Difference-in-Differences with Spatial Spillovers

Kyle Butts

March 12, 2021

Spatial Spillovers

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

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

Estimation is complicated by **Spillover Effects**

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

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

This Paper

In this paper, I

- Present a potential outcomes framework to formalize spillover effects and evaluate ad-hoc adjustments done in the literature
- Propose an estimation strategy that improves on current practices by being more robust to spillovers
- Apply this framework to improve estimation of the local effect of place-based policies in Urban Economics

Bias from Spatial Spillovers

The canonical difference-in-differences estimate is:

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

Two problems in presence of spillover effects:

Bias from Spatial Spillovers

The canonical difference-in-differences estimate is:

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

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

Bias from Spatial Spillovers

The canonical difference-in-differences estimate is:

$$\hat{\tau} = \underbrace{\hat{\mathbb{E}}\left[Y_{i1} - Y_{i0} \mid D_i = 1\right]}_{\substack{\text{Counterfactual Trend } + \ \tau \\ + \text{ Spillover on Treated}}} - \underbrace{\hat{\mathbb{E}}\left[Y_{i1} - Y_{i0} \mid D_i = 0\right]}_{\substack{\text{Counterfactual Trend} \\ + \text{ Spillover on Control}}}$$

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

Outline

1– Formalize spillovers into a potential outcomes framework:

[Clarke (2017), Berg and Streitz (2019), and Verbitsky-Savitz and Raudenbush (2012)]

- I decompose the difference-in-differences estimator into three parts: Direct Effect of Treatment, Spillover onto Treated Units, Spillover onto Control Units
- Show that an indicator for being close to treated units remove *all bias* so long as the indicator contains all units affected by spillovers
- 'Rings' are able to estimate spillover effects while still removing all bias

Outline

2– Apply framework to Urban Economics

- Revisit Kline and Moretti (2014a) analysis of the Tennessee Valley Authority
    - The local effect estimate is contaminated by spillover effects to neighboring counties (Kline and Moretti, 2014b)
    - Large scale manufacturing investment creates an ‘urban shadow’ (Cuberes, Desmet, and Rappaport, 2021; Fujita, Krugman, and Venables, 2001)
- Discuss how framework can reconcile conflicting findings on effect of federal Empowerment Zones (Busso, Gregory, and Kline, 2013; Neumark and Kolko, 2010)

Roadmap of Talk

Theory

Estimation with Spillovers

Application in Urban Economics

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]

- $Y_{it}(D_i, h(\vec{D}, i))$ is the potential outcome of county $i \in \{1, \ldots, N\}$ 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 $h(\vec{D}, i)$ maps the entire treatment vector into an 'exposure mapping' which can be a scalar or a vector.
- No exposure is when $h(\vec{D}, i) = \vec{0}$.

[1] I extend this into an event study framework in the paper, but the intuition is the same as in the 2×2 setting.

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]

- $Y_{it}(D_i, h(\vec{D}, i))$ is the potential outcome of county $i \in \{1, \ldots, N\}$ 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 $h(\vec{D}, i)$ maps the entire treatment vector into an 'exposure mapping' which can be a scalar or a vector.
- No exposure is when $h(\vec{D}, i) = \vec{0}$.

[1] I extend this into an event study framework in the paper, but the intuition is the same as in the 2×2 setting.

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]

- $Y_{it}(D_i, h(\vec{D}, i))$ is the potential outcome of county $i \in \{1, \ldots, N\}$ 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 $h(\vec{D}, i)$ maps the entire treatment vector into an 'exposure mapping' which can be a scalar or a vector.
- No exposure is when $h(\vec{D}, i) = \vec{0}$.

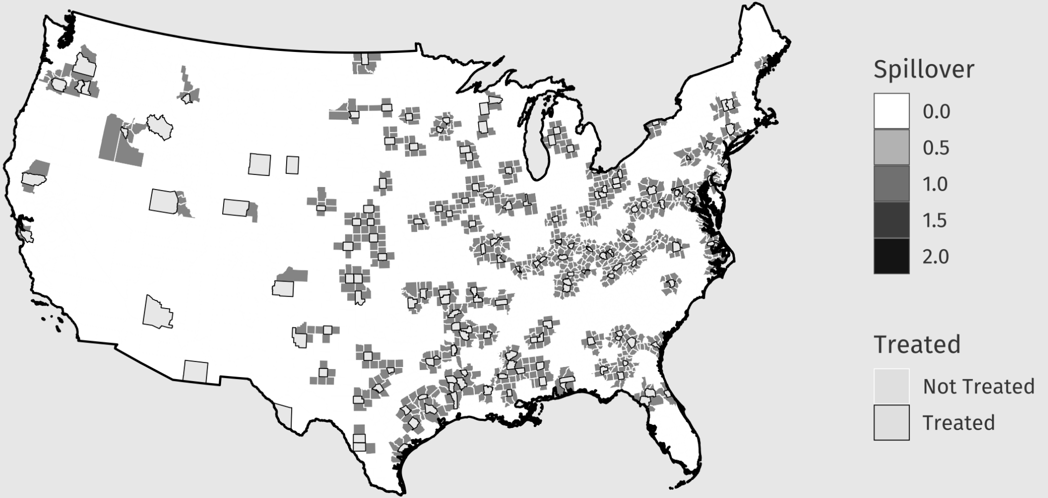
[1] I extend this into an event study framework in the paper, but the intuition is the same as in the 2×2 setting.

