Difference-in-Differences Estimation with Spatial Spillovers

Kyle Butts

Univ. of Colorado, Boulder

May 3, 2021

ABSTRACT

Empirical work often uses treatment assigned following geographic boundaries. When the effects of treatment cross over borders, classical difference-in-differences estimation produces biased estimates for the average treatment effect. In this paper, I introduce a potential outcomes framework to model spillover effects and decompose the estimate's bias in two parts: (1) the control group no longer identifies the counterfactual trend because their outcomes are affected by treatment and (2) changes in treated units' outcomes reflect the effect of their own treatment status and the effect from the treatment status of "close" units. I propose estimation strategies that can remove both sources of bias and semi-parametrically estimate the spillover effects themselves. I extend Callaway and Sant'Anna (2020) to allow for event-study estimates that control for spillovers. To highlight the importance of spillover effects, I revisit analyses of three place-based interventions.

1

1 — Introduction

Empirical work often considers settings where policy is targeted to groups of units by geographic boundaries, but the effect of these treatments spillover onto 'nearby' units.[1] Place-based policies, programs that are targeted to specific locations, are known to have general-equilibrium effects that extend far past the area intended to receive treatment (Kline and Moretti, 2014b). Spillover effects can occur on to both control and treated units. First, there can be spillover effects onto control units. For example, individuals in control areas can travel to the treated areas and receive treatment (e.g. a hospital opening serves nearby residents) or changes in the economy in a treated area can have feedback loops that affect nearby economies (e.g. a new factory opening increases service sector spending). Second, there can be spillover effects onto other treated units. For example, forces of agglomeration can increase effects as treatment concentration increases (e.g. knowledge-sharing between workers), forces of congestion or market competition can decrease the effects of treatment (e.g. bidding up of wages cause employment effects of factory openings to be mitigated), and treatment effectiveness may be improved by multiple jurisidictions learning from one another.

In this paper, I introduce a potential outcome framework that directly models spillover effects. Using this framework, the central theoretical result of my paper is that in the presence of spillovers, the standard difference-in-differences estimate identifies the direct effect of treatment (the estimand of interest) plus two additional bias terms resulting from the spillovers.[2] The intuition for why the two forms of spillovers cause bias is as follows.

First, untreated units that are 'close' to treated units experience effects of treatment and therefore these 'control' units fail to identify the counterfactual trend. When estimating by difference-in-differences, the spillover onto the 'close' control units is averaged into the untreated units' change in outcomes. In this case, the spillover is subtracted from the estimated treatment effect and biases the estimate in the opposite sign of the spillover effect.

For example, consider trying to estimate employment effects of a factory opening in a county.

---

[1] The framework of this paper applies to any setting with a well-defined measures of distance, e.g. geographic distance, economic distance such as supply chains, node distance in a graph, or social relationships in schools or cities.

[2] The term 'direct' effect can refer to either the average treatment effect on the treated or the intent to treat effect. In empirical applications, the 'direct' effect is sometimes called the 'partial equilibrium' effect or the 'local' effect.

Agglomeration economies suggest that neighboring counties would also benefit from knowledge spillovers and improved access to supply chain which would potentially increase economic activity (Duranton and Puga, 2003). The treatment effect estimate is negatively biased because the change in outcome in neighboring counties is higher than it would be absent treatment. Researchers would therefore underestimate the positive benefits of treatment.[3]

Second, changes in treated units' outcomes reflect the effect of their own treatment status and the effect from the treatment status of "close" units.The spillover onto other treated units is averaged into the treated units' change in outcomes. Therefore the bias of treatment effect is the same sign as the sign of the spillover. Continuing with our example, two factory openings in neighboring counties might cause the benefit of each individual factory to increase due to agglomeration forces. The estimated treatment effect will count the 'direct' effect of the factory opening as well as the benefit of the agglomeration economies from nearby factory openings. Therefore the treatment effect will be positively biased due to the positive spillover onto also treated units.

I then use my potential outcome framework to evaluate commonly used empirical strategies found in applied work and provide practical recommendations for researchers. The most common approach is to remove potentially contaminated control units from the sample in order to remove spillover effects on control units (Berg and Streitz, 2019). This approach is not recommended for two reasons. First, the removal of too few control units will leave bias and removal of too many will lower precision of the treatment effect estimate. This results in a bias-variance trade-off that other approaches do not share. Second, this method does not remove spillover effects on treated units and will leave this bias term in the difference-in-differences estimate. Another common approach is to include either an indicator or a set of rings around treated units. This approach will remove all bias from the spillover effects on control units provided that all control units with spillover effects are included in the indicator. Similar to the previous approach, spillover effects on also treated units still add bias to estimates of the direct effect.

Therefore, I propose an estimation strategy that will remove both sources of bias with very

_____

[3] If the policy maker had a welfare function that included both treated and neighboring counties, they would fail to count the positive benefits of the control units and they would underestimate the positive benefits on the treated units. Difference-in-differences estimation would therefore result in double undercounting.

minimal assumptions on the structure of spillovers. To remove *all bias* from the direct effect estimate, a researcher needs to just include an indicator for being close to a treated unit interacted with treatment status, so long as the indicator captures all units affected by spillovers. In my decomposition of the treatment effect estimate, the bias terms are the *average* spillover effects onto treated and control units. An indicator variable interacted with treatment status therefore can estimate these two averages and remove them from the difference-in-differences estimate. Importantly, my method does not require researchers to make any assumptions about how spillover effects propagate across space. The only assumption required is the maximum distance from treated units spillovers can occur.[4]

Since spillover effects are often important causal effects themselves, I next turn to how to estimate spillovers directly. I show with evidence from Monte Carlo simulations, that commonly used semi-parametric estimation strategies capture spillovers well (in a mean square prediction error sense). This involves creating a set of distance bins from treated units (e.g. being 0-20 miles, 20-40 miles, 40-60 miles from treated unit) and interacting them with a treatment indicator.[5] The key consideration required by a researcher is whether the size of spillover effects are additive in the number of nearby treated units or not. For non-additive spillovers, I recommend a set of mutually-exclusive indicators which measures which distance bin that the *closest* treated unit falls within for a given observation. For additive spillovers, I recommend using the *number of* treated units within each distance bin for a given observation.[6]

To show the importance of considering spillover effects, I revisit analyses of place-based policies in urban economics in Section 3.4. I revisit the analysis of the Tennessee Valley Authority by Kline and Moretti (2014a). The Tennessee Valley Authority was a large scale New Deal program that created many new dams for flood-protection and navigation. The construction of large-scale dams lowered the cost of power for industrial firms giving the region an industrial boost (Kitchens,

---

[4] Even if the maximum distance is not large enough, since treatment effects typically decay over distance, most of the bias will still be removed. Including too many units in the indicator will increase the variance in the estimates for average spillover effects as the other control units identify the parallel counterfactual unit.

[5] The choice of the distance bins depends on the economic context and in particular the source of the spillovers.

[6] While the non-additive version of ring indicators will remove all bias from the direct effect of treatment even under misspecification, this is not true of additive rings. Therefore, researchers may want to estimate the direct effect of treatment using non-additive rings and then use additive rings to estimate spillover effects.

2014). The scale of federal investment in the region was large and the pro-manufacturing benefits likely spread further than the Authority's boundary due to the electrification infrastucture and agglomeration economies (Severnini, 2014). I show that estimation by difference-in-differences fails to account for these spillovers and therefore forms biased estimates of the local effect of the Tennessee Valley Authority.

Following this empirical application, I discuss how my framework fits into a larger discussion on identification strategies with place-based policies. I revisit conflicting results about United States' federal Empowerment Zones, a program that creates incentives for businesses to locate in high-poverty neighborhoods. Busso et al. (2013) use Census Tracts from qualified but ultimately rejected applications that are typically far away from accepted Zones as a comparison group and find that significant reductions in poverty. This empirical strategy is based on the idea that since they also qualified for the program, these comparison units would be on parallel trends without being contaminated by spillover effects from treatment. Neumark and Kolko (2010), on the other hand, use census tracts within 1,000 feet of the Zone and find no statistically significant effects. This empirical strategy assumes that the borders are drawn somewhat randomly and therefore being just outside is as good as random. Although, using close units is potentially problematic as they can experience spillover effects from the Zones. My framework can explain the differences in findings if there are positive spillover effects onto census tracts just outside Empowerment Zones.

Last, in Section 4, I extend estimation of the direct effect and spillover effects of treatment into the event study framework by extending the work of Callaway and Sant'Anna (2020). This allows for a very common setting in which treatment turns on for different units in different periods. I first show how to adjust their methods to control for spatial spillovers in estimation of treatment effects. Then, I show how to estimate average spillover effects. My paper is the first paper to study estimation of treatment effects in an event-study framework in the presence of spillover effects.

I demonstrate the method by revisiting the analysis by Bailey and Goodman-Bacon (2015) of Community Health Centers which provided low-cost primary care to impoverished areas. They find that the health centers significantly lowered the mortality rate in treated counties. I find that in this context, since the mortality declines caused by the health centers was due to primary care, spillover effects are near-zero. This result shows the importance of accessability concerns when deciding the location of the health centers.

1.1. Relation to the Literature

There is a large literature on estimation of treatment effects in the presence of spillovers using a 'partial identification' framework where units are in distinct treatment clusters and outcomes depend on the treatment status within the observation's cluster only.[7] Estimation compares units in the partially treated clusters with control units in completely untreated clusters which do not receive spillover effects. This allows standard difference-in-differences estimation of both the direct effect (treatment effect on the treated) and spillover effects (treatment effect on the untreated in the treated clusters).

