Difference-in-Differences with Spatial Spillovers

Kyle Butts

December 2024

Spatial Spillovers

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

$$\tau \equiv \mathbb{E}\left[Y_1(1) - Y_1(0) \mid D = 1\right]$$

Estimation is complicated by **spillover effects**, when the effect of treatment extends over the treatment boundaries (e.g. census tracts)

Example: *Federal empowerment zones*

- Economic activity generated by tax breaks can benefit (+) or steal business (-) from neighboring tracts by agglomeration forces

This Paper

This paper uses the potential outcomes framework from Sävje, Aronow, and Hudgens (2021) and Vazquez-Bare (2023) to formalize spillover effects

Two potential treatment effects of interest:

- **Switching Effect**: holding fixed the treatment of others, what is the effect of switching my treatment
- **Total Effect**: post-hoc analysis of what's the average impact on the treated units of the implemented treatment regime

This Paper

Discuss non-parametric identification of effects in the presence of spillovers

1. Can not identify the switching effect without parameterizing spillovers
2. If spillovers are 'local', then the total effect is identified under a *modified* parallel-trends assumption
    - → Commonly used estimator of adding a dummy for being near to treatment is shown to be consistent
    - → 'Far away' units serve as the counterfactual trend estimate

Contribution

1. Formalize spillovers into a potential outcomes framework:
   → Other papers discuss identification under parametric restrictions (Berg and Streitz, 2019; Clarke, 2017; Verbitsky-Savitz and Raudenbush, 2012)

2. Interference in quasi-experimental settings
   → Complements work on partial interference (Miguel and Kremer, 2004; Sobel, 2006)
   - Generalized strategy of finding 'unexposed groups' to setting where clusters are not distinct

   → Results for general interference typically rely on simple random or block random experiments (Leung, 2020; Vazquez-Bare, 2023)
   - This paper considers quasi-experimental methods, but has to condition on the realized treatment vector

Revisiting 3 empirical applications

Kline and Moretti (2014): employment effects of the Tennessee Valley Authority

- Large scale manufacturing investment creates an 'urban shadow' (Cuberes, Desmet, and Rappaport, 2021; Fujita, Krugman, and Venables, 2001)

Revisiting 3 empirical applications

Kline and Moretti (2014): employment effects of the Tennessee Valley Authority

- Large scale manufacturing investment creates an 'urban shadow' (Cuberes, Desmet, and Rappaport, 2021; Fujita, Krugman, and Venables, 2001)

Chen, Glaeser, and Wessel (2023): housing price impacts of 2017 Opportunity Zones

- Clarify why different identification strategies (rejected applicatns vs. neighboring control units) can find conflicting effects (Busso, Gregory, and Kline, 2013; Neumark and Kolko, 2010, e.g.)

Revisiting 3 empirical applications

Kline and Moretti (2014): employment effects of the Tennessee Valley Authority

- Large scale manufacturing investment creates an 'urban shadow' (Cuberes, Desmet, and Rappaport, 2021; Fujita, Krugman, and Venables, 2001)

Chen, Glaeser, and Wessel (2023): housing price impacts of 2017 Opportunity Zones

- Clarify why different identification strategies (rejected applicatns vs. neighboring control units) can find conflicting effects (Busso, Gregory, and Kline, 2013; Neumark and Kolko, 2010, e.g.)

Bailey and Goodman-Bacon (2015): Extend methodology to staggered treatment timing. Event Study estimates of Community Health Centers find no spillover effects

Theory

Application in Urban Economics

Conclusion

Potential Outcomes Framework

For this presentation, I will assume all treatment occurs at the same time (2-periods or pre-post averages). Staggered treatment-timing is in the paper.

The potential outcome of unit $i \in \{1, \dots, N\}$ at time t with treatment status $D_i \in \{0, 1\}$:

$$Y_{it}(D_i, h_i(\boldsymbol{D}))$$

- $\boldsymbol{D} \in \{0, 1\}^N$ is the vector of all units treatments.
- The function $h_i(\boldsymbol{D})$ maps the entire treatment vector into an 'exposure mapping' which can be a scalar or a vector (Sävje, Aronow, and Hudgens, 2021; Vazquez-Bare, 2023).
  - $\rightarrow$ No exposure is when $h_i(\boldsymbol{D}) = \mathbf{0}$.