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

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

- **Treatment within x miles:**
  $h(\vec{D}, i) = max_j \ 1(d(i,j) \leq x)$ where $d(i,j)$ is the distance between counties i and j.
  - e.g. library access where x is the maximum distance people will travel
  - Spillovers are non-additive

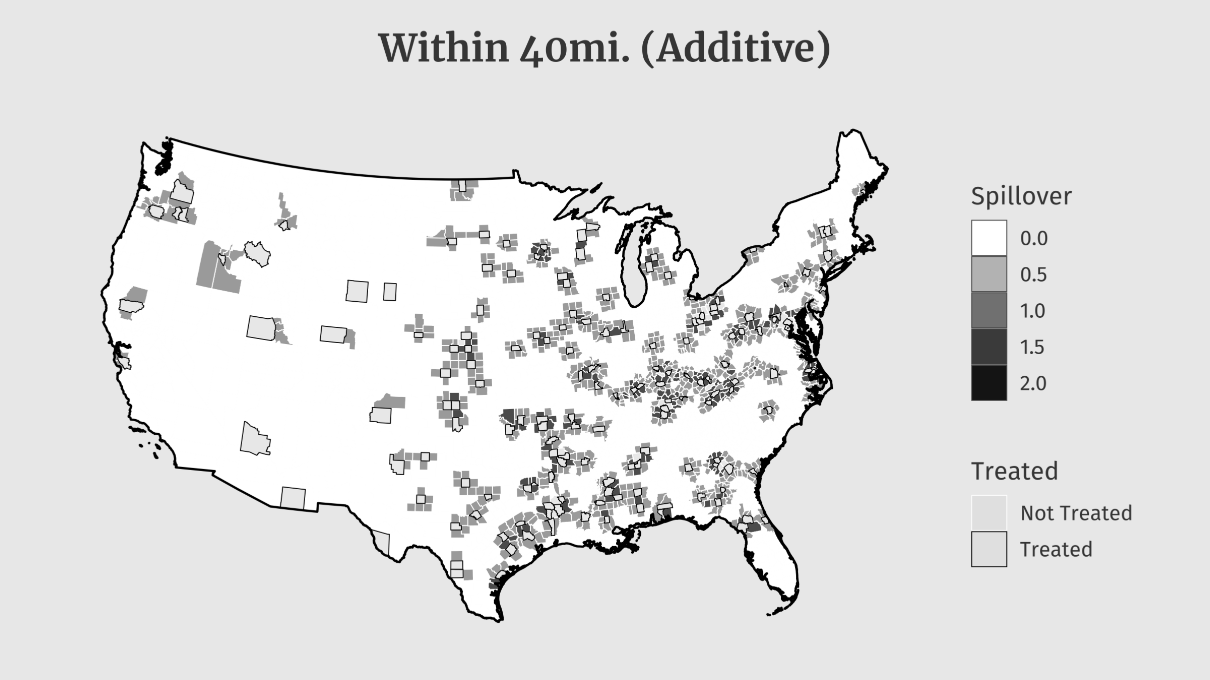
Within 40mi.



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

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

- **Treatment within x miles:**
  $h(\vec{D}, i) = max_j \ 1(d(i,j) \leq x)$ where $d(i,j)$ is the distance between counties i and j.
  - e.g. library access where x is the maximum distance people will travel
  - Spillovers are non-additive
- **Number of Treated within x miles:**
  $h(\vec{D}, i) = \sum_{j=1}^{k} 1(d(i,j) \leq x)$.
  - e.g. large factories opening
  - Agglomeration economies suggest spillovers are additive


Within 40mi. (Additive)
Spillover
0.0
0.5
1.0
1.5
2.0
Treated
Not Treated
Treated

Estimand of Interest:

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

Spillover Effects:

$$\tau_{\text{spillover, treated}} \equiv \mathbb{E}\left[Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(1,0) \mid D_i = 1\right]$$

$$\tau_{\text{spillover, control}} \equiv \mathbb{E}\left[Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0,0) \mid D_i = 0\right]$$

Estimand of Interest

Estimand of Interest:

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

Spillover Effects:

$$\tau_{\text{spillover, treated}} \equiv \mathbb{E}\left[Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(1,0) \mid D_i = 1\right]$$

$$\tau_{\text{spillover, control}} = \mathbb{E}\left[Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0,0) \mid D_i = 0\right]$$

Parallel Trends

I assume a modified version of the parallel counterfactual trends assumption:

$$\mathbb{E}\big[\underbrace{Y_{i,1}(0,\vec{0}) - Y_{i,0}(0,\vec{0})}_{\text{Counterfactual Trend}} \mid D_i = 1\big]$$
$$= \mathbb{E}\big[\underbrace{Y_{i,1}(0,\vec{0}) - Y_{i,0}(0,\vec{0})}_{\text{Counterfactual Trend}} \mid D_i = 0\big],$$

In the *complete absence of treatment* (not just the absence of individual i's treatment):

Changes in outcomes does not depend on treatment status

What does Difference-in-Differences identify?

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

$$\mathbb{E}\left[\hat{\tau}\right] = \underbrace{\mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 1\right] - \mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 0\right]}_{\text{Difference-in-Differences}}$$

$$\begin{aligned} &= \mathbb{E}\left[Y_{i1}(1,0) - Y_{i1}(0,0) \mid D_i = 1\right] \\ &\quad + \quad \mathbb{E}\left[Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(1,0) \mid D_i = 1\right] \\ &\quad - \quad \mathbb{E}\left[Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0,0) \mid D_i = 0\right] \end{aligned}$$

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

What does Difference-in-Differences identify?

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

$$\mathbb{E}\left[\hat{\tau}\right] = \underbrace{\mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 1\right] - \mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 0\right]}_{\text{Difference-in-Differences}}$$

$$\begin{aligned} &= \mathbb{E}\left[Y_{i1}(1,0) - Y_{i1}(0,0) \mid D_i = 1\right] \\ &\quad + \quad \mathbb{E}\left[Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(1,0) \mid D_i = 1\right] \\ &\quad - \quad \mathbb{E}\left[Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0,0) \mid D_i = 0\right] \end{aligned}$$

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

Example of Bias: Large Factory Opening

Control Spillover:

$Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0, 0) = 1$ if county i is within 40 miles of a treated county and 0 otherwise.

Treated Spillover:

$Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(1, 0) = 0.5$ if county i is within 40 miles of a treated county and 0 otherwise.

Then, the bias term is

$$0.5 * \% \text{ of treated counties within 40 miles of another treated county}$$
$$- 1 * \% \text{ of treated counties within 40 miles of a treated county}$$

Partial Identification

Example of Bias: Large Factory Opening

Control Spillover:

$Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0, 0) = 1$ if county i is within 40 miles of a treated county and 0 otherwise.

Treated Spillover:

$Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(1, 0) = 0.5$ if county i is within 40 miles of a treated county and 0 otherwise.