There is a nascent literature exploring estimation of direct and spillover effects which does not require a completely untreated cluster by using a potential outcomes framework. Vazquez-Bare (2019) presents a potential outcomes model that explicitly accounts for "within-group" spillovers in experiments which allows for seperate estimation of direct effect and spillover effects by a simple differences-in-means. I extend this work by focusing on difference-in-differences estimation in non-experimental settings. Sävje et al. (2019) consider estimation of treatment effects by difference-in-differences, but they define treatment effects as the combination of direct effects and spillover effects. My paper develops a strategy to allow researchers to seperately identify treatment effects and spillovers. These estimates can be combined to estimate their definition of average treatment effect. The further advantage is that 'net' treatment effects can be estimated at different levels of exposure which is more relevant for future policy decisions.

I also contribute to the literature that focuses on estimation of treatment effects with spatial spillovers using the difference-in-differences framework (Clarke, 2017; Berg and Streitz, 2019; Verbitsky-Savitz and Raudenbush, 2012; Delgado and Florax, 2015). I contribute to this literature in two ways. First, my paper derives an explicit form for this bias in terms of general potential outcomes which allows my results to capture many different forms of spillovers. For the above papers, if I assume the particular functional forms for potential outcomes, I arrive at the same bias equation as theirs if they have one derived explicitly. Second, my paper also advances the

---

[7] Angelucci and Di Maro (2016) provides an overview of estimation of treatment effects in the presence of "within-group" spillovers. Examples in the literature include: Halloran and Struchiner (1995) consider community-vaccine rates in epidemiology; Miguel and Kremer (2004) consider deworming programs in Kenyan schools; Sobel (2006) considers interference in the Moving to Opportunity Program; and Angrist (2014) studies the context of school peer effects.

literature by considering estimation of direct effects and spillover effects in event-study framework which allows for the common occurence of staggered treatment adoption.

The rest of the paper is structured as follows. Section 2 presents the potential outcomes framework, defines the estimand of interest, and shows the resulting bias from estimating a classical difference-in-differences model. Section 3 focuses on estimation of the direct and spillover effects of treatment. I evaluate currently used solutions in the literature and present a set of recommendations to more robust estimation strategies. Section 3.4 presents an application for evaluating the Tennessee Valley Authority program. Section 4 discusses briefly how to incorporate spillovers into event-study estimation following the insights from Callaway and Sant'Anna (2020). I apply these methods in evaluating spillover effects of U.S. Community Health Centers in Section 4.2.

2 — Potential Outcomes Framework

This section will present a potential outcome framework that formalizes spatial spillovers. Following the canonical difference-in-differences framework, there are T periods and treatment turns on for all treated units at some period t_0 and remain on afterwards.[8] Potential outcomes, denoted $Y_{i,t}(D_i, h(\vec{D}, i))$, for unit i at time t are a function of own treatment-status D_i and, departing from the standard potential outcomes framework, of a function of the entire vector of treatment assignments $h(\vec{D}, i)$ where $\vec{D} \in \{0, 1\}^n$ denotes the n-dimensional vector of all unit treatments. The function $h(\vec{D}, i)$ is referred to as an 'exposure mapping' and is a non-negative scalar- or vector- valued function. The exposure mapping measures the intensity at which unit i is affected by spatial spillovers. When unit i is sufficiently 'far' away, it has no exposure to spatial spillovers and $h(\vec{D}, i) = \vec{0}$. The exposure mapping formalizes the SUTVA violation by summarizing how outcomes are affected by other unit's treatment assignment. To help better understand the exposure mapping function, I give three examples that are commonly used in the literature.

Example 1. *First, $h(\vec{D}, i)$ can be an indicator variable that equals one only if there is a treated unit within $\bar{d}$ miles of unit i.[9] Let $d(i, j)$ be a distance measure which tells the distance unit i is*

---

[8] This framework is extended to staggered treatment adoption in Section 4.

[9] Similarly a dummy for counties that share contiguous borders is commonly used.

from unit j. In this case

$$h(\vec{D}, i) = \max_{i \neq j} D_j * \mathbf{1}[d(i,j) < \bar{d}] \tag{1}$$

This exposure mapping is a realistic specification when it is assumed that spillover effects do not decay over distance until $\bar{d}$ and the intensity of spillovers does not depend on the number of neighboring units treated. For example, this exposure mapping likely applies in the context of new library creation (Berkes and Nencka, 2020). In this case the distance $\bar{d}$ would be the maximum distance people would travel to a nearby library. The benefits of access to a neighboring town library does not depend on whether people can access 1 or more nearby libraries, so the binary exposure mapping is a good approximation to spillovers in this context.

Example 2. *Second, $h(\vec{D}, i)$ can be a function that equals the number of units treated within distance $\bar{d}$, i.e.*

$$h(\vec{D}, i) = \sum_j D_j * \mathbf{1}[d(i,j) < \bar{d}]. \tag{2}$$

This exposure mapping is no longer binary, so the intensity of spillovers is additive in the number of nearby units treated. In the context of large store openings, this exposure mapping captures the additive nature of agglomeration economies. As more nearby counties receive new stores, the force of agglomeration and therefore the spillover effect increases (e.g. Basker (2005)).

Example 3. *Last, $h(\vec{D}, i)$ is a spatial decay function where exposure decreases with distance. In this case, spillover intensity is the sum across all treated observations' decay term, i.e.*

$$h(\vec{D}, i) = \sum_{j \neq i} D_j e^{-\alpha d(i,j)}. \tag{3}$$

This exposure mapping allows for the intensity of spillovers to depend on distance to treatment and also is additive in the number of nearby units treated.[10] The speed that spillover effects decay over distance depends on the parameter α that can be calibrated by the researcher or estimated using a non-linear least squares routine. In the literature on R&D investment, for example, Keller (2002) uses a modified version of this exposure mapping where D_j is country j's R&D expenditure. This

---

[10] However, this specification assumes that all units are affected by all other units. This creates problems with inference because it implies potential correlation between all units' error terms. For this reason, this function often is summed over only the k-nearest neighbors or over units within $\bar{d}$ miles.

specification features exponential decay over distance which captures the theoretical insight that research spending in closer locations has a larger effect than further away spending.

After chosing an exposure mapping that makes sense in the economic context, a researcher must also specify the functional form of the potential outcomes. Typically, the exposure mapping enters into the regression linearly and potentially the coefficient is allowed to differ by treatment status to reflect that spillovers on control and treated units can be different phenomenon.

2.1. Spatial Spillovers

With the potential outcomes defined, I define 'spillover effects' as:

$$Y_{i,t}(D_i, h(\vec{D}, i)) - Y_{i,t}(D_i, \vec{0}).$$

The spillover effect measures the difference in potential outcomes between being exposed at intensity $h(\vec{D}, i)$ and not being exposed. This effect can differ for either treated or control units as the magnitude or even the nature of spillovers might differ between treated and untreated units. Then, the average spillover effect onto control/treated units averages over potential heterogeneity in the effect size of spillovers and over heterogeneity in exposure intensity $h(\vec{D}, i)$:

$$\tau_{\text{spill,control}} \equiv \mathbb{E}\left[Y_{i,t}(D, h(\vec{D}, i)) - Y_{i,t}(D, \vec{0}) \mid D_i = D\right].$$

To emphasize, the average spillover effect onto conrol/treated units averages over each unit's exposure mapping which can be heterogeneous across units. For example, spillover effects can decay over distance as in Equation (3) and the average spillover effect averages across different exposure intensities.

It is important to clarify what I am assuming is the estimand of interest researchers would like to estimate when using difference-in-differences because there are multiple potential estimands defined in the literature. I assume that what the 'average treatment effect' is trying to measure in this context is what I will call the 'direct effect of treatment':

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

which measures the effect of being treated in the absence of exposure to spillovers.

My definition of the direct effect of treatment differs from Sävje et al. (2019) where they define the average treatment effect as

$$\mathbb{E}\left[Y_{i,t}(1, h(\vec{D}, i)) - Y_{i,t}(0, h(\vec{D}, i)) \mid D_i = 1\right],$$

where the expectation is over individuals and over different levels of exposures. The difference between their treatment effect and mine becomes more clear by adding and subtracting terms and rearranging:

$$
\begin{aligned}
Y_{i,t}(1, &h(\vec{D}, i)) - Y_{i,t}(0, h(\vec{D}, i)) \\
&= Y_{i,t}(1, h(\vec{D}, i)) - Y_{i,t}(1, 0) + Y_{i,t}(1, 0) - Y_{i,t}(0, 0) + Y_{i,t}(0, 0) - Y_{i,t}(0, h(\vec{D}, i)) \\
&= \underbrace{Y_{i,t}(1, 0) - Y_{i,t}(0, 0)}_{\text{Direct Effect}} + \underbrace{Y_{i,t}(1, h(\vec{D}, i)) - Y_{i,t}(1, 0)}_{\text{Spillover on Treated}} - \underbrace{(Y_{i,t}(0, h(\vec{D}, i)) - Y_{i,t}(0, 0))}_{\text{Spillover on Control}}
\end{aligned}
$$

The individual treatment effect Sävje et al. (2019) use is the direct effect of treatment plus the difference between spillover on treated units and spillover on control units.

