Difference-in-Differences with Spatial Spillovers

Kyle Butts

April 30, 2022

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**, when the effect of treatment extends over the treatment boundaries (states, counties, etc.).

Example: Amazon Shipping Center on employment

- A shipping center opening in county *c* has positive employment effects on **nearby control counties**
- Having nearby counties with shipping centers raises wages and therefore reduces the employment effect on **treated counties**

This Paper

In this paper, I...

- Present a potential outcomes framework to formalize treatment and spillover effects and discuss potential estimands of interest
- Discuss estimation of effects that are *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 occur in the 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 occur in the presence of spillover effects:

- **Spillover onto Control Units:** Nearby “control” units fail to estimate counterfactual trends

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 occur in the presence of spillover effects:

- **Spillover onto Control Units:** Nearby "control" units fail to estimate counterfactual trends
- **Spillover onto other Treated Units:** Treated units are also affected by nearby units and therefore combine "direct" effects with spillover effects

Outline

1– Formalize spillovers into a potential outcomes framework:

[**Clarke˙2017**, **Berg˙Streitz˙2019**, and **Verbitsky-Savitz˙Raudenbush˙2012**]

→ I decompose the difference-in-differences estimator into three parts: Direct Effect of Treatment, Spillover onto Treated Units, Spillover onto Control Units

→ Discuss estimands of interest and classify which are identifiable and which are not

Outline

2– Apply framework to Urban Economics

→ Revisit **Kline˙Moretti˙2014a** analysis of the Tennessee Valley Authority

- The local effect estimate is contaminated by spillover effects to neighboring counties (**Kline˙Moretti˙2014b**)
- Large scale manufacturing investment creates an 'urban shadow' (**Cuberes˙Desmet˙Rappaport˙2021**; **Fujita˙Krugman˙Venables˙2001**)

→ Discuss how framework can reconcile conflicting findings on effect of federal Empowerment Zones (**Busso˙Gregory˙Kline˙2013**; **Neumark˙Kolko˙2010**)

→ Event Study estimates of Community Health Centers find highly localized efffects (**Bailey˙Goodman˙Bacon˙2015**)

Roadmap

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 $h(\vec{D}, i) = \vec{0}$.

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

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 $h(\vec{D}, i) = \vec{0}$.

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

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

- **Treatment within x miles:**

  $h(\vec{D}, i) = max_j \, 1\left(d(i,j) \leq x\right)$ where $d(i,j)$ is the distance between counties i and j.

  $\rightarrow$ e.g. library access where x is the maximum distance people will travel

  $\rightarrow$ Spillovers are non-additive, i.e. spillover effects do not depend on number of nearby treated areas

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.

  $\rightarrow$ e.g. library access where x is the maximum distance people will travel

  $\rightarrow$ Spillovers are non-additive

- **Number of Treated within x miles:**

  $h(\vec{D}, i) = \sum_{j=1}^{k} 1(d(i,j) \leq x)$.

  $\rightarrow$ e.g. Amazon shipping center

  $\rightarrow$ Agglomeration economies suggest spillovers are additive

Estimands

Treatment Effect without Spillovers

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

- The average effect of switching on unit i's treatment

Estimands

Switching Effect

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

- Keep everyone's treatment constant and toggle unit i's treatment effect. Average across all units with exposure h.
- This is policy relevant: what will happen if my county turns on treatment.
- This requires knowledge of $h_i(\vec{D})$ in order to find control units to estimate $Y_{i1}(0, h_i(\vec{D}))$.

Estimands

Switching Effect

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

- Keep everyone's treatment constant and toggle unit i's treatment effect. Average across all units with exposure h.
- This is policy relevant: what will happen if my county turns on treatment.
- This requires knowledge of $h_i(\vec{D})$ in order to find control units to estimate $Y_{i1}(0, h_i(\vec{D}))$.

Estimands

Switching Effect

$$\tau_{\text{switch}}(h) = \mathbb{E}\left[Y_{i1}(1, h_i(\vec{D})) - Y_{i0}(0, \vec{0}) \mid D_i = 1, h_i(\vec{D}) = h\right]$$
$$- \mathbb{E}\left[Y_{i1}(0, h_i(\vec{D})) - Y_{i0}(0, \vec{0}) \mid D_i = 1, h_i(\vec{D}) = h\right]$$

Identification requires two things:

1. Knowledge of $h_i(\vec{D})$ in order to estimate the second term with control units.
2. Spillover effect homogeneity (so second term would be the same for treated units).

Estimands

Total Effect

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

- Toggle entire vector of treatment effects. Average across all treated units.
- This is helpful for post-hoc policy analysis: what was the average effect on treated units of implementing $\vec{D}$.

