

Roadmap of Talk

Theory

Simulations

Application: Tennessee Valley Authority

Conclusion

Monte Carlo Simulations

- Simulated the following Data Generating Process

$$y_{it} = -2 + \mu_t + \mu_i + \tau_{\text{direct}} D_{it} + \tau_{\text{spillover, control}} (1 - D_{it}) \text{Near}_{it} \\ + \tau_{\text{spillover, treated}} D_{it} \text{Near}_{it} + \varepsilon_{it}$$

- Observation i is a U.S. county, year $t \in \{2000, \dots, 2019\}$, treatment D_{it} turns on in 2010 and is assigned randomly or through a spatial process.
- Near_{it} is an indicator for having your county centroid be within 40 miles of a treated unit.
- $\tau_{\text{direct}} = 2$, $\tau_{\text{spillover, control}} = 1$ and $\tau_{\text{spillover, treat}} = -0.5$.

Source of Bias 1: Control Units

Simulation 1:

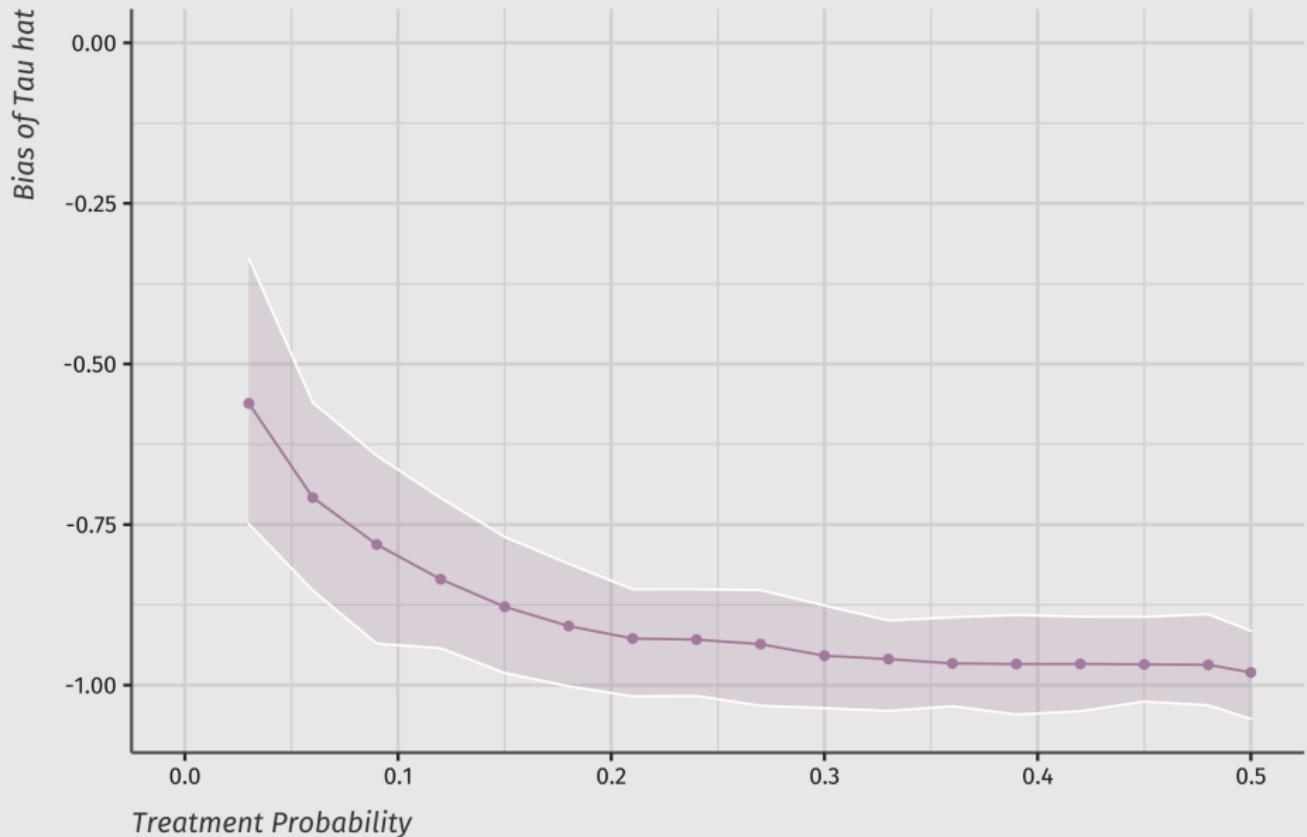
- The first simulation changes the probability of treatment, D_{it} .
- Assume, for now, there is only spillover on the control, i.e.
 $\tau_{\text{spillover,treat}} = 0$.
- Estimate the following equation:

$$y_{it} = \alpha + \mu_t + \mu_i + \tau D_{it} + \epsilon_{it}$$

- The bias of our estimate is $Bias = \hat{\tau} - 2$

As more units are treated, the bias increases

Monte Carlo simulations with 100 trials per simulation

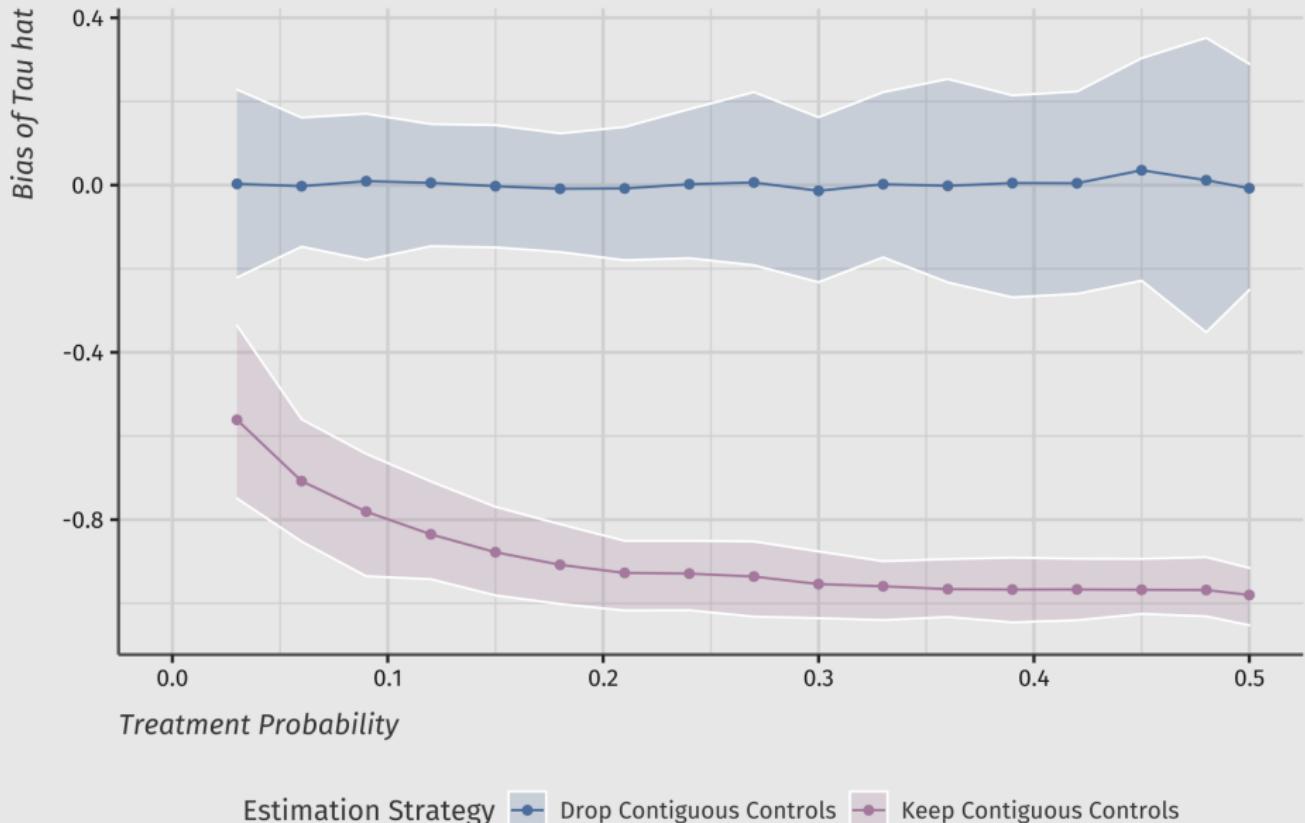


Solution: Removing “contaminated” controls

A common solution to the problem of spillover is to restimate without neighboring control units and see what happens to the bias. This simulation assumes the correct “contaminated” control units, i.e. only control units with $h_i(\vec{D}) \neq 0$, which is not possible for researchers.