I prefer my definition of the direct effect because it allows for seperate identification of the direct effect of treatment and the spillover effects themselves. The alternative estimand averages over exposure intensities and therefore is difficult to be used by a policy maker in a particular location. That is, a county deciding whether to implement a policy will want to consider the treatment effect plus the difference in spillover effects at *a particular* exposure intensity. The estimand proposed by Sävje et al. (2019) returns the average of spillovers across levels of exposure and can not be used to predict spillovers at specific exposure intensities while my preferred method of identifying the three effects seperately can. Estimates of the direct effect and spillover effect can be later combined to estimate the average treatment effect as defined by Sävje et al. (2019).

2.2. Bias in Difference-in-Differences Estimation

In this section, I identify the two sources of bias in difference-in-differences estimation of the direct effect of treatment. Researchers typically estimate the canonical two-way fixed effects model,

$$y_{i,t} = \tau D_{i,t} + \mu_i + \mu_t + \epsilon_{i,t}, \tag{4}$$

where $D_{i,t} = D_i \mathbf{1}(t = 1)$. The estimator $\hat{\tau}$ is a biased estimate for τ_{direct} in the presence of spillovers. To show this, I first present the equivalent to the parallel counterfactual trends assumption in the context of the new potential outcome framework.

Assumption 1 (Parallel Counterfactual Trends).

$$\mathbb{E}\left[\bar{Y}_{i,Post}(0,\vec{0}) - \bar{Y}_{i,Pre}(0,\vec{0}) \mid D_i = 1\right] = \mathbb{E}\left[\bar{Y}_{i,Post}(0,\vec{0}) - \bar{Y}_{i,Pre}(0,\vec{0}) \mid D_i = 0\right],$$

where $\bar{Y}_{i,W}$ is the average outcome in either the post- or pre-periods.

This assumption states that in the absence of treatment and with zero exposure (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. This generalizes to the classic parallel counterfactual trends when SUTVA is satisfied because then every unit has zero exposure.

Given that Assumption 1 holds, the estimate $\hat{\tau}$ from (4) can be decomposed as the direct effect and the two sources of spillover bias. The proof is given in Appendix A.

Proposition 1 (Decomposition of Difference-in-Differences Estimate).
If Assumption 1 holds, the expectation of the estimate $\hat{\tau}$ from (4) is

$$\mathbb{E}[\hat{\tau}] = \underbrace{\mathbb{E}\left[\bar{Y}_{i,Post} - \bar{Y}_{i,Pre} \mid D_i = 1\right] - \mathbb{E}\left[\bar{Y}_{i,Post} - \bar{Y}_{i,Pre} \mid D_i = 0\right]}_{\text{Difference-in-Differences}}$$

$$= \mathbb{E}\left[\bar{Y}_{i,Post}(1,\vec{0}) - \bar{Y}_{i,Post}(0,\vec{0}) \mid D_i = 1\right] + \mathbb{E}\left[\bar{Y}_{i,Post}(1,h(\vec{D},i)) - \bar{Y}_{i,Post}(1,\vec{0}) \mid D_i = 1\right]$$

$$\quad - \mathbb{E}\left[\bar{Y}_{i,Post}(0,h(\vec{D},i)) - \bar{Y}_{i,Post}(0,\vec{0}) \mid D_i = 0\right]$$

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

The intuition behind the biases are as follows. First, the change in outcomes among treated units combines the direct effect of treatment with the spillover effects caused by nearby treated units. Therefore the first difference adds the average spillover effect onto the treated units, $\tau_{\text{spill,treated}}$. Second, the change in outcomes among control units combines the parallel counterfactual trend with the average spillover effect onto control units. Since $\hat{\tau}$ is found by subtracting the change in outcomes among the control units, we subtract the average spillover effect onto the control, $\tau_{\text{spill,control}}$.

Readers that are judging estimates in the presence of spillovers can use the following heuristics to sign the bias. If they suspect spillovers onto control units, the bias is the opposite sign of the spillover. If they suspect spillovers onto treated units, the bias is the same sign of the spillover. The overall bias is the sum of the two.

3 — Estimating 2×2 Difference-in-Differences with Spillovers

3.1. Unbiased Estimation of Direct Effect

Under mild conditions, it is possible to control directly in a regression for the two sources of spillover effects and unbiasedly estimate the direct effect of treatment. A pair of indicators that contains all control/treated units respectively will estimate the average spillover effect on control/treated units and remove these sources of bias. The assumptions required to do this is that the researcher needs to be able to identify the maximum distance for which units are affected by spillovers and that there be unexposed units remaining.

Assumption 2 (Spillovers Are Local). *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, i, such that $\min_{j:\, D_j=1} d(i, j) > \bar{d}$.

Part (i) of Assumption (2) requires that spillovers are 'local' in that units are no longer exposed to spillovers after some maximum distance $\bar{d}$. Part (ii) of Assumption (2) requires that there exists control units with no exposure in order to estimate the counterfactual trend and treated units with no exposure in order to estimate the direct effect. Let $S_i = \mathbf{1}(\min_{j:\, D_j=1} d(i, j) \leq \bar{d})$ be an indicator for being within $\bar{d}$ miles of the closest treated unit after treatment.

In order to remove both sources of bias, we modify the two-way fixed effects model to include this indicator interacted with treatment status:

$$y_{i,t} = \tau D_{i,t} + \tau_{\text{spill,control}}(1 - D_{i,t})S_i + \tau_{\text{spill,treated}} D_{i,t} S_i + \mu_i + \mu_t + \epsilon_{i,t}. \tag{5}$$

Proposition 2 (Unbiased Estimate for Direct Effect). *If Assumption 1 holds, the expectation of the estimate $\hat{\tau}$ from (5) is*

$$\mathbb{E}\left[\hat{\tau}\right] = \mathbb{E}\left[\bar{Y}_{i,Post} - \bar{Y}_{i,Pre} \mid D_i = 1, S_i = 0\right] - \mathbb{E}\left[\bar{Y}_{i,Post} - \bar{Y}_{i,Pre} \mid D_i = 0, S_i = 0\right]$$

$$= \tau_{direct}$$

The expectation of the two spillover effect estimates are

$$\mathbb{E}\left[\hat{\tau}_{spill,conrol}\right] = \mathbb{E}\left[\bar{Y}_{i,Post}(0, h(\vec{D}, i)) - \bar{Y}_{i,Post}(0, \vec{0}) \mid D_i = 0, S_i = 1\right]$$

and

$$\mathbb{E}\left[\hat{\tau}_{spill,treated}\right] = \mathbb{E}\left[\bar{Y}_{i,Post}(1, h(\vec{D}, i)) - \bar{Y}_{i,Post}(1, \vec{0}) \mid D_i = 1, S_i = 1\right]$$

The intuition behind this is that the indicator interacted with treatment status serve as unbiased estimates of the *average* spillover effect onto treated and control units and therefore remove these terms from the direct effect estimate, $\hat{\tau}$.[11] This proposition shows that a researcher does not need to know how the spillovers occur over space in order to have an unbiased estimate for treatment effects. An indicator for being 'near' treatment ineracted with treatment status is enough to have an unbiased estimate for the treatment effect.[12]

A natural question is why wouldn't researchers make the distance really large in order to guarantee they have an unbiased estimate of τ_{direct}? The problem with this is that control units with $S_i = 0$ identify the counterfactual trend and treated units with $S_i = 0$ identify the direct effect. Therefore leaving few units with $S_i = 0$ will yield more variable estimates. On the other hand, having units that experience spillovers not included in the S_i indicator will leave some bias in the estimate. Although, since spillover effects typically will grow weaker over distance, the bias may not be particularly large if distant units are mistakenly treated as if they have zero exposure. Therefore there is a bias-variance trade-off in extending the extent of S_i that should be balanced by researchers guided by their particular economic context.

---

[11] The proof uses standard Frisch-Waugh-Lovell style logic on a set of indicator variables and the fact that if $S_i = 0$ then $h(\vec{D}, i) = \vec{0}$. The derivation follows closely the work of Clarke (2019).

[12] In contexts with different measures of 'distance', note that this result only requires an indicator that captures all affected units. For example, thinking about spillovers between industries, this requires the researcher to correctly identify which industries are affected by spillovers and which are not.

A common approach in the literature is to drop control units adjacent to treated units to avoid the problems of spillovers in identification. Proposition 1 and Assumption 2 would suggest that this can be effective in removing (at least part of) the average spillover effect on control units. However, since researchers do not know how far the spillovers extend there is difficulty in implementing this method effectively. On the one hand, removal of too few units will keep some control units that are experiencing spillover effects which would leave the estimated treatment effect biased. On the other hand, removing too many units will leave few observations and therefore increase the variance of the estimates. There is therefore a bias-variance trade off inherent to this method that can be easily improved by creating an indicator for these units rather than dropping them from the sample. In addition to the unnecessary loss of precision induced by this method, the removal of problematic control observations, will not remove spillover effects onto treated units. Therefore if treated units are affected by the treatment status of nearby units, then the spillover effect on treated units which would remain as a source of bias.[13]

The second part of proposition 2 shows that equation (5) can also be used to estimate average spillover effects. The spillover effects themselves are potentially relevant for policy makers. For example, if the policy maker cares about treated areas and neighboring areas, then they should consider if the direct effect comes at a cost to nearby areas (negative spillovers) or if the benefits are being undercounted by only considering the treated area (positive spillovers). Similarly, a policy maker trying to determine the net benefits for a treated location must consider the direct effect and the spillover effect received by treated units. This could vary by location if the size of benefits created by the treatment depend on the number of treated units nearby.