Then, the bias term is

$$0.5 * \% \text{ of treated counties within 40 miles of another treated county}$$
$$- 1 * \% \text{ of treated counties within 40 miles of a treated county}$$

Partial Identification

Roadmap of Talk

Theory

Estimation with Spillovers

Application in Urban Economics

Conclusion

Monte Carlo Simulations

- Data Generating Process

$$y_{it} = \mu_t + \mu_i + \tau_{\text{direct}} D_{it} + \tau_{\text{spillover, control}} (1 - D_{it}) \text{Within 40mi.}_{it}$$
$$+ \tau_{\text{spillover, treated}} D_{it} \text{Within 40mi.}_{it} + \varepsilon_{it}$$

- Observation i is a U.S. county, year $t \in \{2000, \ldots, 2019\}$, treatment D_{it} turns on in 2010 and is assigned randomly or through a spatial process.
- Within 40mi._{it} is an indicator for having your county center of population be within 40 miles of a treated unit.
- $\tau_{\text{direct}} = 2$, $\tau_{\text{spillover, control}} = 1$ and $\tau_{\text{spillover, treat}} \in \{0, 0.5\}$.

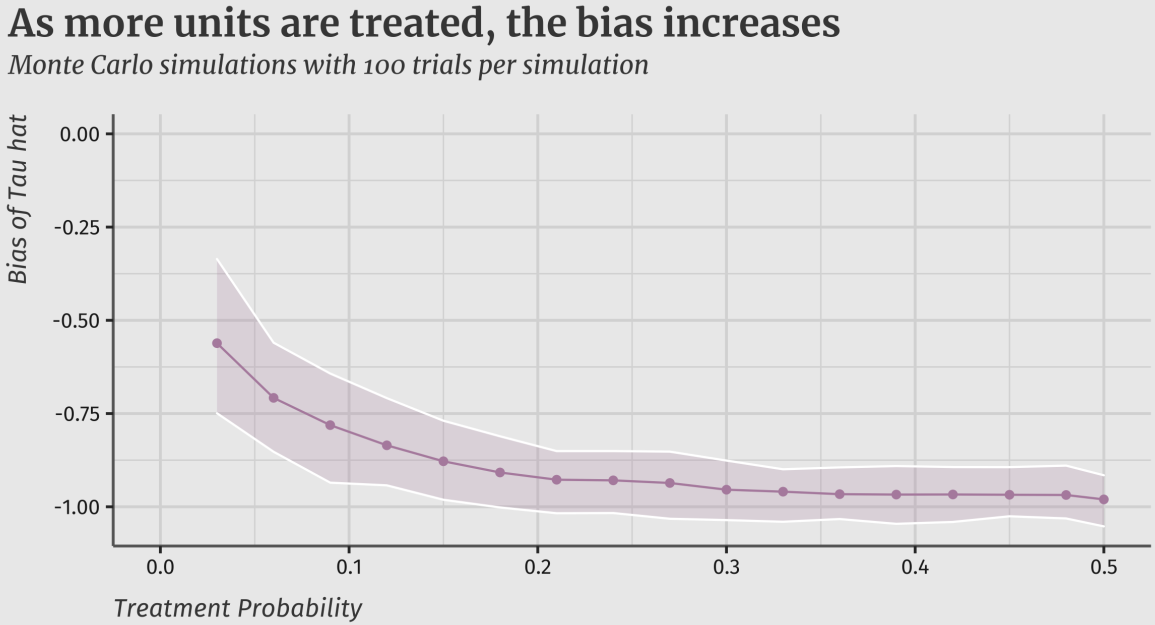
Source of Bias 1: Control Units

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 via TWFE:

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

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


As more units are treated, the bias increases
Monte Carlo simulations with 100 trials per simulation
Bias of Tau hat
0.00
-0.25
-0.50
-0.75
-1.00
0.0
0.1
0.2
0.3
0.4
0.5
Treatment Probability

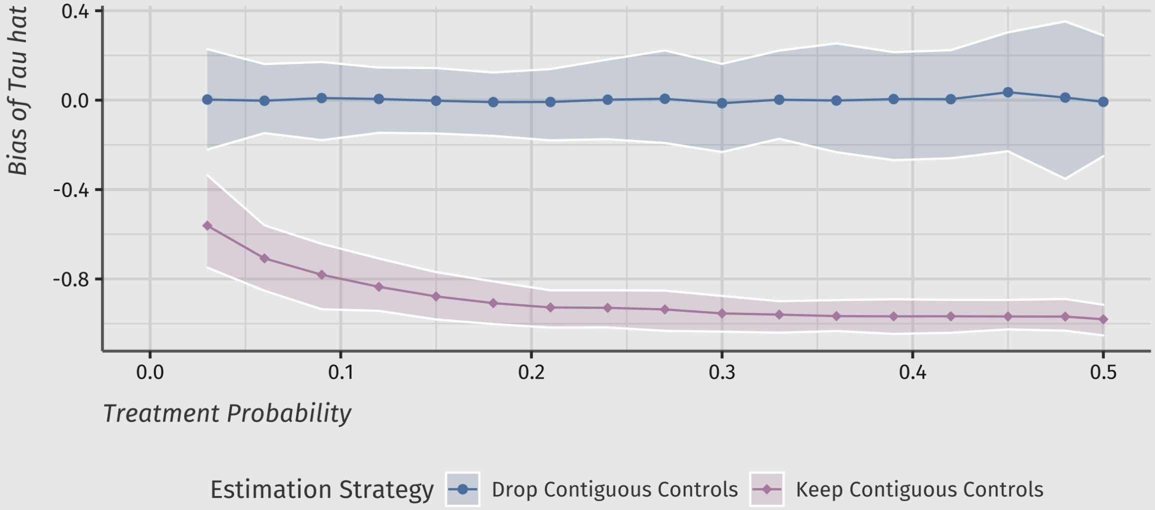
Solution: Removing “contaminated” controls

A common solution to the problem of spillover is to restimate on a subsample with neighboring control units removed.

This simulation drops only units with $\text{Within 40mi.}_{it} = 1$. The exact exposure mapping is typically not observable for researchers, so this is best case scenario.