Dropping control units removes bias, but increases variance

Monte Carlo simulations with 100 trials per simulation

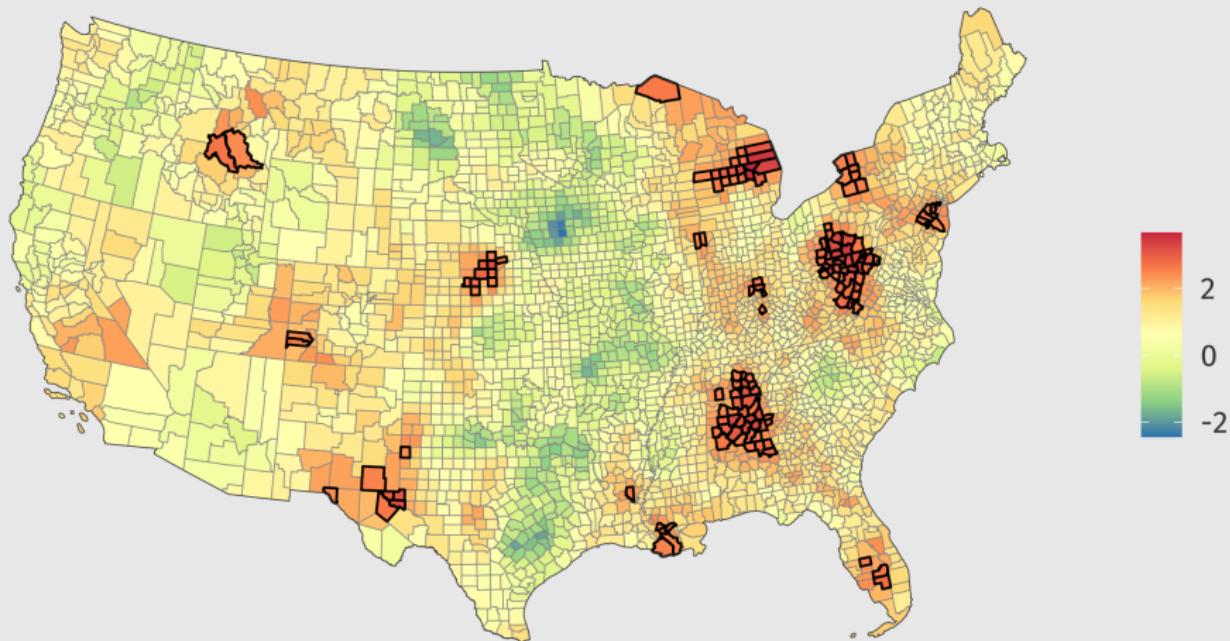


Source of Bias 2: Treated Units

Simulation 2:

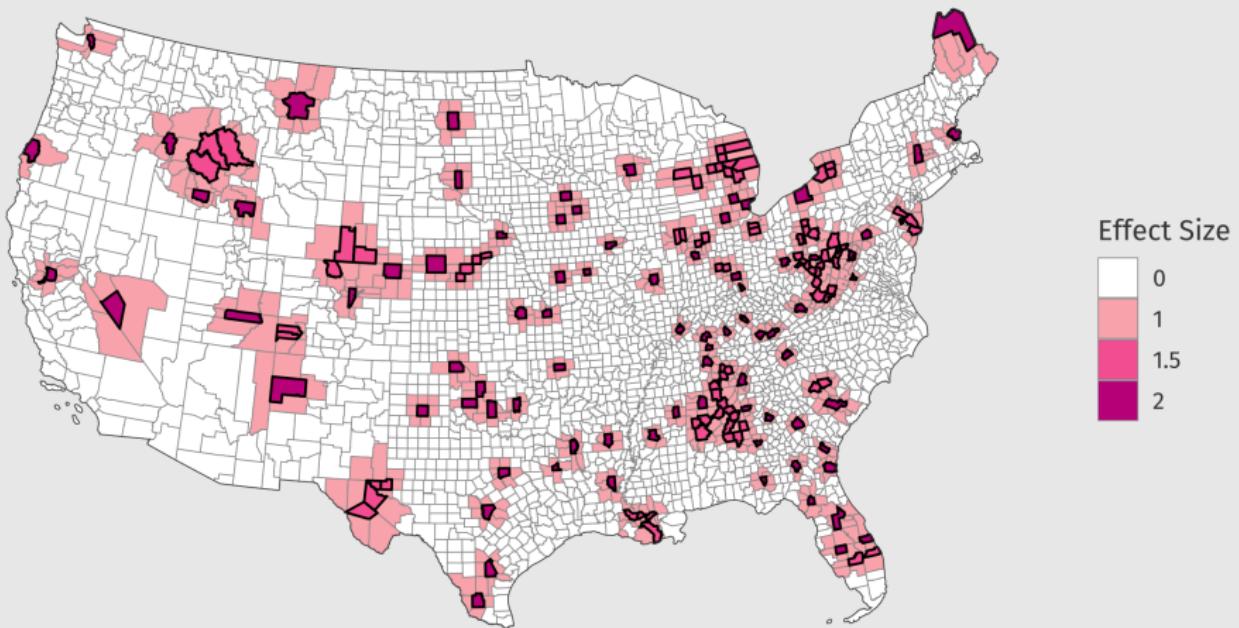
- I now add in the bias from spillover of treated units on to other treated units.
- This source of bias only occurs when treated units are located next to each other, so the magnitude of bias depends on spatial autocorrelation of D_{it} .
- Using a method from geosciences, I generate correlated “fields” across the US.
- Then, for the counties in the top 10% of values (hot spots in the field), I increase the probability of treatment.
- The unconditional probability remains fixed at 10%.