Estimands

Total Effect

$$\tau_{\text{total}} = \mathbb{E}\left[Y_{i1}(1, h_i(\vec{D})) - Y_{i0}(0, \vec{0}) \mid D_i = 1\right]$$
$$- \mathbb{E}\left[Y_{i1}(0, \vec{0}) - Y_{i0}(0, \vec{0}) \mid D_i = 1\right]$$

Identification much simpler:

- Need to identify control units without spillover effects

Estimands

Direct Effect

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

- Toggle treatment effect for individuals with no exposure.
- This is the switching effect with $h = 0$.
- Identified under mild assumptions.

Estimands

Direct Effect

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

- Toggle treatment effect for individuals with no exposure.
- This is the switching effect with $h = 0$.
- Identified under mild assumptions.

Estimands

Direct Effect

$$\tau_{\text{direct}} = \mathbb{E}\left[Y_{i1}(1, \vec{0}) - Y_{i0}(0, \vec{0}) \mid D_i = 1\right] - \mathbb{E}\left[Y_{i1}(0, \vec{0}) - Y_{i0}(0, \vec{0}) \mid D_i = 1\right]$$

Identification much simpler:

- Need to identify treated/control units without spillover effects

Estimands

Spillover Effects

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

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

What does Difference-in-Differences identify?

Parallel Trends

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

Assumption: *Parallel Counterfactual Trends*

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

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

Changes in outcomes do not depend on treatment status

What does Difference-in-Differences identify?

Decomposition

With the parallel trends assumption, I decompose the difference-in-differences estimate as follows:

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

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

$$= \tau_{\text{total}} - \tau_{\text{spillover, control}}$$

What does Difference-in-Differences identify?

Decomposition

With the parallel trends assumption, I decompose the difference-in-differences estimate as follows:

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

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

$$= \tau_{\text{total}} - \tau_{\text{spillover, control}}$$

What does Difference-in-Differences identify?

Decomposition

With the parallel trends assumption, I decompose the difference-in-differences estimate as follows:

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

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

$$= \tau_{\text{total}} - \tau_{\text{spillover, control}}$$

Identification of Direct/Total Effects

Assumption: *Spillovers are Local*

Let $d(i,j)$ be the distance between units i and j. There exists a distance $\bar{d}$ such that

(i) For all units i,

$$\min_{j:D_j=1} d(i,j) > \bar{d} \implies h(\vec{D}, i) = \vec{0}.$$

(ii) There are treated and control units such that $\min_{j:\,D_j=1} d(i,j) > \bar{d}$.

Identification of Total Effect

With assumption that spillovers are local, define S_{it} to be an indicator equal to one in the post period for all units with $h(\vec{D}, i) \neq \vec{0}$ (and potentially some units with $= \vec{0}$).

Estimation of the following equation:

$$y_{it} = \mu_t + \mu_i + \tau D_{it} + \tau_{\text{spill,control}} S_{it} * (1 - D_{it}) + \varepsilon_{it}$$

- $\hat{\tau}$ is consistent for τ_{total}
- $\hat{\tau}_{\text{spill, control}}$ is not consistent for average spillover effects.

Identification of Direct Effect

With assumption that spillovers are local, define S_{it} to be an indicator equal to one in the post period for all units with $h(\vec{D}, i) \neq \vec{0}$ (and potentially some units with $= \vec{0}$).

Estimation of the following equation:

$$y_{it} = \mu_t + \mu_i + \tau D_{it} + \tau_{\text{spill,treat}} S_{it} * D_{it} + \tau_{\text{spill,control}} S_{it} * (1 - D_{it}) + \varepsilon_{it}$$

- $\hat{\tau}$ is consistent for τ_{direct}
- $\hat{\tau}_{\text{spill}}$'s are not consistent for average spillover effects.

Roadmap

Tennessee Valley Authority

Kline˙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 Questions:

- What is the total effect of TVA investments on manufacturing and agricultural employment?
- Do these effects come at the cost of other counties?

Tennessee Valley Authority

Kline˙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 Questions:

- What is the total effect of TVA investments on manufacturing and agricultural employment?
- Do these effects come at the cost of other counties?

Identification

Kline˙Moretti˙2014a run the county-level difference-in-differences specification:

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

- y are outcomes for agricultural employment and manufacturing employment.
- 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)

Specification including spillovers

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

- Ring(d) is a set of indicators for being in the following distance bins (in miles) from the Tennessee Valley Authority:

$$d \in \{(0, 50], (50, 100], (100, 150](150, 200]\}$$

Table: Effects of Tennessee Valley Authority on Decadel Growth, 1940-2000

	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.	TVA between 150-200 mi.
Dependent Var.	(1)	(2)	(3)	(4)	(5)	(6)
Agricultural employment	-0.0514^{***}	-0.0739^{***}	-0.0371^{***}	-0.0164	-0.0298^{***}	-0.0157^{*}
	(0.0114)	(0.0142)	(0.0002)	(0.0114)	(0.0096)	(0.0088)
Manufacturing employment	0.0560^{***}	0.0350	-0.0203^{***}	-0.0245	-0.0331^{*}	-0.0296^{**}
	(0.0161)	(0.0218)	(0.0006)	(0.0282)	(0.0189)	(0.0142)