However, the estimates are not trivial to interpret as the "typical" magnitude of spillover effects. This is because $\hat{\tau}_{\text{spill,control}}$ and $\hat{\tau}_{\text{spill,treat}}$ average spillover effects across all units with $S_i = 1$ which could contain units that experience no spillover effects. For example, if only units really close to treatment receive spillover effects and S_i contains a lot of units that experience no spillover effects, then $\hat{\tau}_{\text{spill, control}}$ could be estimated near zero even though there are spillover effects.

An improvement on this is to use a set of concentric ring indicators instead of a single large ring indicator to more flexibly estimate the spillover effects (e.g 0-20 miles, 20-40 miles and 40-60 miles from the nearest treated unit). Since the set of rings is collinear with a single large ring, the

---

[13] Appendix B contains a set of Monte Carlo simulations that highlight the problems more rigorously.

Figure 1 — Rings Improve Estimation of Spillover Effects

(a) Single Ring

(b) Multiple Rings
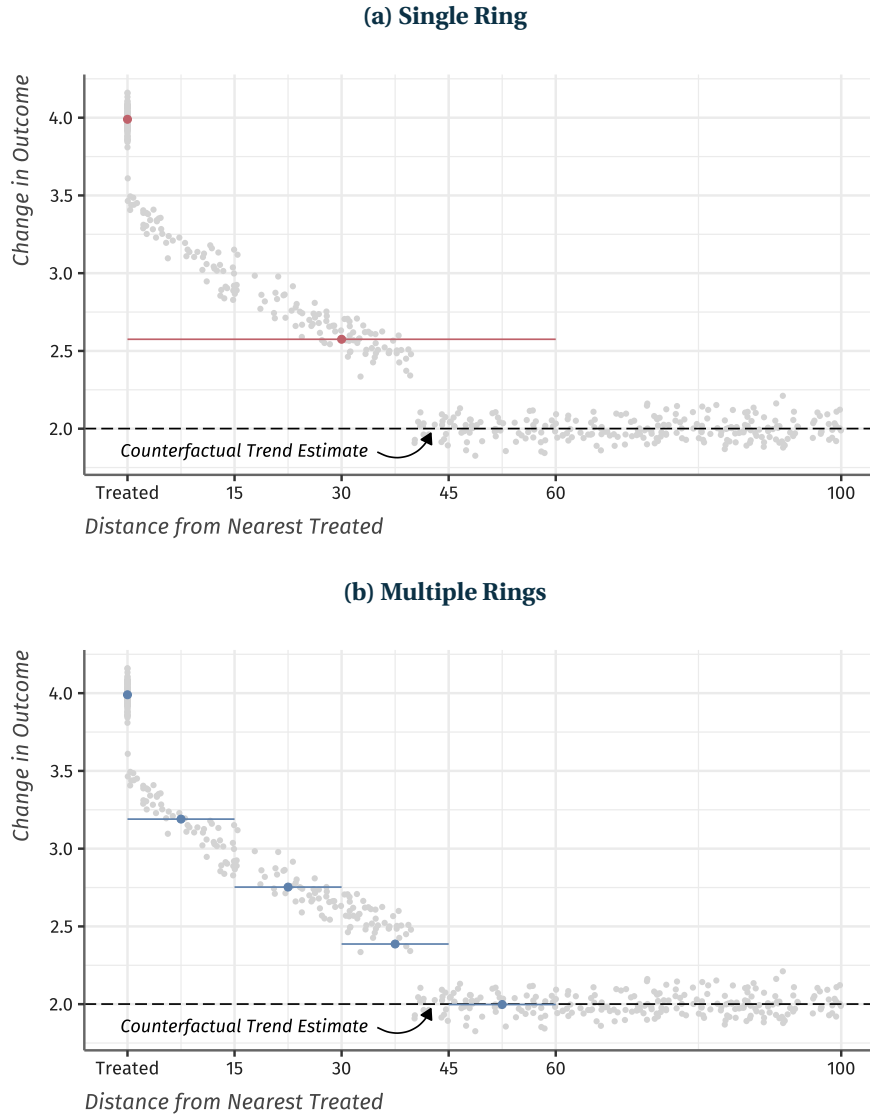
Notes: Panel (a) shows the estimate of a single large ring indicator following equation (5). Panel (b) shows a set of concentric ring indicators following equation (5) where S_i is replaced with a set of ring indicators.

direct effect estimate retains its unbiased properties, while spillover effects can potentially be estimated more accurately. A simple example of how rings improves estimation of spillover effects can be seen in Figure 1. The figure displays observations of $\bar{Y}_{i,Post} - \bar{Y}_{i,Pre}$ on the y-axis and the distance to the nearest treated unit on the x-axis and treated units are placed at a zero distance for visualization purposes. The example has a counterfactual trend of 2 so any difference from 2

is due to either treatment/spillover effects and the error term. Panel (a) uses a single indicator for being between 0 and 60 miles from treatment and the difference between the bar and the estimated counterfactual trend is the estimated spillover effect which in this case is about 0.6. This number masks that units within the first 15 miles have a treatment effect of 1 or larger and units between 40 and 60 miles have 0 spillover effects.

Figure (b) on the other hand uses 4 indicators for every 15 mile interval (0-15, 15-30, 30-45, and 45-60 miles). This improves on the estimate of a single ring in two ways. Multiple rings are better able to capture the evolution of spillover effects over distance. Furthermore, the ring that is included past the true maximum distance where spillover effects occur has a point estimate near zero. The rings method acts a semi-parametric estimator in the sense that as the number of rings increase and their width shrinks, the estimates would trace out the spillover effect function (subject to their being units outside of the rings). However, precision of estimates would decrease as the number of rings increase and therefore there is a bias-variance trade-off in the number and width of rings. Clarke (2017) proposes a cross-validation technique for optimally choosing the number of rings and their techniques that uses the out-of-sample mean square prediction error as a way to estimate the bias-variance tradeoff. Interested readers can see the paper for technical and implementation details.

One disadvantage of using a set of indicator for *the nearest* treated unit is that some spillover effects are additive in the number of nearby treated units. In this case, summarizing exposure by the distance to the closest treated unit fails to capture important information on the location of other treated units and the coefficients on the ring indicators might be interpreted incorrectly. One potential solution is to count the number of treated units within each ring instead of creating indicators for nearest treated unit. This additive version of rings however no longer removes all bias from the direct effect of treatment unless the exposure mapping is correctly specified. Therefore, if spillover effects are expected to be additive in nature, seperate estimation of the direct effect and spillover effects would be the preferred method.

3.2. Monte Carlo Simulations

3.2.1. Bias Under Misspecification of Spillovers

To illustrate the above discussion, I turn to a set of Monte Carlo simulations. The first exercise is to consider a set of different data generating processes and see how well estimates of the direct effect perform when controlling for spillovers using potentially misspecified exposure mappings. In the general form, I generate data with unit and time fixed effects and an error term that is uncorrelated with $D_{i,t}$ for many different exposure mappings $h(\vec{D}, i)$. I generate data for unit i is a US county at time $t \in \{1, \ldots, 20\}$ using the following data-generating process:

$$y_{i,t} = \mu_t + \mu_i + 2D_{i,t} + \beta_{\text{spill, control}}(1 - D_{i,t})h(\vec{D}, i) + \varepsilon_{i,t}, \tag{6}$$

where $\mu_t \sim N(0.2t, 0.1^2)$ and $\mu_i \sim N(6, 2^2)$ respectively and the error term is IID and distributed as $\varepsilon_{i,t} \sim N(0, 2^2)$.

For simplicity, I also remove spillovers onto treated units, but the results of which specifications perform best in the simulation are the same. In order to keep the magnitude of bias constant across specifications, I normalize the average spillover magnitude to be equal across specifications.[14] For each true data-generating process, I estimate (6) with the correct spillover and misspecified alternatives $\tilde{h}(\vec{D}, i)$.

The set of spillover specifications included in the data-generating process are 'Within $40/80$mi.' which equals 1 if the counties' center of population is within $40/80$ miles to a treated county's center of population. 'Within $40/80$mi. (Additive)' is the number of treated county's within $40/80$ miles of the control county. 'Decay' is given by $\max_j D_j e^{-0.02d(i,j)} * 1(d(i,j) < 80)$ and 'Decay (Additive)' is given by $\sum_{j \neq i} D_j \exp^{-0.02d(i,j)}$.

The last set of spillover specifications is commonly used in the literature as a semi-parametric estimator and are referred to as 'Rings'. These consist of a set of indicators for falling within distance bins from the nearest treated unit (e.g. indicators for being between 0-20, 20-40, 40-60, and 60-80 miles to closest treated unit). The last specification is an additive version of 'Rings' which is the number of treated units within each distance bin. This maintains a lot of the intuitive advantages of the 'Rings' specification but parameterizes the spillovers to be additive in the number of nearby treated units. The advantage of this, as seen below, is enhanced performance

---

[14] This results in a constant bias for TWFE estimates across data-generating processes.