Dropping control units removes bias, but increases variance

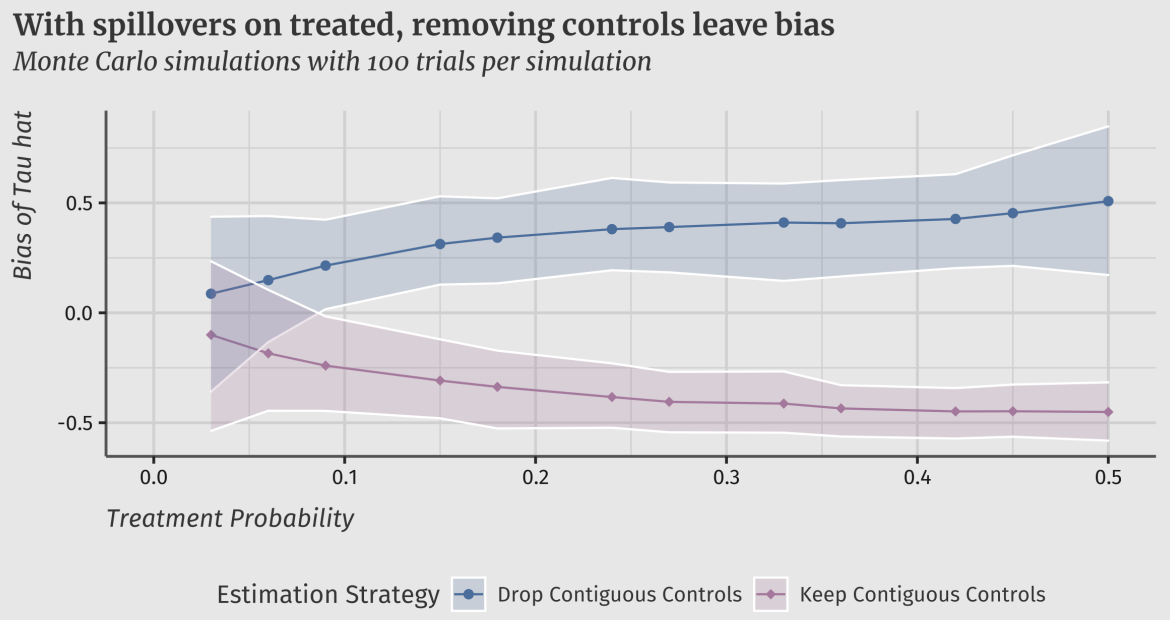
Monte Carlo simulations with 100 trials per simulation



Why Removing Controls Doesn't Work Well

1. Lose precision in treatment effect estimate by dropping control units
2. If there are spillovers on treated units, removing controls won't remove both sources of bias.

I repeat the same simulations but with $\tau_{\text{spillover, treat}} = 0.5$


With spillovers on treated, removing controls leave bias
Monte Carlo simulations with 100 trials per simulation
Bias of Tau hat
0.5
0.0
-0.5
0.0
0.1
0.2
0.3
0.4
0.5
Treatment Probability
Estimation Strategy
Drop Contiguous Controls
Keep Contiguous Controls

Solution: Parametrize potential outcomes

Another option is to directly control for spillovers by parametrizing the exposure mapping.

This is not observable to a researcher which raises the question:

Are there specifications that work well even if the spillovers are misspecified?

Robustness to Misspecification

Generate data using the same data-generating process as before but with different spillover functions:

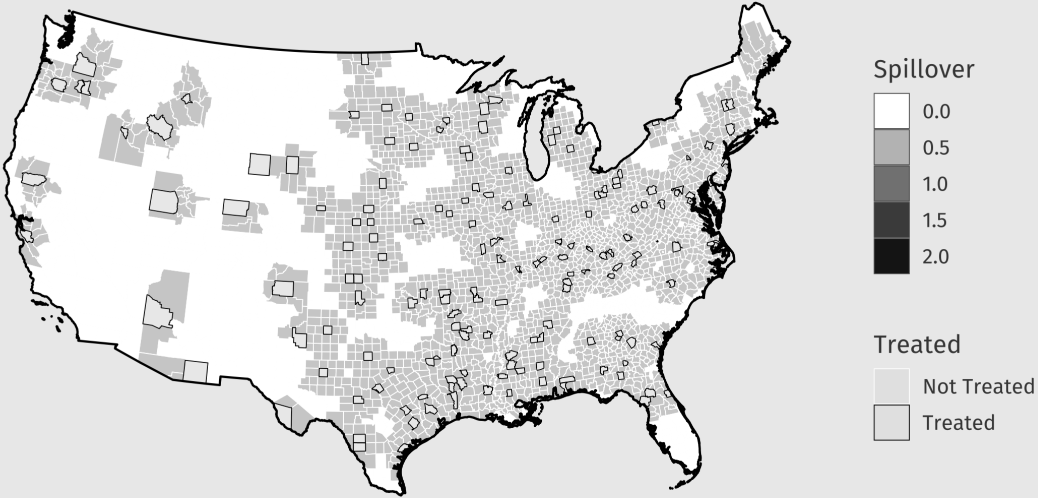
$$y = \mu_t + \mu_i + 2D_{it} + \beta_{\text{spill,control}} * (1 - D_{it})h(\vec{D}, i) + \varepsilon_{it}$$

Then, I estimate each data-generating process using (potentially) misspecified $\tilde{h}(\vec{D}, i)$ and report the average estimate bias.