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

Table: Effects of Tennessee Valley Authority on Decadel Growth, 1940-2000

	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.	TVA between 150-200 mi.
Dependent Var.	(1)	(2)	(3)	(4)	(5)	(6)
Agricultural employment	-0.0514^{***}	-0.0739^{***}	-0.0371^{***}	-0.0164	-0.0298^{***}	-0.0157^{*}
	(0.0114)	(0.0142)	(0.0002)	(0.0114)	(0.0096)	(0.0088)
Manufacturing employment	0.0560^{***}	0.0350	-0.0203^{***}	-0.0245	-0.0331^{*}	-0.0296^{**}
	(0.0161)	(0.0218)	(0.0006)	(0.0282)	(0.0189)	(0.0142)

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

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˙Young˙2019**).

Identification Strategies and Place-Based Policies

- **Busso˙Gregory˙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˙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˙Kolko˙2010** are attenuated towards zero

Identification Strategies and Place-Based Policies

- **Busso˙Gregory˙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˙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˙Kolko˙2010** are attenuated towards zero

Identification Strategies and Place-Based Policies

- **Busso˙Gregory˙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˙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˙Kolko˙2010** are attenuated towards zero

Roadmap

Gardner (2021) Overview

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

The problem with estimating this by OLS is that the treatment variable becomes residualized $\tilde{D}_{it}$ and this leads to all sorts of problems... (see new diff-in-diff literature)

Gardner (2021) recommends a two-step approach:

1. Estimate μ_i and μ_t using never-treated/not-year-treated observations (D_{it} = 0). Then subtract off $\hat{\mu}_i$ and $\hat{\mu}_t$.
2. Then, regress $y_{it} - \hat{\mu}_i - \hat{\mu}_t \equiv \tilde{y}_{it}$ on τD_{it} (or event study leads/lags). This estimate is unbiased because D_{it} is non-residualized (standard errors require adjusting).

Gardner (2021) Overview

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

The problem with estimating this by OLS is that the treatment variable becomes residualized $\tilde{D}_{it}$ and this leads to all sorts of problems... (see new diff-in-diff literature)

Gardner (2021) recommends a two-step approach:

1. Estimate μ_i and μ_t using never-treated/not-year-treated observations (D_{it} = 0). Then subtract off $\hat{\mu}_i$ and $\hat{\mu}_t$.
2. Then, regress $y_{it} - \hat{\mu}_i - \hat{\mu}_t \equiv \tilde{y}_{it}$ on τD_{it} (or event study leads/lags). This estimate is unbiased because D_{it} is non-residualized (standard errors require adjusting).

Gardner (2021) Overview

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

The problem with estimating this by OLS is that the treatment variable becomes residualized $\tilde{D}_{it}$ and this leads to all sorts of problems... (see new diff-in-diff literature)

Gardner (2021) recommends a two-step approach:

1. Estimate μ_i and μ_t using never-treated/not-year-treated observations (D_{it} = 0). Then subtract off $\hat{\mu}_i$ and $\hat{\mu}_t$.
2. Then, regress $y_{it} - \hat{\mu}_i - \hat{\mu}_t \equiv \tilde{y}_{it}$ on τD_{it} (or event study leads/lags). This estimate is unbiased because D_{it} is non-residualized (standard errors require adjusting).

Controlling for Spillovers in Staggered Treatment Timing

$$y_{it} = \mu_i + \mu_t + \tau D_{it} + \tau_{\text{spill,control}} S_{it} * (1 - D_{it}) + \tau_{\text{spill,treat}} S_{it} * D_{it} + \varepsilon_{it}$$

Adjust two-step approach:

1. Estimate μ_i and μ_t using observations that are not yet treated/affected by spillovers ($D_{it} = 0$ and $S_{it} = 0$). Then subtract off $\hat{\mu}_i$ and $\hat{\mu}_t$.
2. Then, regress $\tilde{y}_{it}$ on $\tau D_{it} + \tau_{\text{spill,control}} S_{it} * (1 - D_{it}) + \tau_{\text{spill,treat}} S_{it} * D_{it}$ (or interacted event study leads/lags). This estimate is unbiased because D_{it} is non-residualized (standard errors require adjusting).

Controlling for Spillovers in Staggered Treatment Timing

$$y_{it} = \mu_i + \mu_t + \tau D_{it} + \tau_{\text{spill,control}} S_{it} * (1 - D_{it}) + \tau_{\text{spill,treat}} S_{it} * D_{it} + \varepsilon_{it}$$

Adjust two-step approach:

1. Estimate μ_i and μ_t using observations that are not yet treated/affected by spillovers ($D_{it} = 0$ and $S_{it} = 0$). Then subtract off $\hat{\mu}_i$ and $\hat{\mu}_t$.