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

Treatment within x miles:

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

- e.g. library access where x is the maximum distance people will travel
- Spillovers are non-additive, i.e. spillover effects do not depend on number of nearby treated areas

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

Treatment within x miles:

$h_i(\boldsymbol{D}) = max_j \, 1(d(i,j) \leq x)$ where $d(i,j)$ is the distance between units 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:

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

- e.g. Amazon shipping center
- Agglomeration economies suggest spillovers are additive

Treatment Effect without Spillovers

Without spillovers, there is only a singular treated/untreated state, so the average treatment effect on the treated is unique:

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

- With spillovers, there are many different potential outcome contrasts to consider.

Estimand: Switching Effect

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

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 I turn on treatment for my unit

Estimand: Switching Effect

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

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 I turn on treatment for my unit
- The effect can depend on how exposed you are, h
  - $\rightarrow$ E.g. If treatment is more effective when your neighbors are treated (no "slippage")

Estimand: Switching Effect

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

Estimand: Switching Effect

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

Identification requires three things:

1. Knowledge of $h_i(\boldsymbol{D})$ in order to estimate the second term with control units
2. For similarly exposed treated and control units, need parallel trends

Estimand: Switching Effect

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

Identification requires three things:

1. Knowledge of $h_i(\boldsymbol{D})$ in order to estimate the second term with control units
2. For similarly exposed treated and control units, need parallel trends
3. For similarly exposed treated and control units, need spillover effect homogeneity, so second term would be the same for treated units (Callaway, Goodman-Bacon, and Sant'Anna, 2021)

Estimand: Total Effect

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

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 $\boldsymbol{D}$.

Estimand: Total Effect

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

Identification much simpler:

1. Need to identify control units without spillover effects
2. For the non-exposed control units, need parallel trends
   - → So that the average change in Y for the non-exposed units equals the average change in Y for the treated units

Estimands: Spillover Effect

For each unit, we define the spillover effect for that unit as:

$$Y_{i1}(d, h_i(\boldsymbol{D})) - Y_{i1}(d, 0)$$

- Average over a subset of (control) units; e.g. neighboring census tracts to empowerment zone

Modified Parallel Trends

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

Assumption: *Parallel Counterfactual Trends*

In the *complete absence of treatment* (not just the absence of individual i's treatment), changes in outcomes do not depend on treatment status:

$$\mathbb{E}[Y_{i,1}(0, \mathbf{0}) - Y_{i,0}(0, \mathbf{0}) \mid D_i = 1] = \mathbb{E}[Y_{i,1}(0, \mathbf{0}) - Y_{i,0}(0, \mathbf{0}) \mid D_i = 0]$$

What does Difference-in-Differences identify?

With the parallel trends assumption, the difference-in-differences estimate can be decomposed as follows:

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

$$= \tau_{\text{total}} - \tau_{\text{spill}},$$

where $\tau_{\text{spill}} = \mathbb{E}[Y_{i1}(d, h_i(\boldsymbol{D})) - Y_{i1}(d, 0) \mid D_i = 0]$ is the average spillover on all control units

- Biased in the opposite direction of spillover effects onto control units.

Identification of Total Effect

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,

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

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

Modified Parallel Trends

Assumption: *Modified Parallel Counterfactual Trends*

For a chosen $\bar{d}$, define S_i to be a dummy for units within $\bar{d}$ of a treated unit. To identify the total effect, we need

$$\mathbb{E}[Y_{i,1}(0,\mathbf{0}) - Y_{i,0}(0,\mathbf{0}) \mid D_i = 1] = \mathbb{E}[Y_{i,1}(0,\mathbf{0}) - Y_{i,0}(0,\mathbf{0}) \mid D_i = 0, S_i = 0].$$

Modified Parallel Trends

The modified parallel trends assumption is worth remarking on in terms of applied work:

- Since unobservables typically vary smoothly over space, the best counterfactual units are those near to treatment (e.g. Baum-Snow and Ferreira (2015))
- Using far-away units can prevent bias from spillovers, but can possibly adds bias from non-parallel trends

Identification of Total Effect

With assumption that spillovers are local and modified parallel trends, we can either form the difference-in-differences estimate.