Kriging Example



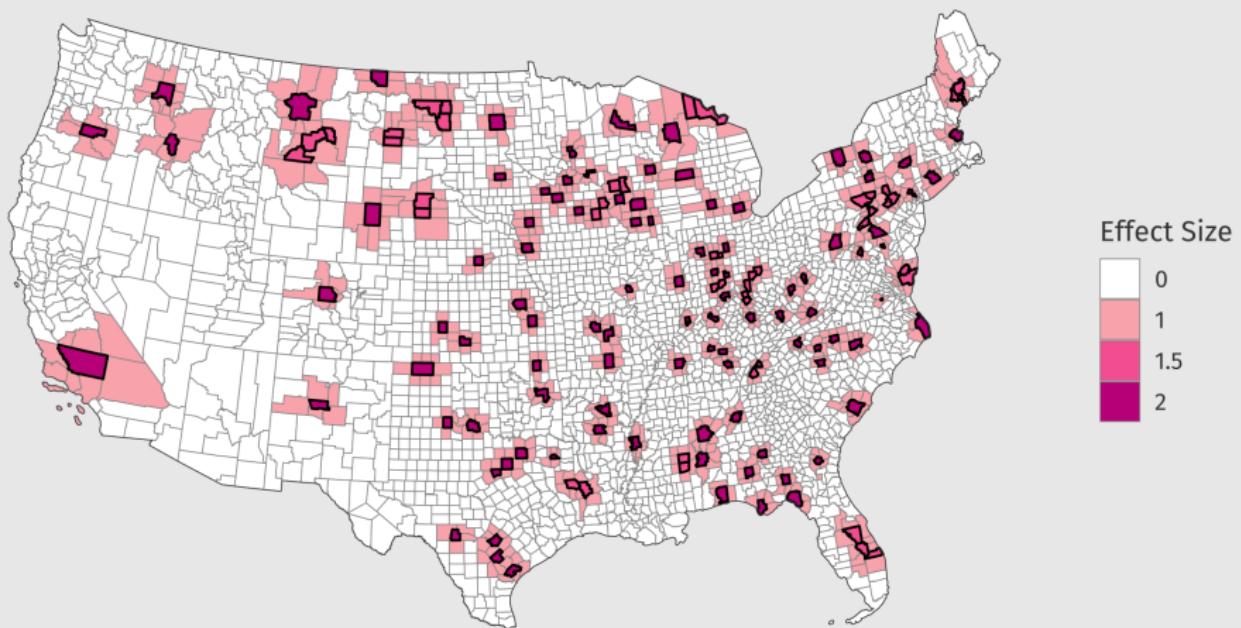
Large Spatial Autocorrelation

$\Pr(\text{treatment}) = 5\%;$ Bias = $-0.25 - 0.24 = -0.49$
Spatial Autocorrelation Measure = 1.4