<div align="center">

Table 1 — Bias from Misspecification of Spillovers

</div>

Specification	Data-Generating Process					
	Within 40mi	**Within 80mi**	**Within 40mi. (Additive)**	**Within 80mi. (Additive)**	**Decay**	**Decay (Additive)**
TWFE (No Spillovers)	0.263	0.263	0.263	0.263	0.263	0.263
	[0.091]	[0.091]	[0.091]	[0.091]	[0.091]	[0.091]
Within 40mi.	0.000	0.219	0.164	0.000	0.182	0.149
	[0.022]	[0.070]	[0.049]	[0.022]	[0.055]	[0.044]
Within 80mi.	0.000	0.000	0.000	0.000	0.000	0.000
	[0.024]	[0.024]	[0.024]	[0.024]	[0.024]	[0.024]
Within 40mi. (Additive)	0.049	0.227	0.179	−0.001	0.183	0.149
	[0.025]	[0.074]	[0.054]	[0.022]	[0.055]	[0.044]
Within 80mi. (Additive)	0.043	0.142	0.107	−0.004	−0.001	−0.003
	[0.025]	[0.043]	[0.034]	[0.023]	[0.023]	[0.023]
Decay	−0.151	0.079	0.000	−0.166	0.024	−0.024
	[0.046]	[0.030]	[0.024]	[0.051]	[0.024]	[0.024]
Decay (Additive)	−0.015	0.155	0.095	−0.077	0.026	−0.001
	[0.023]	[0.047]	[0.032]	[0.029]	[0.024]	[0.023]
Rings (0-20, 20-30, 30-40)	0.000	0.219	0.164	0.000	0.182	0.149
	[0.022]	[0.070]	[0.049]	[0.022]	[0.055]	[0.044]
Rings (0-20, 20-30, 30-40, 40-60, 60-80)	0.000	0.000	0.000	0.000	0.000	0.000
	[0.024]	[0.024]	[0.024]	[0.024]	[0.024]	[0.024]
Rings (0-20, 20-30, 30-40, 40-60, 60-80) (Additive)	0.044	0.142	0.108	−0.001	−0.001	−0.001
	[0.025]	[0.043]	[0.035]	[0.023]	[0.023]	[0.023]

Notes. Each cell corresponds to the results from 1000 simulations of the data generating process specified by the column header and estimated by equation (6) with $\tilde{h}(\vec{D}, i)$ given by the row label. The non-bracketed number represents the mean bias from the simulations and the bracketed number represents the mean square error given by $\sum_{i=1}^{n} \frac{(\hat{\tau}_i - \tau_0)^2}{n}$

of estimating spillover effects that are additive in nature. However, if the true specification is non-additive, then the additive specification no longer will remove all the bias from the direct effect estimate.

The results of the simulations are produced in Table 1. For each entry in the table, the column label corresponds to the true exposure mapping used in the data-generating process and the row label corresponds to the exposure mapping, $\tilde{h}(\vec{D}, i)$, used in estimation. The corresponding

cell gives the mean bias from estimating equation (6) with the $\tilde{h}(\vec{D}, i)$ given by the row. The bracketed number below the mean bias is the mean-squared error of the estimate which is given by $\sum_{i=1}^{n} \frac{(\hat{\tau}_{i,\text{direct}} - \tau_0)^2}{n}$. The mean squared error combines the square of the bias term, the variance of the estimator, and the variance of the error term which does not depend on the equation specification. This measure gives a good sense of the trade-off with whether decreasing bias comes at the cost of increasing the variance of the estimator.

The first thing to note is that all specifications remove a large portion of the bias relative to traditional two-way fixed effect estimate and lower the mean-squared error. In particular, correctly specifying the data-generating process removes almost all the bias in the direct effect estimate and greatly lowers the mean-squared error.

The second result is that both the non-additive 'Rings' and the 'Within' specification perform best among all the misspecified spillovers in terms of removing bias, so long as they are wide enough to capture all the spillovers. This is due to the fact that indicator variables will correctly identify the average spillovers onto control units if they cover all the effected control units (see Proposition 2). They also do so without increasing the variance of the estimator relative to the true specification, as shown by comparing the mean-squared errors. However, if the indicator is not wide enough as is the case with Within 40mi. for the 80mi. data-generating process, bias remains in the estimates. Therefore, if researchers are primarily interested in identifying the direct effect, then either 'Rings' or 'Within' specifications with a large cutoff distance are preferred as they remove almost all bias in all settings.

3.2.2. Spillover Effect Estimation

Now, I turn to analyzing how the specifications perform at estimating the spillover effects themselves. For the next table, I predict the spillover effect for each control unit from the point estimates on $\tilde{h}(\vec{D}, i)$. Then as a measure for how well spillovers are estimated, I will calculate one minus the mean squared prediction error normalized by the sum of squared spillovers (for comparison across data-generating processes),

$$1 - \frac{\sum_{i:D_i=0}(\beta_{\text{spill, control}}h(\vec{D}, i) - \hat{\beta}_{\text{spill, control}}\tilde{h}(\vec{D}, i))^2}{\sum_{i:D_i=0}(\beta_{\text{spill, control}}h(\vec{D}, i))^2}.$$

This will produce a percentage of spillovers that are explained by the specification, so numbers closer to 100% represent better modelling of the spillovers.

Table 2 — Percent of Spillovers Predicted by Specification

Specification	Within 40mi	Within 80mi	Within 40mi. (Additive)	Within 80mi. (Additive)	Decay	Decay (Additive)
TWFE (No Spillovers)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Within 40mi.	99.3%	25.0%	58.8%	85.4%	38.4%	55.6%
Within 80mi.	39.5%	96.4%	85.7%	33.9%	71.5%	67.6%
Within 40mi. (Additive)	85.1%	20.5%	51.4%	99.5%	40.5%	60.6%
Within 80mi. (Additive)	45.5%	61.0%	70.5%	47.2%	98.3%	93.5%
Decay	60.1%	81.8%	97.4%	52.7%	75.5%	81.9%
Decay (Additive)	60.4%	55.6%	78.0%	63.8%	93.2%	98.4%
Rings (0-20, 20-30, 30-40)	98.4%	22.8%	58.2%	85.7%	37.1%	55.8%
Rings (0-20, 20-30, 30-40, 40-60, 60-80)	96.7%	91.9%	92.1%	84.2%	72.7%	78.4%
Rings (0-20, 20-30, 30-40, 40-60, 60-80) (Additive)	83.3%	56.6%	72.7%	97.6%	94.9%	94.8%

The header group "Data-Generating Process" spans all six data columns.

Notes. Each cell corresponds to the results from 1000 simulations of the data generating process specified by the column header and estimated by equation (6) with $\tilde{h}(\vec{D}, i)$ given by the row label. The number corresponds to the mean square prediction error of the spillover effect for each control unit normalized by the total variance of spillover effects, given by

$$1 - \frac{\sum_{i:D_i=0}(\beta_{\text{spill, control}}h(\vec{D}, i) - \hat{\beta}_{\text{spill, control}}\tilde{h}(\vec{D}, i))^2}{\sum_{i:D_i=0}(\beta_{\text{spill, control}}h(\vec{D}, i))^2}.$$

Results of this exercise are presented in Table 2. For each specification, correctly specifying the spillovers produces the best estimates for spillovers as expected. Among misspecified exposure mapping, 'Rings' that are large enough to capuhre all the spillovers perform better than all other specifications in all non-additive data-generating processes. On the other hand, when the spillovers are additive, the predicted spillovers from the non-additive ring specifications do not as accurately measure these spillovers as they did before.[15] Non-additvie specifciations perform somewhat poorly at estimating spillovers in the additive data-generating processes, 'Within (Additive)' and 'Decay (Additive)'. The 'Rings (Additive)' specification sucessfully predicts spillovers in both additive data-generating processes.

Therefore there are two practical pieces of guidance from these simulations. First, for their

---

[15] Although the non-additive Rings do estimate between 70% and 80% of the spillover effects while maintaining complete removal of bias.

ability to remove all bias, 'Rings' should be the default used by researchers for estimating the direct effect of treatment. Second, the most important aspect of correctly estimating spillovers is to consider whether they are additive in the number of treated units or not. Researchers should use a version of Rings either way.

3.3. Guide for Applied Researchers

The above discussion yields to practical recommendations based on three questions, the answers to which will ultimately depend on the economic context being studied.

1. Researchers should consider whether control units and/or treated units experience effects from the treatment status of other units. For example, in the context of library construction, treated units likely do not experience spillovers while close control units do. In the context of factory openings which affects nearby labor markets regardless of treatment status, both control and treated units likely experience spillovers.

2. Researchers must decide how far the spillovers extend and whether the effect decays over distance. This should again be guided by the economic context of the research question. For example, things such as driving distance or access via public transportation can be helpful in deciding which units are 'close enough' to experience spillovers.

   Importantly, Proposition 2 shows that the researcher does not need to specify the exact functional form of spillovers in order to have unbiased estimates of the treatment effect provided that spillovers are 'local' as in Assumption 2. If only the direct effect of treatment is of interest, than a single ring indicator that contains units within the maximum distance will suffice. On the other hand, if spillover effects are of interest themselves, then a set of concentric rings is preferred for the reasons given above.

3. Last, it is important for estimation to determine if the spillovers are additive or non-additive in nature. If the number of nearby treated units matter, then the exposure mapping should sum over nearby treated units in some way. For example, in the context of library construction, the number of neighbors with libraries does not change the size of the spillover for it is access to *any* library that matters. In the context of factory openings, though, agglomeration economies imply that the spillovers are additive in nature. As illustrated in the Monte Carlo

simulations, in order to accurately estimate spillovers requires correctly using a additive or non-additive version of the concentric rings.