More simply by running this regression:

$$y_{it} = \mu_t + \lambda_i + \tau D_i \mathbb{1}(t = 1) + \tau_{\text{spill}} S_i (1 - D_i) \mathbb{1}(t = 1) + \varepsilon_{it}$$

- $\hat{\tau}$ is consistent for τ_{total}
- Standard errors that account for spatial correlation are likely preferred (Conley, 1999)

h

$$y_{it} = \mu_t + \lambda_i + \tau D_i \mathbb{1}(t=1) + \tau_{\text{spill}} S_i (1 - D_i) \mathbb{1}(t=1) + \varepsilon_{it}$$

By strengthening the modified parallel trends assumption, $\hat{\tau}_{\text{spill}}$ estimates the average spillover effect among the nearby control units ($D_i = 0, S_i = 1$):

$$\mathbb{E}[Y_{i,1}(0, \mathbf{0}) - Y_{i,0}(0, \mathbf{0}) \mid D_i = 0, S_i = 0] = \mathbb{E}[Y_{i,1}(0, \mathbf{0}) - Y_{i,0}(0, \mathbf{0}) \mid D_i = 0, S_i = 1]$$

h

$$y_{it} = \mu_t + \lambda_i + \tau D_i \mathbb{1}(t = 1) + \tau_{\text{spill}} S_i (1 - D_i) \mathbb{1}(t = 1) + \varepsilon_{it}$$

By strengthening the modified parallel trends assumption, $\hat{\tau}_{\text{spill}}$ estimates the average spillover effect among the nearby control units ($D_i = 0, S_i = 1$):

$$\mathbb{E}[Y_{i,1}(0, \mathbf{0}) - Y_{i,0}(0, \mathbf{0}) \mid D_i = 0, S_i = 0] = \mathbb{E}[Y_{i,1}(0, \mathbf{0}) - Y_{i,0}(0, \mathbf{0}) \mid D_i = 0, S_i = 1]$$

Can chop up S_i into a set of concentric rings, to estimate spillover as a function of distance

- requires parallel trends for each ring (Butts, 2023)

Theory

Application in Urban Economics

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

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. (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 units?

Identification

Kline and Moretti (2014) run the unit-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 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 units

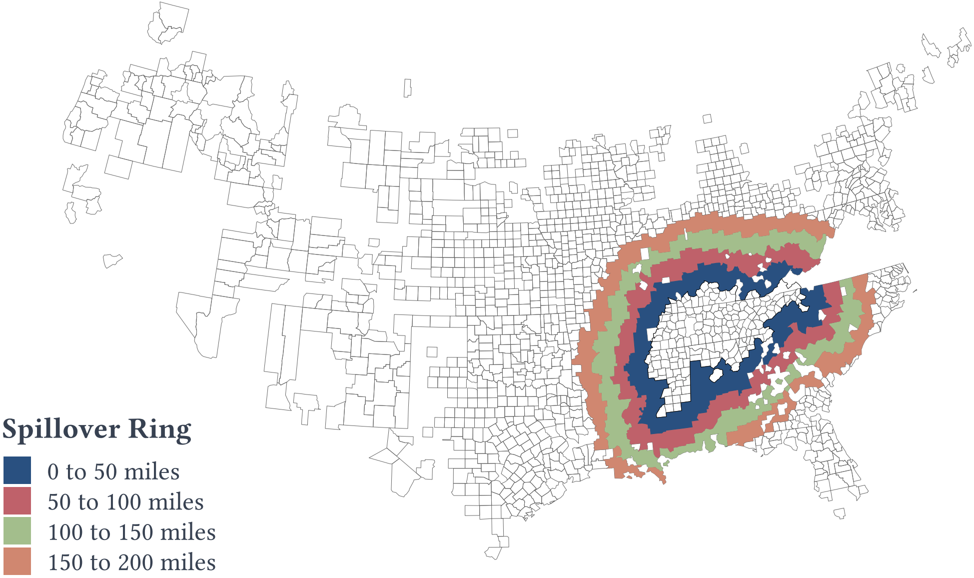
- **Agriculture:**
  - → Employees might be drawn to hire wages for new manufacturing jobs in Tennessee Valley (negative spillover on control units)
- **Manufacturing:**
  - → Cheap electricity might be available to nearby units (positive spillover on control units)
  - → Manufacturing jobs that would 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