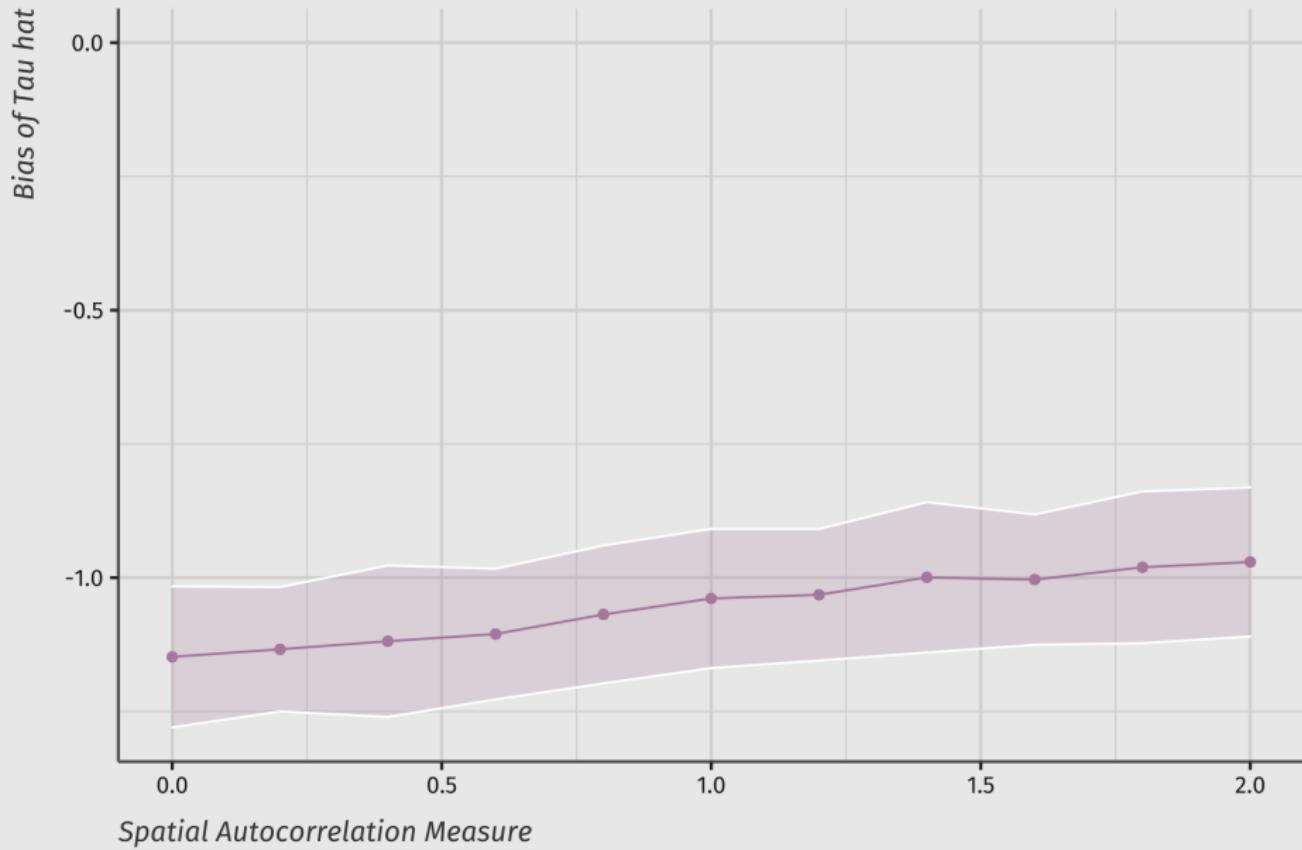
Small Spatial Autocorrelation

$Pr(\text{treatment}) = 5\%;$ Bias = $-0.15 - 0.25 = -0.4$
Spatial Autocorrelation Measure = 0.4



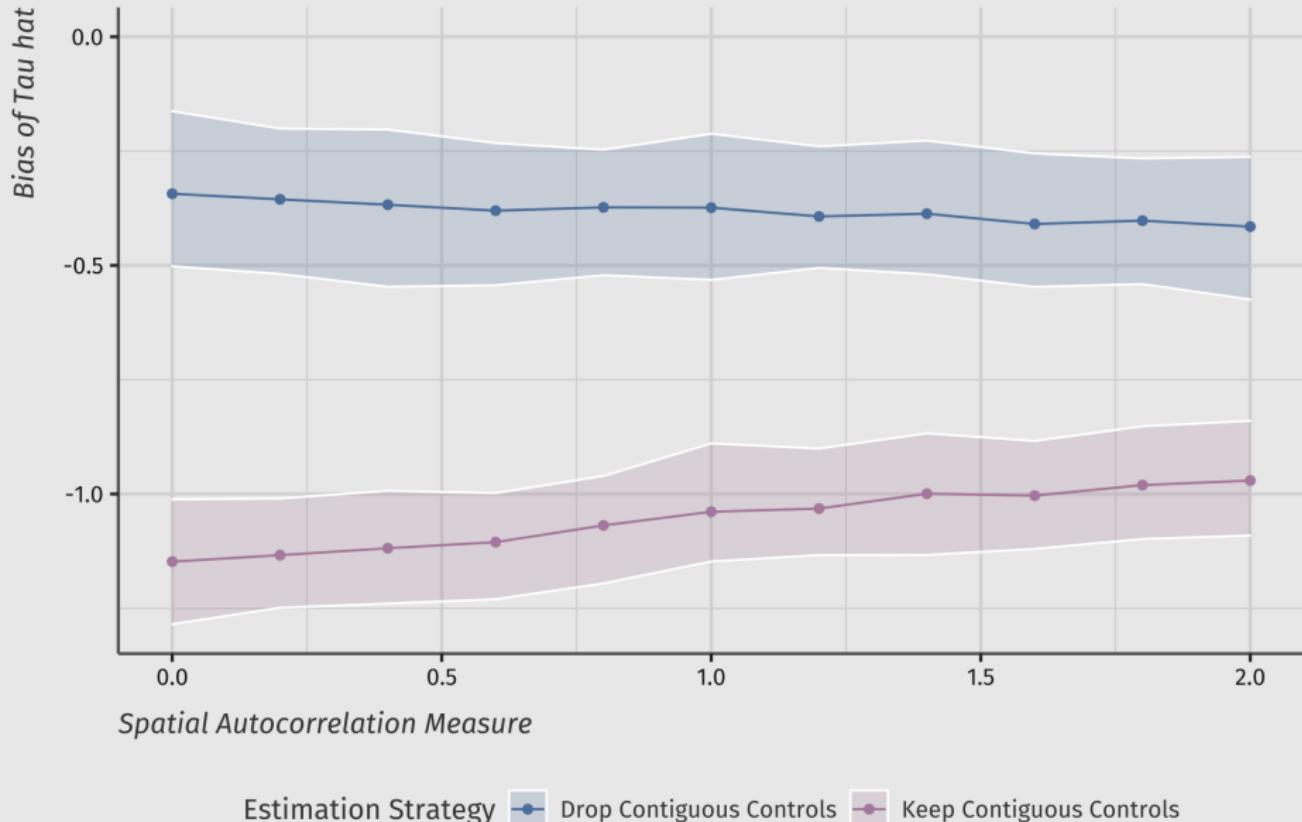
Bias as spatial correlation changes

Monte Carlo simulations with 100 trials for each simulation



Dropping control units no longer effectively removes all bias

Monte Carlo simulations with 100 trials per simulation



Spillovers as estimand of interest

Until now, we assumed our estimand of interest is τ_{direct} .