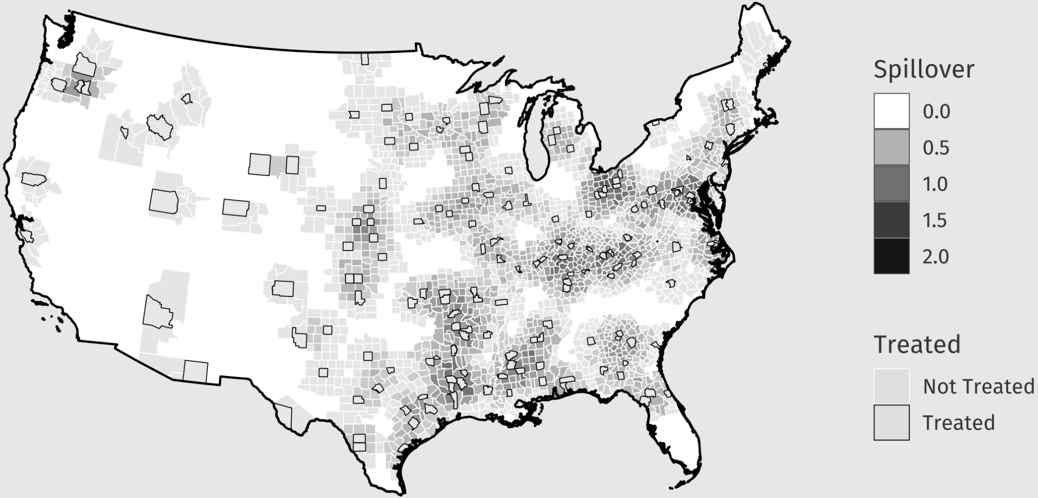
Specifications of $h(\vec{D}, i)$

- *Within* $40/80$*mi.*:
    - Inidicator for nearest treated unit being within $40/80$ miles.
- *Within* $40/80$*mi. (Additive)*:
    - Number of treated units being within $40/80$ miles.

Within 80mi.

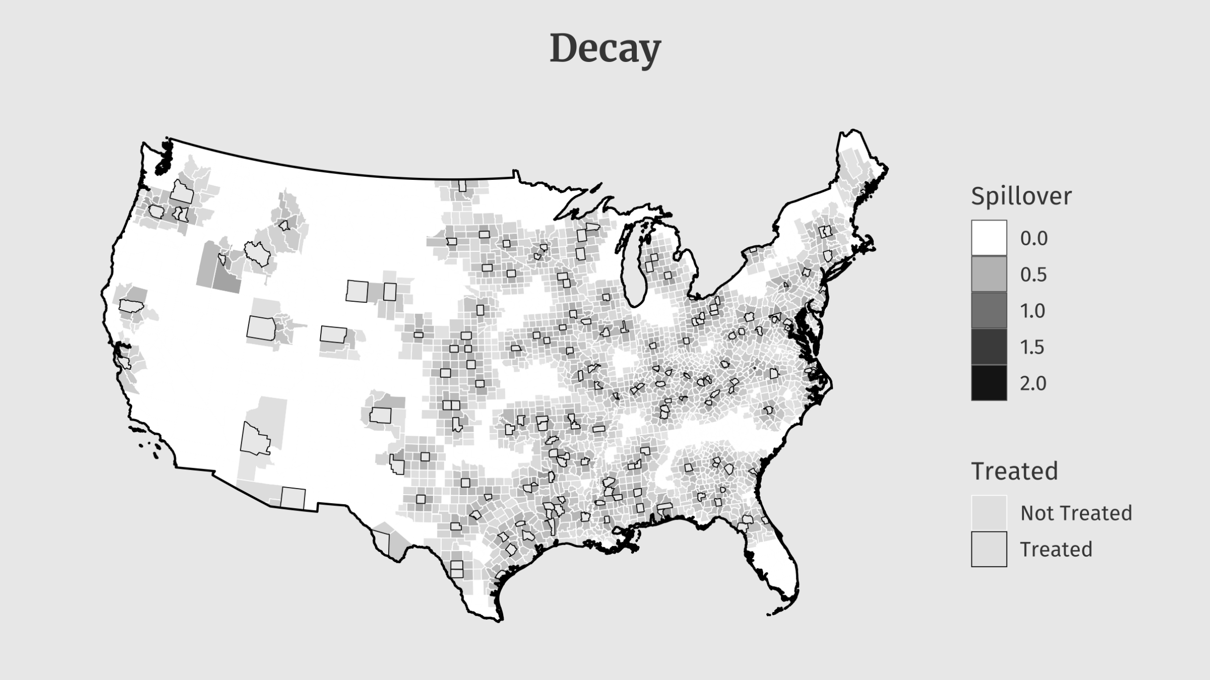


Within 80mi. (Additive)

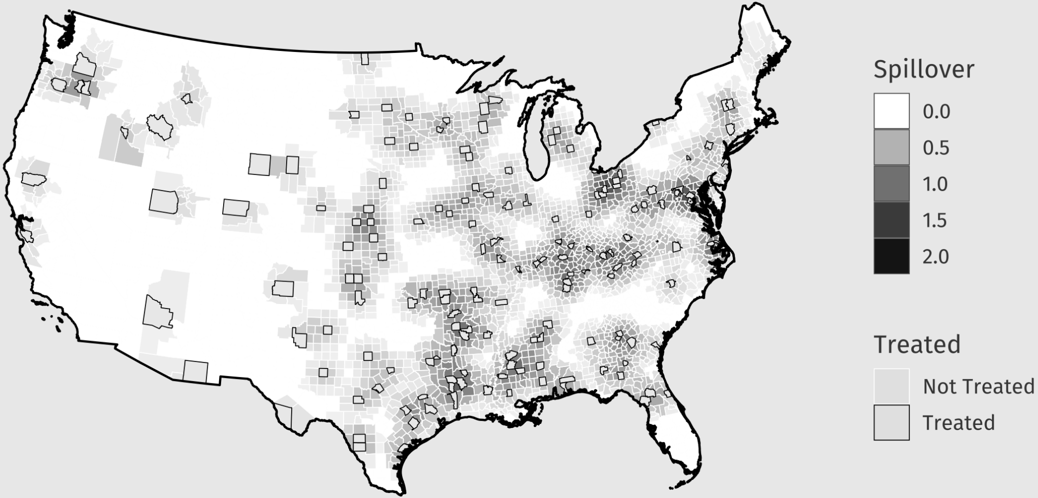


Specifications of $h(\vec{D}, i)$

- *Within* $40/80$*mi.*:
  - Inidicator for nearest treated unit being within $40/80$ miles.
- *Within* $40/80$*mi. (Additive)*:
  - Number of treated units being within $40/80$ miles.
- *Decay*:
  - $\max_j D_j * e^{-0.02d(i,j)} * 1(d(i,j) < 80)$
- *Decay (Additive)*:
  - $\sum_j D_j * e^{-0.02d(i,j)} * 1(d(i,j) < 80)$