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

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

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


Spillover Ring
0 to 50 miles
50 to 100 miles
100 to 150 miles
150 to 200 miles

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 emp.	-0.0514^{***}	-0.0739^{***}	-0.0371^{***}	-0.0164	-0.0298^{***}	-0.0157^{*}
	(0.0114)	(0.0142)	(0.0002)	(0.0114)	(0.0096)	(0.0088)

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

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 emp.	-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 emp.	0.0560^{***}	0.0350	-0.0203^{***}	-0.0245	-0.0331^{*}	-0.0296^{**}
	(0.0161)	(0.0218)	(0.0006)	(0.0282)	(0.0189)	(0.0142)

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

Identification Strategies and Place-Based Policies

The literature on federal Enterprise Zones, a place-based policy that gives tax breaks to businesses that locate within the boundary, has found conflicting results using different identification strategies:

- 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.

Identification Strategies and Place-Based Policies

Revisit analysis of 2017 federal Opportunity Zones on rental prices from Chen, Glaeser, and Wessel (2023):

- They use both strategies and find different magnitudes of effects

To explain the differences, I use the not-selected group as the control group and include two rings for being within 1/2 mi. and being between 1/2 and 1 mi. of opportunity zone

Effects of Opportunity Zones on Annual Home Price Growth

Control Group:	**Not-Selected**	**Neighboring**	**Not-Selected**
	(1)	(2)	(3)
Treat × Post	0.3033*	0.6478***	0.1788
	(0.1661)	(0.2457)	(0.1692)
< 1/2mi. × Post			-1.057***
			(0.3618)
< 1mi. × Post			-0.7430***
			(0.1922)

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

Community Health Centers

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

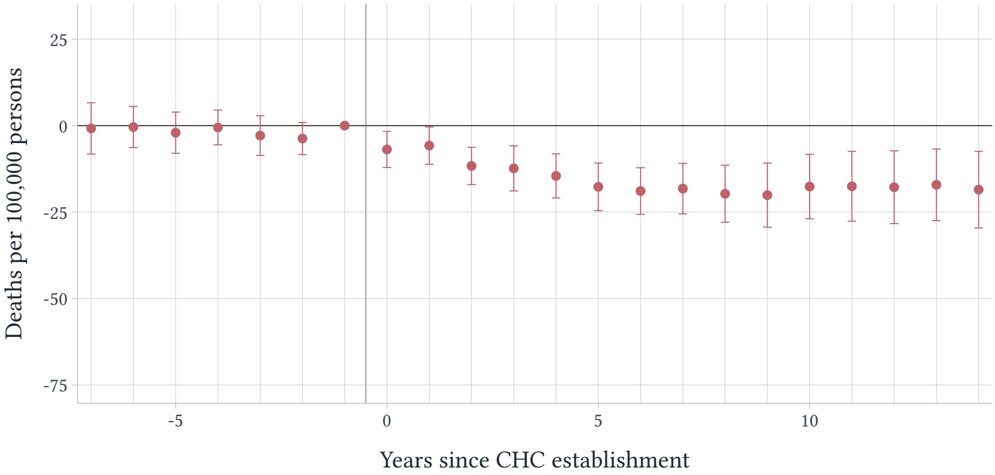
Research Question:

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

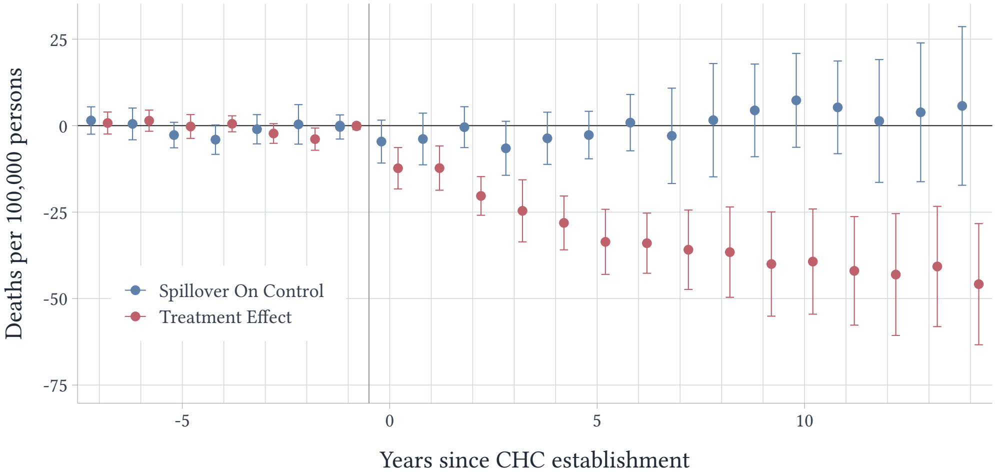
In the paper, I extend the imputation estimator for staggered treatment-timing (Borusyak, Jaravel, and Spiess, 2024; Gardner, 2021)

- Show multiple 'treatment' problem from Goldsmith-Pinkham, Hull, and Kolesár (2022) is not a problem with an imputation estimator

Effects of Establishment of Community Health Centers



Direct and Spillover Effects of Community Health Centers



Theory

Application in Urban Economics

Conclusion

Conclusion

- Discuss non-parametric identification of treatment effects in the presence of spatial spillovers
  - → Switching Effect requires parameterization of exposure mapping
  - → Total Effect is identified under minimal assumptions
- 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
- Extend estimator to a 'modern' event-study estimator that is robust to treatment effect heterogeneity

References I

Bailey, Martha J. and Andrew Goodman-Bacon (Mar. 2015). “The War on Poverty’s Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans”. *American Economic Review* 105.3, pp. 1067–1104. ISSN: 0002-8282. DOI: 10.1257/aer.20120070.

Baum-Snow, Nathaniel and Fernando Ferreira (2015). “Causal Inference in Urban and Regional Economics”. *Handbook of Regional and Urban Economics*. Vol. 5. Elsevier, pp. 3–68. ISBN: 978-0-444-59533-1. DOI: 10.1016/B978-0-444-59517-1.00001-5.

Berg, Tobias and Daniel Streitz (2019). *Handling Spillover Effects in Empirical Research*. Working Paper, p. 59.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess (2024). “Revisiting Event Study Designs: Robust and Efficient Estimation”. *Review of Economic Studies*. Forthcoming.

References II

Busso, Matias, Jesse Gregory, and Patrick Kline (Apr. 2013). "Assessing the Incidence and Efficiency of a Prominent Place Based Policy". *American Economic Review* 103.2, pp. 897–947. ISSN: 0002-8282. DOI: 10.1257/aer.103.2.897.

Butts, Kyle (2023). "JUE Insight: Difference-in-differences with geocoded microdata". *Journal of Urban Economics* 133, p. 103493.

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna (July 2021). "Difference-in-Differences with a Continuous Treatment". *arXiv:2107.02637 [econ]*. arXiv: 2107.02637.

Chen, Jiafeng, Edward Glaeser, and David Wessel (2023). "JUE Insight: The (non-) effect of opportunity zones on housing prices". *Journal of Urban Economics* 133, p. 103451.