3.4. Application in Place-Based Policy Analysis

To illustrate the importance of accounting for spatial spillovers in the estimation of treatment effects, I revisit the analysis of the Tennessee Valley Authority (TVA) in Kline and Moretti (2014a). The TVA program was a large-scale federal investment started in 1934 that focused on construction of dams and transportation canals in an attempt to modernize the Tennessee Valley's economy. By the end of WWII, the TVA became the largest single power supplier in the country and significantly lowered the cost of wholesale energy for factories.[16] With over \$20 Billion (in 2000 dollars) spent which is hundreds of dollars transferred per person in the Authority, the impacts are very likely to extend past the authority's borders. Kline and Moretti (2014a) analysze a range of outcome variables, but for succinctness I use only (the log of) agricultural and manufacturing employment. Since the TVA primarily improved manufacturing industries through large-scale electrification, the authors predict that employment will grow in manufacturing and shrink in agricultural as workers switch to higher-paying manufacturing jobs.

The analysis in Kline and Moretti (2014a) begins by comparing changes in county-level outcomes from 1940 to either 1960 (short-run effects) or 2000 (long-run effects) between treated counties in the Authority and control counties outside. The primary specification is

$$y_{c,t} - y_{c,1940} = \alpha + \text{TVA}_c \tau + X_{c,1940} \gamma + (\varepsilon_{c,t} - \varepsilon_{c,1940}), \tag{7}$$

where c denotes county, TVA_c is an indicator variable for being in the Authority, Post_t is an indicator for being in the post-period, and y is a set of outcome variables (in logs).[17] Pre-treatment

---

[16] More details on the program are found in Kline and Moretti (2014a). The effects on wholseale electricity are discussed in Kitchens (2014).

[17] The two-periods difference-in-differences regression is equivalent to a first-difference regression. The authors use an Oaxaca-Blinder estimator on the first differences and the results of Kline (2011) show that this estimator is equivalent to a weighted difference-in-differences estimate. The weights are the propensity scores from a linear probability model where the outcome variable is being in the TVA. The weights can be seen in Figure III of their text, but generally counties close to the TVA are slightly upweighted relative to uniform weights. Their estimator does not differ much from the standard difference-in-differences results since the weights are not that different from uniform weights.

control variables, $X_{c,1940}$, are interacted with Post$_t$ to allow for places to be on different long-term trends.[18] To improve the likelihood of the paralell trends assumption, they run a logistic regression to predict being in the TVA based on their set of control variables $X_{c,1940}$ and keep only observations in the top 75% of predicted probability. The counties used in the sample are presented in Figure 2.

Kline and Moretti (2014a) estimate (7) to identify the 'local effect' of the TVA – what I am calling the 'direct effect'. However, their point estimates compare, in part, changes in outcomes in TVA counties with changes in outcomes for neighboring counties that likely were impacted by the large-scale program. Their identification strategy depends *explicitly* on a parallel counterfactual trends assumption after controlling for covariates and *implicitly* on the SUTVA assumption.

In the paper, the authors discuss the nature of spillovers that can occur. For agriculture employment, the authors claim that improved wages in the Authority will draw agriculture workers out of nearby counties. Hence they predict a negative spillover. For manufacturing, the sign is ambiguous. There could be positive spillovers if electrification brought cheap power and agglomeration economies to the neighboring areas.[19] However, manufacturing could decline if firms chose to locate in the Authority that would have, in the absence of the program, decided to locate in nearby counties.[20] My methodology will allow me to empirically test these predictions in the data and remove their bias from the treatment effect estimates.

By comparing counties inside the Authroity to counties on the other side of the border, the authors likely underestimate the negative effect on agricultural employment while the bias in the manufacturing effect is theoretically ambiguous. The authors do recognize the problem of these comparisons and remove counties that share borders with the authority's, but due to the scale of the program, the spillovers are likely to extend further than this. As discussed in Section 3.1, bias in their estimate will likely remain even after dropping contiguous counties. The estimation strategy I present keeps the observations near the TVA while controlling for spillover effects in a more rigorous manner.

I extend their analysis to control for spatial spillovers in the difference-in-differences specifica-

---

[18] See footnote 8 in Kline and Moretti (2014a) for a full listing of control variables.

[19] Duranton and Puga (2003) give theoretical motivations for different sources of agglomeration.

[20] Cuberes et al. (2021) refer to the positive spillover effects as 'urban access' and the negative spillover effecte as 'urban shadows' and find descriptive evidence of these effects changing importance over time.

22

Figure 2 — TVA Effective Sample and Spillover Variables
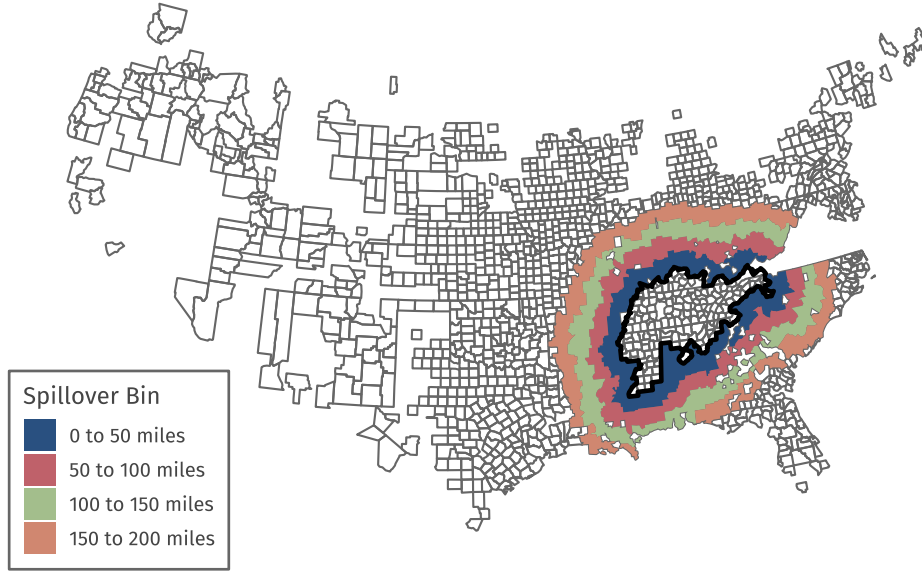
Notes: The above figure plots all the counties used in the estimation. Counties that fall within the distance intervals $\{(0, 50], (50, 100], (100, 150], (150, 200]\}$ measured in miles are colored by their respective bin.

tion. To parametrize the exposure mapping, I include a set of indicator variables for when a county is within a certain distance interval from the Authority. Specifically, I use the following intervals Dist = $\{(0, 50], (50, 100], (100, 150], (150, 200]\}$ measured in miles and define Between(d) as an indicator for being within the interval $d \in$ Dist away from the Authority. Figure 2 displays the four spillover variables by filling in each distance bin in a different color. The magnitude and potentially the sign of spillovers can change with distance, hence I allow the effect to be seperately identified in bins. I keep the number of bins small to improve precision of spillover estimates.

The specification with spillovers is given as follows:

$$y_{i,t} - y_{i,1940} = \alpha + \text{TVA}_i \tau + \sum_{d \in \text{Dist}} \text{Between}(d)\delta_d + X_{i,1940}\beta + (\varepsilon_{i,t} - \varepsilon_{i,1940}), \tag{8}$$

where $t \in \{1960, 2000\}$. The coefficients δ_d estimate the average spillover effect onto control units at different distances from the Authority. From proposition 2, $\hat{\tau}$ will be an unbiased estimate for the 'local' effect so long as spillovers do not occur past 200 miles from the TVA.

Table 3 — Effects of Tennessee Valley Authority on Decadel Growth

	Diff-in-Diff	Diff-in-Diff with Spillovers				
			TVA between 0-50 mi.	TVA between 50-100 mi.	TVA between 100-150 mi.	TVA between 150-200 mi.
	TVA	TVA				
Dependent Var.	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: 1940-2000						
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)
Panel B: 1940-1960						
Agricultural employment	0.0940***	0.0856*	−0.0062	−0.0042	−0.0303	−0.0039
	(0.0275)	(0.0444)	(0.0474)	(0.0448)	(0.0404)	(0.0298)
Manufacturing employment	0.0894***	0.0993**	0.0228	0.0225	−0.0055	−0.0066
	(0.0324)	(0.0432)	(0.0510)	(0.0561)	(0.0348)	(0.0253)

Notes. Each row corresponds to an outcome variable. Each cell is the point estimate and the standard error for the variable described in the column title. All standard errors are Conley standard errors with a correlation cutoff of 200 miles following Conley (1999). The column labeled 'Diff-in-Diff' estimates (7) by OLS and is similar to the estimate reported in Kline and Moretti (2014a). The final four columns labeled 'Diff-in-Diff with Spillovers' are estimates from (8).

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

The results of the long-run analysis from 1940 to 2000 are presented in Panel A Table 3. The first column lists which dependent variable (measured in logs) was used in the row. The following columns contain point estimates for τ and δ_d's in different specifications. The point estimates can be interpreted as decadel growth rates in outcomes. The column labeled difference-in-differences uses an ordinary least squares estimator for the specification without spillovers, equation (7). This estimate finds a decline in agricultural employment of about 5.1% per decade and an increase in manufacturing employment of about 5.6% per decade.

Turning to the specification that includes spillovers, equation (8), column (2) contains a point estimate for τ and and columns (3)-(6) contain point estimates of the spillover effects δ_d. For agricultural employment, the point estimates show there was a decline in agriculture employment in control units near the Authority. For control units between 0 and 50 miles, column (4) indicates

a decline in agricultural employment of 3.7% per decade. Between 50 and 100 miles the point estimate is -1.6% per decade, between 100 and 150 miles the point estimate is -3% per decade, and between 150 and 2000 miles the point estimate is -1.6% per decade. This is likely due to the fact that higher paying manufacturing jobs within the Authority drew farm-worker migrants from nearby counties. Because the spillovers onto the control counties is negative, the original difference-in-differences estimator was positively biased. The new point estimate indicates a decline of agricultural employment of about 7.4% per decade compared to 5.1% in the standard difference-in-difference specification.