Decay
Spillover
0.0
0.5
1.0
1.5
2.0
Treated
Not Treated
Treated

Decay (Additive)



Specifications of $h(\vec{D}, i)$

- *Within* $40/80$*mi.*:
  - Inidicator for nearest treated unit being within $40/80$ miles.
- *Within* $40/80$*mi. (Additive)*:
  - Number of treated units being within $40/80$ miles.
- *Decay*:
  - $\max_j D_j * e^{-0.02d(i,j)} * 1(d(i,j) < 80)$
- *Decay (Additive)*:
  - $\sum_j D_j * e^{-0.02d(i,j)} * 1(d(i,j) < 80)$
- *Rings*:
  - Set of concentric rings. For each ring, indicator for nearest treated unit being within that distance bin
- *Rings (Additive)*:
  - Set of concentric rings. For each ring, number of treated units being within that distance bin

Rings (0-20, 20-30, 30-40, 40-60, 60-80)

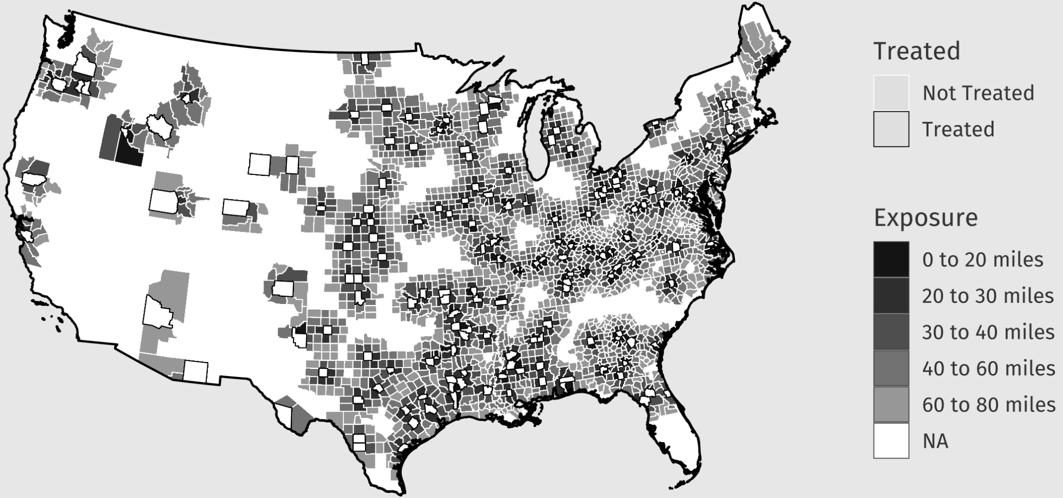


Table 1: Bias from Misspecification of Spillovers

	Data-Generating Process					
Specification	**Within 40mi.**	**Within 80mi.**	**Within 40mi. (Additive)**	**Within 80mi. (Additive)**	**Decay 80mi.**	**Decay 80mi. (Additive)**
TWFE (No Spillovers)	0.258	0.258	0.258	0.258	0.258	0.258
Within 40mi.	−0.005	0.213	−0.005	0.176	0.159	0.143
Within 80mi.	−0.009	−0.009	−0.009	−0.009	−0.009	−0.009
Within 40mi. (Additive)	0.043	0.221	−0.006	0.177	0.174	0.143
Within 80mi. (Additive)	0.034	0.134	−0.012	−0.009	0.099	−0.010
Decay 80mi.	−0.159	0.070	−0.174	0.014	−0.009	−0.033
Decay 80mi. (Additive)	−0.023	0.148	−0.084	0.019	0.088	−0.008
Rings (0-20, 20-30, 30-40)	−0.005	0.213	−0.005	0.176	0.159	0.143
Rings (0-20, 20-30, 30-40, 40-60, 60-80)	−0.009	−0.009	−0.009	−0.009	−0.009	−0.009
Rings (0-20, 20-30, 30-40, 40-60, 60-80) (Additive)	0.036	0.134	−0.008	−0.008	0.100	−0.009

Inidicator that captures all affected unit removes all bias

Table 1: Bias from Misspecification of Spillovers

	Data-Generating Process					
Specification	**Within 40mi.**	**Within 80mi.**	**Within 40mi. (Additive)**	**Within 80mi. (Additive)**	**Decay 80mi.**	**Decay 80mi. (Additive)**
TWFE (No Spillovers)	0.258	0.258	0.258	0.258	0.258	0.258
Within 40mi.	−0.005	0.213	−0.005	0.176	0.159	0.143
Within 80mi.	−0.009	−0.009	−0.009	−0.009	−0.009	−0.009
Within 40mi. (Additive)	0.043	0.221	−0.006	0.177	0.174	0.143
Within 80mi. (Additive)	0.034	0.134	−0.012	−0.009	0.099	−0.010
Decay 80mi.	−0.159	0.070	−0.174	0.014	−0.009	−0.033
Decay 80mi. (Additive)	−0.023	0.148	−0.084	0.019	0.088	−0.008
Rings (0-20, 20-30, 30-40)	−0.005	0.213	−0.005	0.176	0.159	0.143
Rings (0-20, 20-30, 30-40, 40-60, 60-80)	−0.009	−0.009	−0.009	−0.009	−0.009	−0.009
Rings (0-20, 20-30, 30-40, 40-60, 60-80) (Additive)	0.036	0.134	−0.008	−0.008	0.100	−0.009

Inidicator that captures all affected unit removes all bias

Table 1: Bias from Misspecification of Spillovers

	Data-Generating Process					
Specification	**Within 40mi.**	**Within 80mi.**	**Within 40mi. (Additive)**	**Within 80mi. (Additive)**	**Decay 80mi.**	**Decay 80mi. (Additive)**
TWFE (No Spillovers)	0.258	0.258	0.258	0.258	0.258	0.258
Within 40mi.	−0.005	0.213	−0.005	0.176	0.159	0.143
Within 80mi.	−0.009	−0.009	−0.009	−0.009	−0.009	−0.009
Within 40mi. (Additive)	0.043	0.221	−0.006	0.177	0.174	0.143
Within 80mi. (Additive)	0.034	0.134	−0.012	−0.009	0.099	−0.010
Decay 80mi.	−0.159	0.070	−0.174	0.014	−0.009	−0.033
Decay 80mi. (Additive)	−0.023	0.148	−0.084	0.019	0.088	−0.008
Rings (0-20, 20-30, 30-40)	−0.005	0.213	−0.005	0.176	0.159	0.143
Rings (0-20, 20-30, 30-40, 40-60, 60-80)	−0.009	−0.009	−0.009	−0.009	−0.009	−0.009
Rings (0-20, 20-30, 30-40, 40-60, 60-80) (Additive)	0.036	0.134	−0.008	−0.008	0.100	−0.009

Inidicator that captures all affected unit removes all bias

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 treatment of neighbors?