Clarke, Damian (2017). "Estimating Difference-in-Differences in the Presence of Spillovers". *Munich Personal RePEc Archive*, p. 52.

Conley, T.G. (1999). "GMM estimation with cross sectional dependence". *Journal of Econometrics* 92.1, pp. 1–45. ISSN: 0304-4076. DOI: https://doi.org/10.1016/S0304-4076(98)00084-0.

Cuberes, David, Klaus Desmet, and Jordan Rappaport (Mar. 2021). "Urban Growth Shadows". *Journal of Urban Economics*, p. 103334. ISSN: 00941190. DOI: 10.1016/j.jue.2021.103334.

Fujita, Masahisa, Paul R Krugman, and Anthony Venables (2001). *The spatial economy: Cities, regions, and international trade.* MIT press.

Gardner, John (2021). *Two-Stage Difference-in-Differences.* Tech. rep.

Goldsmith-Pinkham, Paul, Peter Hull, and Michal Kolesár (2022). *Contamination bias in linear regressions.* Tech. rep. National Bureau of Economic Research.

References IV

Kline, Patrick and Enrico Moretti (Feb. 2014). "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority". *The Quarterly Journal of Economics* 129.1, pp. 275–331. DOI: 10.1093/qje/qjt034.

Leung, Michael P (2020). "Treatment and spillover effects under network interference". *Review of Economics and Statistics* 102.2, pp. 368–380.

Miguel, Edward and Michael Kremer (Jan. 2004). "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities". *Econometrica* 72.1, pp. 159–217. ISSN: 0012-9682, 1468-0262. DOI: 10.1111/j.1468-0262.2004.00481.x.

Neumark, David and Jed Kolko (July 2010). "Do enterprise zones create jobs? Evidence from California's enterprise zone program". *Journal of Urban Economics* 68.1, pp. 1–19. ISSN: 00941190. DOI: 10.1016/j.jue.2010.01.002.

References V

Sävje, Fredrik, Peter M. Aronow, and Michael G. Hudgens (Apr. 2021). "Average treatment effects in the presence of unknown interference". *The Annals of Statistics* 49.2. ISSN: 0090-5364. DOI: 10.1214/20-AOS1973.

Sobel, Michael E (Dec. 2006). "What Do Randomized Studies of Housing Mobility Demonstrate?: Causal Inference in the Face of Interference". *Journal of the American Statistical Association* 101.476, pp. 1398–1407. ISSN: 0162-1459, 1537-274X. DOI: 10.1198/016214506000000636.

Vazquez-Bare, Gonzalo (2023). "Identification and estimation of spillover effects in randomized experiments". *Journal of Econometrics* 237.1, p. 105237.

Verbitsky-Savitz, Natalya and Stephen W. Raudenbush (Jan. 2012). "Causal Inference Under Interference in Spatial Settings: A Case Study Evaluating Community Policing Program in Chicago". *Epidemiologic Methods* 1.1. ISSN: 2161-962X. DOI: 10.1515/2161-962X.1020.

Event Study

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) 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 $\hat{\mu}_t$.

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

$$y_{it} = \mu_i + \mu_t + \tau d_{it} + \tau_{\text{spill}} S_i (1 - D_i) 1(t = 1) + \varepsilon_{it}$$

Adjust two-step approach:

Controlling for Spillovers in Staggered Treatment Timing

$$y_{it} = \mu_i + \mu_t + \tau d_{it} + \tau_{\text{spill}} S_i (1 - D_i) 1(t = 1) + \varepsilon_{it}$$

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 $S_i 1(t = 1) = 0$). Then subtract off $\hat{\mu}_i$ and $\hat{\mu}_t$.
2. Then, regress $\tilde{y}_{it}$ on $\tau d_{it} + \tau_{\text{spill}} S_i * (1 - D_i) 1(t = 1)$ (or interacted event study leads/lags). This estimate is unbiased because $D_i 1(t = 1)$ is non-residualized (standard errors require adjusting).