2. Then, regress $\tilde{y}_{it}$ on $\tau D_{it} + \tau_{\text{spill,control}} S_{it} * (1 - D_{it}) + \tau_{\text{spill,treat}} S_{it} * D_{it}$ (or interacted event study leads/lags). This estimate is unbiased because D_{it} is non-residualized (standard errors require adjusting).

Community Health Centers

Bailey˙Goodman˙Bacon˙2015 study the creation of federal community health centers between 1965 and 1974.

Research Question:

- Do low-/no-cost health services lower the mortality rate of the treated counties?
- *New Question:* Do these effects spread to neighboring counties?

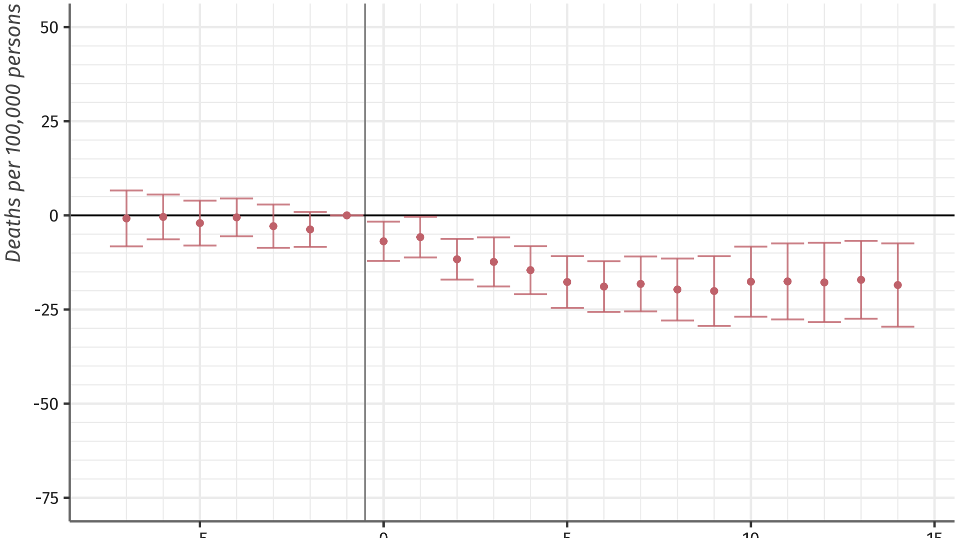

Figure: Effects of Establishment of Community Health Centers

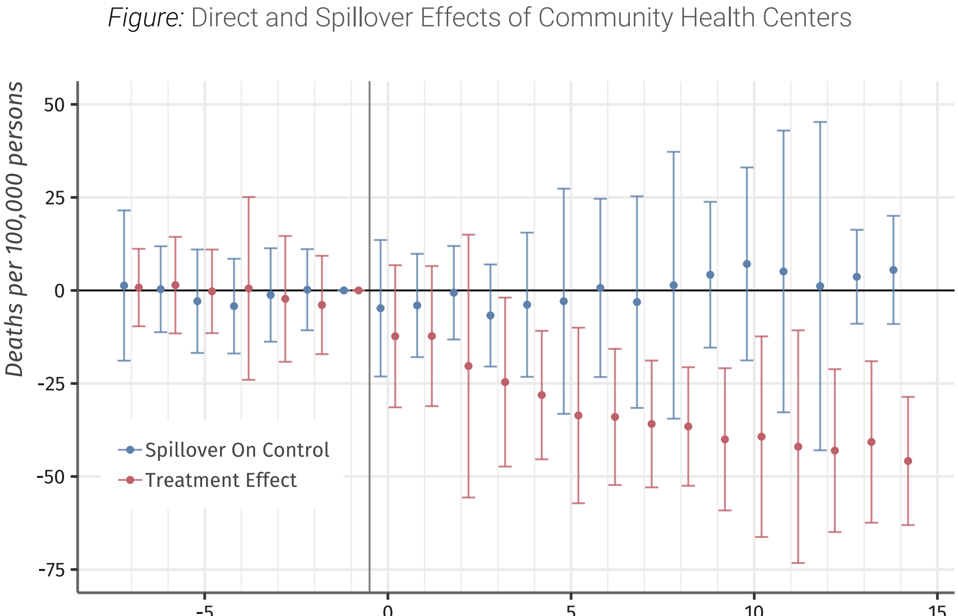

Figure: Direct and Spillover Effects of Community Health Centers
Deaths per 100,000 persons
50
25
0
-25
-50
-75
-5
0
5
10
15
Spillover On Control
Treatment Effect

Roadmap

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.