For manufacturing, our point estimates for spillovers are consistently negative, though imprecisely estimated in some columns. The spillover estimates suggest that neighboring counties experienced potentially negative spillover effects in the long-run. Since there are negative spillover effects present, the new point estimate in column (2) of 3.5% is significantly smaller than the original estimate of 5.6%. The spillover estimates are evidence that in the long-run, urban shadow forces dominate the benefits of urban access. As you move away from the Tennessee Valley, the point estimates become more negative which suggest that the benefits of urban access vanish over distance faster than the costs of urban shadow.

To see how spillover effects from a large-scale place-based policy develop over time, Panel B in Table 3 presents results for the effects of the Tennessee Valley in the short-run using outcome data in 1960. Unlike in the long-run, areas near the Tennessee Valley did not experience significant declines in agricultural employment in the short-run. Since our long-run analysis finds significant increases in high paying manufacturing employment in the Tennessee Valley, this result is consistent with long-run migration costs being lower than short-run costs.

For manufacturing, there are potentially positive increases in manufacturing employment within 100 miles of the Tennessee Valley authority and near zero effects between 100 and 200 miles. In the short-run, it appears that the effects of urban access and the cheap wholesale electricity dominated the effects of urban shadow. The effect of urban shadows can potentially be smaller in the short-run if operating firms are unlikely to relocate. Long-run effects can be larger as entrant firms change their location decision and operating firms are slowly replaced which is what the long-run spillover effect estimates suggest.

These results show that including spillovers in the estimation of direct treatment effects

is potentially important and can lead to *significant* differences in treatment effect estimates. Analysis of place-based policies that do not account for the fact that treatment effects can spillover beyond the borders of treated areas can potentially be biased. More, the spillover effects caused by place-based policies change over time as frictions can create delays in reoptimizing behavior.

3.5. Identification Strategies in Place-Based Analysis

More generally, my framework provides important insights into identification strategies when analysizing the effects of place-based policies. There are two ways I contribute. First, Baum-Snow and Ferreira (2015) recognize the problem of spatial spillovers causing problems in identification and point to aggregation of units as a way to alleviate to the problem (e.g. aggregating census tracts to metropolitan areas). However, this approach collapses the three seperate effects which might each be of interest into a singular aggregate effect. I have shown that all three effects of place-based policies can be estimated under very general assumptions about the spillovers. Therefore, the methods propose simple ways to estimate treatment effects with non-aggregated data.

Second, my framework provides insight for different identification strategies often used in the literature. Since place-based policies are targeted to specific distressed areas, comparison units are often hard to find. Researchers have developed many identification strategies to find comparison units that are similar in terms of unobservables. First, researchers use border discontinuities to compare treated units to units just on the otehr side of the border. The rationale for this is that unobservables that could be correlated with treatment would evolve continuously across the border and so neighobring units would serve as valid comparisons. A different identification strategy compares approved and close but ultimately rejected applicants (e.g.). Since both the treated and comparison groups likely are simliar due to the comparison group was close to being treated, unobservables should be balanced. A key difference in these identification strategies is the distance comparison units are from the treated area and therefore the former is more prone to bias due to spillovers than the latter.

For example, consider the analysis of the effects of federal Empowerment Zones. The federal Empowerment Zones are specially designated areas in high-unemployment areas. The program gives businesses located in the zone tax incentives and the goal of the program is to reduce

unemployment and poverty. In the literature, there are a set of conflicting results with some papers suggesting that the Empowerment Zones do indeed reduce poverty rates and others finding near-zero effects.[21]

Busso et al. (2013) compare census tracts in Empowerment Zones to census tracts that qualified and were rejected from the program. The rejected tracts are not typically geographically near accepted Empowerment Zones which removes the concern of spatial spillovers affecting control units' change in outcomes. They find large significant reductions in poverty rates. Meanwhile, Neumark and Kolko (2010) compare census tracts in Empowerment Zones to census tracts within 1,000 feet of the Zone. These control counties are likely the ones that experience the largest spillover effects. They find near-zero effects on poverty.

My framework can reconcile 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. These two identification strategies of using rejected applicants or using bordering units as the control group are very common in the urban literature more generally (Baum-Snow and Ferreira, 2015). My paper suggests that the former is the preferred strategy if spillovers occur onto nearby control units. Identification strategies that rely on geographically close control units as having similar unobservables should be used cautiously. If treatment effects spillover treatment border, then estimated treatment effects can be biased.

4 — Estimating Event Study with Spillovers

The intuition from how spillovers cause biases in estimates of the direct effect of treatment extend into the setting where there is staggered adoption in treatment. When treatment turns on for different units at different times, two-way fixed effect estimates can be viewed as a weighted sum of 2×2 difference-in-differences estimates.[22] Therefore the bias terms will be an identically weighted sum of the bias terms from the 2×2 estimates, assuming that the parallel counterfactual trends assumption (1) holds. However, since weights on some of the 2×2 estimates can be

---

[21] See Table 1 of Neumark and Young (2019) for a summary of the various treatment effect estimates in the literature.

[22] Various forms for these weights are described in Goodman-Bacon (2018), Sun and Abraham (2020), and de Chaisemartin and D'Haultfoeuille (2019). I do not recharacterize the weights in this article and guide interested readers to the source articles themselves.

negative, the sign of the spillover effects do not determine the sign of the weighted average of the spillover effects. This makes the bias from spillovers much more difficult to sign.

In order to estimate treatment effects in the presence of spatial spillovers and staggered treatment timing, I will propose an estimation strategy that follows Callaway and Sant'Anna (2020) while incorporating spillovers directly into estimation. However, I do not cover all the technical assumptions and point the reader to the main text for these details. To begin, Callaway and Sant'Anna define a 'cohort' g as the set of units that recevie treatment starting in period g. Let G_g be an idicator equal to one for all units where treatment starts at period g and C an indicator equal to one if the unit never receives treatment. Each cohort, g, will serve as the treated group for a set of 2×2 difference-in-differences. For each period t they define the 'group-time average treatment effect', $ATT(g,t) = \mathbb{E}\left[Y_t(g) - Y_t(0) \mid G_g = 1\right]$, where G_g is an indicator for starting treatment in year g. The potential outcome is indexed by g to allow for treatment effects to differ depending on the initial treatment year.

Therefore in the context of spillovers, our equivalent term is the 'group-time average direct effect' which will be defined as

$$ATT_{\text{direct}}(g,t) = \mathbb{E}\left[Y_t(g,0) - Y_t(0,0) \mid G_g = 1\right],$$

where the second term in potential outcomes represents an exposure mapping of $h(\vec{D}, i) = 0$. Similar to above, we modify the parallel trends assumptions given by Callaway and Sant'Anna to hold in the absence of exposure.

Assumption 3 (Parallel Counterfactual Trends on a "Never-Treated" Group)**.**

$$\mathbb{E}\left[Y_t(0,0) - Y_{t-1}(0,0) \mid G_g = 1\right] = \mathbb{E}\left[Y_t(0,0) - Y_{t-1}(0,0) \mid C = 1\right]$$

Assumption 4 (Parallel Counterfactual Trends on "Not-Yet-Treated" Groups)**.**

$$\mathbb{E}\left[Y_t(0,0) - Y_{t-1}(0,0) \mid G_g = 1\right] = \mathbb{E}\left[Y_t(0,0) - Y_{t-1}(0,0) \mid D_s = 0, G_g = 0\right]$$

The group-time average direct effect is the building block of analysis in the framework proposed by Callaway and Sant'Anna. Estimation of the group-time average direct treatment effect, $ATT_{\text{direct}}(g,t)$ depends on which parallel trends assumption a researcher use. The difference-in-difference estimate uses period t outcomes as the post-period and some period $g - \delta$ as

28

the pre-period where δ is the number of periods before initial treatment.[23] In the context of spillovers, the canoncial difference-in-differences estimation will have to be adjusted using a parameterized exposure mapping to prevent the bias detailed in the 2×2 setting. As shown in Section 3, this can be done using a large 'Within' indicator that is wide enough to capture all of the spillovers. If a researcher is using Assumption 4, then estimation will be done with a subsample containing treated units from group g and all units that are in groups $g' > t$. If a researcher is using Assumption 3, then estimation will be done using a supsample containing only control units and units in group g.

Then, as detailed in C&S, the $ATT(g, t)$ can be aggregated using different weights depending on the pararmeter of interest (see Table 1 of Callaway and Sant'Anna (2020)). These aggregations take the form of

$$\theta = \sum_g \sum_t w(g, t) ATT_{\text{direct}}(g, t)$$

Different weights produce different estimates for different average effects of interest, e.g. event study estimates, group-level average treatment effects (across time), or an aggergate treatment effect. Estimation and different aggregation of the direct effects can be done using the 'did' package produced by C&S where the the 'Within' indicator is included as a covariate. It is important to note that exposure can change over time as new units are treated, so researchers should make sure to use the correct values for exposure mapping for each year t.