However, the two other spillover effects are of interest as well:

- $\tau_{\text{spillover, control}}$: Do the benefits of a treated county come at a cost to neighbor counties?
- $\tau_{\text{spillover, treated}}$: Does the estimated effect change based on others treatment? (This is what you should consider if you are a policy maker)

To estimate the spillover effects, we have to parameterize $h_i(\vec{D})$ function and the potential outcomes function $Y_i(D_i, h_i(\vec{D}))$.

Roadmap of Talk

Theory

Simulations

Application: Tennessee Valley Authority

Conclusion

Tennessee Valley Authority

Kline and Moretti (2014) look at the long-run impacts of the Tennessee Valley Authority (TVA).

- The TVA was a large-scale federal investment started in 1934 that focused on improving manufacturing economy.
- The program focused on electrification through dams and transportation canals in order to improve the manufacturing economy
- Hundreds of dollars spent annually per person

Identification

Kline and Moretti (2014) run the county-level difference-in-differences specification:

$$y_{c,2000} - y_{c,1940} = \alpha + \text{TVA}_c \tau + X_{c,1940} \beta + (\varepsilon_{c,2000} - \varepsilon_{c,1940}) \quad (1)$$

- y are outcomes for agricultural employment, manufacturing employment, and median family income.
- TVA_c is the treatment variable
- $X_{c,1940}$ allow for different long-term trends based on covariates in 1940.

They trim the sample using a logit regression to predict treatment using $X_{c,1940}$ and then keep control units in the top 75% of predicted probability.

Spillovers in the TVA Context

In our context, there is reason to believe spillovers can occur to nearby counties

- Agricultural employees might be drawn to hire wages for new manufacturing jobs in Tennessee Valley
- Manufacturing jobs that would have been created in the control units in the absence of treatment might move to the Tennessee Valley
- Since manufacutring jobs pay higher than agricultural jobs, the above two spillover effects can cause changes in median family income

Specification including spillovers

$$\Delta y_c = \alpha + \text{TVA}_i \tau + \sum_{d \in \text{Dist}} \text{Between}(d) \delta_d + X_{i,1940} \beta + \Delta \varepsilon_c \quad (2)$$

- $\text{Between}(d)$ is a set of indicators for being in the following distance bins (in miles) from the Tennessee Valley Authority:

$$d \in \{(0, 50], (50, 100], (100, 150]\}$$

Effective Sample and Spillover Variables

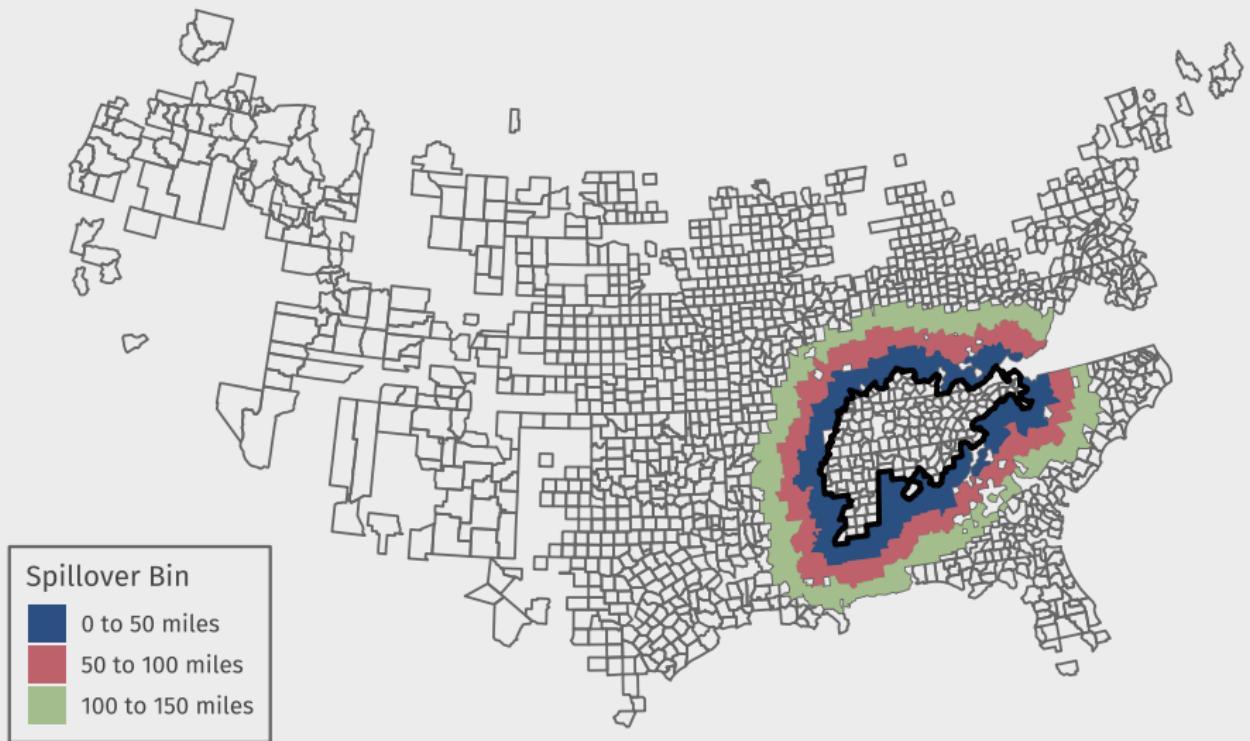


Table 2: Effects of Tennessee Valley Authority on Decadal Growth

Dependent Var.	Diff-in-Diff		Diff-in-Diff with Spillovers		
	TVA	TVA	TVA between 0-50 mi.	TVA between 50-100 mi.	TVA between 100-150 mi.
Agricultural employment	-0.0514*** (0.0087)	-0.0678*** (0.0102)	-0.0310** (0.0123)	-0.0112 (0.0094)	-0.0252*** (0.0084)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 2: Effects of Tennessee Valley Authority on Decadal Growth

Dependent Var.	Diff-in-Diff		Diff-in-Diff with Spillovers		
	TVA	TVA	TVA between 0-50 mi.	TVA between 50-100 mi.	TVA between 100-150 mi.
Agricultural employment	-0.0514*** (0.0087)	-0.0678*** (0.0102)	-0.0310** (0.0123)	-0.0112 (0.0094)	-0.0252*** (0.0084)
Manufacturing employment	0.0560*** (0.0187)	0.0461** (0.0210)	-0.0104 (0.0205)	-0.0128 (0.0257)	-0.0248* (0.0147)

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$.

Table 2: Effects of Tennessee Valley Authority on Decadal Growth

Dependent Var.	Diff-in-Diff		Diff-in-Diff with Spillovers		
	TVA	TVA	TVA between 0-50 mi.	TVA between 50-100 mi.	TVA between 100-150 mi.
Agricultural employment	-0.0514*** (0.0087)	-0.0678*** (0.0102)	-0.0310** (0.0123)	-0.0112 (0.0094)	-0.0252*** (0.0084)
Manufacturing employment	0.0560*** (0.0187)	0.0461** (0.0210)	-0.0104 (0.0205)	-0.0128 (0.0257)	-0.0248* (0.0147)
Median family income	0.0191*** (0.0067)	0.0196** (0.0077)	0.0056 (0.0079)	-0.0062 (0.0060)	-0.0011 (0.0032)

^{*} $p < 0.1$; ^{**} $p < 0.05$; ^{***} $p < 0.01$.

Roadmap of Talk

Theory

Simulations

Application: Tennessee Valley Authority

Conclusion

To-do:

- Include Monte-Carlo simulations on misspecification of spillovers
- Extend bias to settings with differential timing of treatment

Thank you!