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

Estimation of Spillover Effects

To see which specifications can predict spillover effects well, I estimate the spillover effects for each control unit, $\hat{\beta}_{\text{spill, control}} * \tilde{h}(\vec{D}, i)$.

Then calculate

$$1 - \frac{\overbrace{\sum_{i:D_i=0} (\beta_{\text{spill, control}} h(\vec{D}, i) - \hat{\beta}_{\text{spill, control}} \tilde{h}(\vec{D}, i))^2}^{\text{Mean Square Prediction Error}}}{\underbrace{\sum_{i:D_i=0} (\beta_{\text{spill, control}} h(\vec{D}, i))^2}_{\text{Normalization}}}$$

This gives the proportion of spillovers explained by $\tilde{h}(\vec{D}, i)$

Table 2: Percent of Spillovers Predicted by Specification

	Data-Generating Process					
Specification	**Within 40mi.**	**Within 80mi.**	**Within 40mi. (Additive)**	**Within 80mi. (Additive)**	**Decay 80mi.**	**Decay 80mi. (Additive)**
TWFE (No Spillovers)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Within 40mi.	99.4%	25.9%	85.6%	38.8%	59.5%	56.1%
Within 80mi.	39.8%	96.2%	34.3%	71.7%	85.6%	68.0%
Within 40mi. (Additive)	85.3%	21.2%	99.5%	40.6%	52.0%	60.7%
Within 80mi. (Additive)	45.8%	61.8%	47.2%	98.4%	71.0%	93.6%
Decay 80mi.	60.1%	82.5%	52.7%	75.8%	97.5%	82.2%
Decay 80mi. (Additive)	60.7%	56.9%	63.8%	93.5%	79.0%	98.7%
Rings (0-20, 20-30, 30-40)	98.4%	23.7%	85.9%	37.5%	58.9%	56.2%
Rings (0-20, 20-30, 30-40, 40-60, 60-80)	96.6%	91.7%	84.2%	72.7%	91.9%	78.4%
Rings (0-20, 20-30, 30-40, 40-60, 60-80) (Additive)	83.5%	57.4%	97.6%	95.0%	73.5%	94.9%

Donuts perform best at estimating spillover effects.

It is important to get Additive vs. Non-Additive correct.

Table 2: Percent of Spillovers Predicted by Specification

	Data-Generating Process					
Specification	**Within 40mi.**	**Within 80mi.**	**Within 40mi. (Additive)**	**Within 80mi. (Additive)**	**Decay 80mi.**	**Decay 80mi. (Additive)**
TWFE (No Spillovers)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Within 40mi.	99.4%	25.9%	85.6%	38.8%	59.5%	56.1%
Within 80mi.	39.8%	96.2%	34.3%	71.7%	85.6%	68.0%
Within 40mi. (Additive)	85.3%	21.2%	99.5%	40.6%	52.0%	60.7%
Within 80mi. (Additive)	45.8%	61.8%	47.2%	98.4%	71.0%	93.6%
Decay 80mi.	60.1%	82.5%	52.7%	75.8%	97.5%	82.2%
Decay 80mi. (Additive)	60.7%	56.9%	63.8%	93.5%	79.0%	98.7%
Rings (0-20, 20-30, 30-40)	98.4%	23.7%	85.9%	37.5%	58.9%	56.2%
Rings (0-20, 20-30, 30-40, 40-60, 60-80)	96.6%	91.7%	84.2%	72.7%	91.9%	78.4%
Rings (0-20, 20-30, 30-40, 40-60, 60-80) (Additive)	83.5%	57.4%	97.6%	95.0%	73.5%	94.9%

Donuts perform best at estimating spillover effects.

It is important to get Additive vs. Non-Additive correct.

Table 2: Percent of Spillovers Predicted by Specification

Specification	Within 40mi.	Within 80mi.	Within 40mi. (Additive)	Within 80mi. (Additive)	Decay 80mi.	Decay 80mi. (Additive)
	Data-Generating Process					
TWFE (No Spillovers)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Within 40mi.	99.4%	25.9%	85.6%	38.8%	59.5%	56.1%
Within 80mi.	39.8%	96.2%	34.3%	71.7%	85.6%	68.0%
Within 40mi. (Additive)	85.3%	21.2%	99.5%	40.6%	52.0%	60.7%
Within 80mi. (Additive)	45.8%	61.8%	47.2%	98.4%	71.0%	93.6%
Decay 80mi.	60.1%	82.5%	52.7%	75.8%	97.5%	82.2%
Decay 80mi. (Additive)	60.7%	56.9%	63.8%	93.5%	79.0%	98.7%
Rings (0-20, 20-30, 30-40)	98.4%	23.7%	85.9%	37.5%	58.9%	56.2%
Rings (0-20, 20-30, 30-40, 40-60, 60-80)	96.6%	91.7%	84.2%	72.7%	91.9%	78.4%
Rings (0-20, 20-30, 30-40, 40-60, 60-80) (Additive)	83.5%	57.4%	97.6%	95.0%	73.5%	94.9%

Donuts perform best at estimating spillover effects.

It is important to get Additive vs. Non-Additive correct.

Roadmap of Talk

Theory

Estimation with Spillovers

Application in Urban Economics

Conclusion

Tennessee Valley Authority

Kline and Moretti (2014a) 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. (Hundreds of dollars spent anually per person)
- The program focused on large-scale dam construction that brought cheap wholesale electricity to the region

Research Question: What are the local effects of TVA on manufacturing and agricultural economies? Do these effects come at the cost of other counties?

Identification

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

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

- 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

- **Agriculture:**
  - Employees might be drawn to hire wages for new manufacturing jobs in Tennessee Valley (negative spillover on control units)
- **Manufacturing:**
  - Cheap electricity might be available to nearby counties (positive spillover on control units)
  - Manufacturing jobs that would have been created in the control units in the absence of treatment might move to the Tennessee Valley (negative spillover on control units)
- **Income:**
  - Since manufacutring jobs pay higher than agricultural jobs, the above two spillover effects can cause changes in median family income

Specification including spillovers

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

- Ring(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

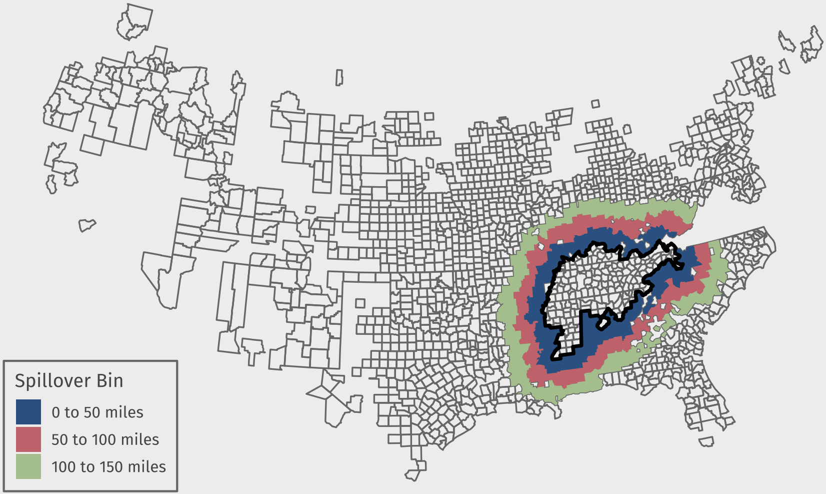


Table 3: Effects of Tennessee Valley Authority on Decadel Growth

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

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

Table 3: Effects of Tennessee Valley Authority on Decadel Growth

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

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

Table 3: Effects of Tennessee Valley Authority on Decadel Growth

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

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

Identification Strategies and Place-Based Policies

The literature on federal Enterprise Zones, place-based policy that gives tax breaks to businesses that locate within the boundary, has found conflicting results, suggesting positive or near-zero effects of the program (Neumark and Young, 2019).

- Busso, Gregory, and Kline (2013) compare census tracts in Empowerment Zones to census tracts that qualified and were rejected from the program. They find significant large reduction of poverty.
- Neumark and Kolko (2010) compare census tracts in Empowerment Zones to census tracts within 1,000 feet of the Zone. They find near-zero effects on poverty.

My framework can explain both of these results. If census tracts just outside the Empowerment Zones also benefit from the policy, then the estimates of Neumark and Kolko (2010) are attenuated towards zero

Identification Strategies and Place-Based Policies

The literature on federal Enterprise Zones, place-based policy that gives tax breaks to businesses that locate within the boundary, has found conflicting results, suggesting positive or near-zero effects of the program (Neumark and Young, 2019).

- Busso, Gregory, and Kline (2013) compare census tracts in Empowerment Zones to census tracts that qualified and were rejected from the program. They find significant large reduction of poverty.
- Neumark and Kolko (2010) compare census tracts in Empowerment Zones to census tracts within 1,000 feet of the Zone. They find near-zero effects on poverty.

My framework can explain both of these results. If census tracts just outside the Empowerment Zones also benefit from the policy, then the estimates of Neumark and Kolko (2010) are attenuated towards zero

Identification Strategies and Place-Based Policies

The literature on federal Enterprise Zones, place-based policy that gives tax breaks to businesses that locate within the boundary, has found conflicting results, suggesting positive or near-zero effects of the program (Neumark and Young, 2019).

- Busso, Gregory, and Kline (2013) compare census tracts in Empowerment Zones to census tracts that qualified and were rejected from the program. They find significant large reduction of poverty.
- Neumark and Kolko (2010) compare census tracts in Empowerment Zones to census tracts within 1,000 feet of the Zone. They find near-zero effects on poverty.

My framework can explain both of these results. If census tracts just outside the Empowerment Zones also benefit from the policy, then the estimates of Neumark and Kolko (2010) are attenuated towards zero

Identification Strategies and Place-Based Policies

The literature on federal Enterprise Zones, place-based policy that gives tax breaks to businesses that locate within the boundary, has found conflicting results, suggesting positive or near-zero effects of the program (Neumark and Young, 2019).

- Busso, Gregory, and Kline (2013) compare census tracts in Empowerment Zones to census tracts that qualified and were rejected from the program. They find significant large reduction of poverty.
- Neumark and Kolko (2010) compare census tracts in Empowerment Zones to census tracts within 1,000 feet of the Zone. They find near-zero effects on poverty.

My framework can explain both of these results. If census tracts just outside the Empowerment Zones also benefit from the policy, then the estimates of Neumark and Kolko (2010) are attenuated towards zero

Roadmap of Talk

Theory

Estimation with Spillovers

Application in Urban Economics

Conclusion

Conclusion

- I decomposed the TWFE estimate into the direct effect and two spillover terms
- I showed that a set of concentric rings allows for estimation of the direct effect of treatment and they are able to model spillovers well
- For place-based policies, I show the importance of considering spatial spillovers when estimating treatment effects
- More generally, identification strategies that use very close control units in order to minimize differences in unobservables should consider the problems with treatment effect spillovers.

Roadmap of Talk

Appendix

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:

$$\begin{aligned}\Delta &\equiv [\underline{\Delta}, \overline{\Delta}] \\ &= \left[\min \tau_{\text{spill,treated}} - \tau_{\text{spill,control}},\ \max \tau_{\text{spill,treated}} - \tau_{\text{spill,control}}\right]\end{aligned}$$

The partially identified set is formed by

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

Bias-adjusted Inference

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

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

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

Back to Example of Bias