4.1. Estimation of Spillover Effects

Estimation of the spillover effects is potentially much more difficult as estimation of continuous variables or multiple treatment variables in the presence of staggered treatment timing is still a work in progress and past the scope of this article. However, for a 'Within' indicator, estimation can be done in much the same way as the direct effect but with more caution when constructing the samples for each 2×2 estimate. The 'Within' indicator can be thought of as the treatment indicator and the sample would consist of only the never-treated units. The advantage of the 'Within' indicator is that it will remove all bias from the direct effects estimates if the indicator extends far enough and will give an estimate of the average spillover effect. However, as mentioned above, a

---

[23] It is recommended that $g - \delta$ be the closest period before anticipation can occur. For example, if no anticipation is present in the setting, then $\delta = 1$ would be the recommended choice.

'Within' indicator does a poor job of estimating (a) the shape of spillover effects over distance and (b) spillover effects that are additive in the number of nearby treated units. Estimation of different exposure mappings is a task left for future research.

Groups in this case will be defined by the period which 'Within' first equals 1 for a given observation and estimation of $ATT(g,t)$ should use all observations where 'Within' is equal to zero for period t and before. Then, estimation can be done assuming that all control units are on the same counterfactual trends in the absence of spillovers which is satisfied by Assumption 3. Therefore, estimation and aggregation can be done in much the same way as for the direct effect but on just the subsample of control units.

4.2. Application on Community Health Centers

As an application of the above methods, I extend the analysis of Bailey and Goodman-Bacon (2015). The authors study the creation of federal community health centers between 1965 and 1974 that provided *primary* care to low-income communities. They test the hypothesis if access to low/no- cost health care services decreased mortality rates on the treated population. To answer this question, the authors use a common event-study framework to compare outcomes in treated counties to all other US counties by estimating the following specification

$$Y_{jt} = \theta_j + X_{jt}\beta + \sum_{y=-7}^{-2} \pi_y D_j 1(t - T_j^* = y) + \sum_{y=0}^{15} \tau_y D_j 1(t - T_j^* = y) + \varepsilon_{jt}, \tag{9}$$

where T_j^* is the year the county establishes a community health center, D_j is an indicator for being treated, θ_j is county fixed effects, and X_{jt} contains a set of controls.[24] The coefficents π_y can be interpreted as tests of parallel pre-trends and τ_y can be interepreted as the treatment effect of a community health center y years after establishment.

The result are in Figure 3. Estimates of $\hat{\tau}_y$ show that counties that received health centers had significantly lower mortality rates than the other counties in the US. In years following the establishment of the community health centers, the authors find a reduction of between 15-30 deaths per 100,000 residents compared to a baseline adjusted mortality rate of 929 deaths per 100,000 residents.

---

[24] Controls include 1960 county characteristic trends, state-year fixed effects, and urban-group fixed effects. A full list can be found on page 1080 of Bailey and Goodman-Bacon (2015).

Figure 3 — Effects of Establishment of Community Health Centers
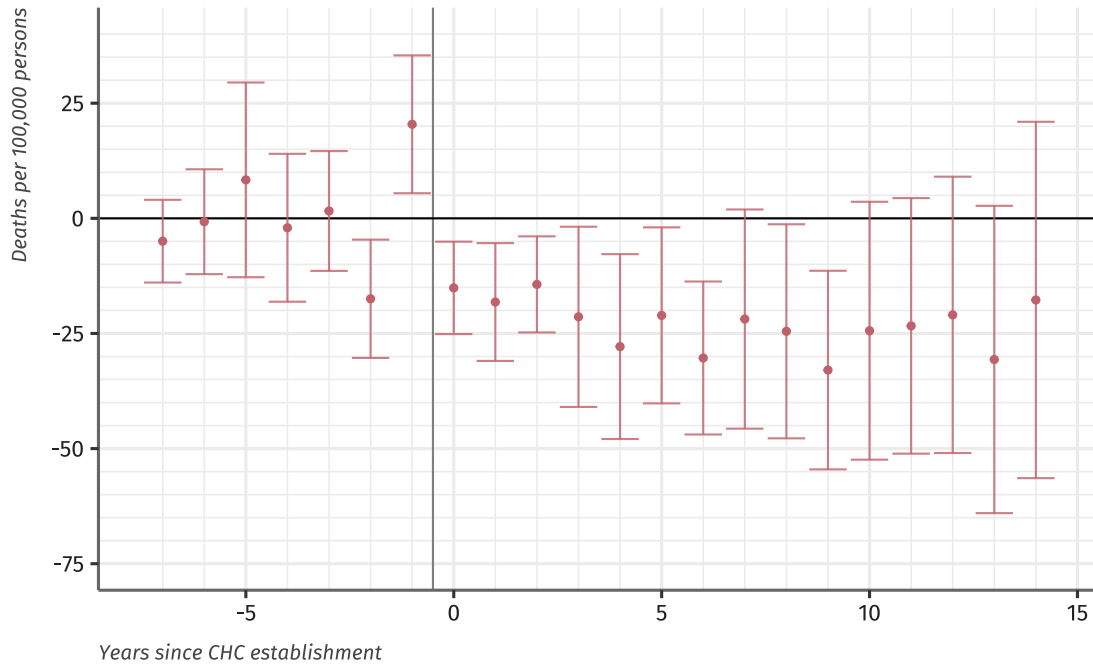

Notes: This figure recreates figure 5 from Bailey and Goodman-Bacon (2015). Details on the event study specificaiton and controls can be found in the text.

There are theoretical reasons to think spillovers may or may not exist in this context. On the one hand, individuals outside the county can potentially travel to the community health centers to receive care. This would create a negative spillover effect on mortality rates in nearby counties which would bias their estimates towards zero. On the other hand, Bailey and Goodman-Bacon (2015) document evidence that their estimated effects are not due to emergency but rather primary care services. In this case, it is less likely low-income individuals would travel very far to receive primary care and hence the spillover effects is potentially near zero. My methodology can provide an answer to the question of how far do individuals travel for low/no -cost primary care.

As in the method detailed above, I use the subsample of counties that do not receive a community health center and use an indicator for being within 25 miles of a treated county as the treatment variable. Then, I follow Callaway and Sant'Anna (2020) in estimating event-study

Figure 4 — Direct and Spillover Effects of Community Health Centers
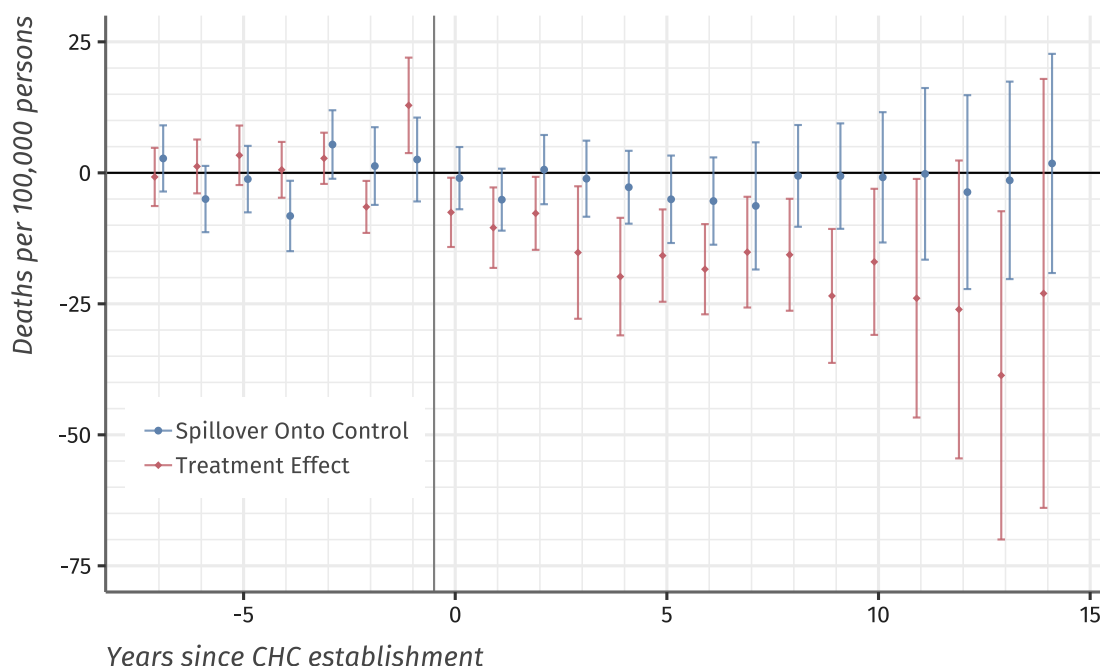
Notes: This figure plots event study estimates for the treatment effect and the spillover effect on control units within 25 miles of treatment at different periods relative to establishment year. The estimates are generated using the 'did' package produced from the paper Callaway and Sant'Anna (2020).

coefficients. The results are presented in Figure 4. The confidence intervals labeled with circles represent point estimates for the average spillover effect on control units within 25 miles. No spillover effect is estimated to be significantly different from zero which suggests that the effects of community health centers are very local. Since there are near zero spillover effects, the direct effect estimates marked in Figure 4 as diamonds maintain the same shape as the author's original estimates with estimates between 15-30 fewer deaths per 100,000 persons.

The spillover effects results provide evidence that low-income individuals will not travel far to receive primary care. Practically, this suggests that community health centers should be targeted to be as accessible as possible for poor individuals as they are unable to travel far to access the services.

5 — Conclusion

This paper has considered the common environment where treatment is assigned via administrative boundary while the effects of treatment spread across these borders. In this context, difference-in-differences estimation will identify a combination of the direct effect of treatment and two additional terms resulting from spillovers. I have proposed a potential outcomes framework that formalizes spillovers.

I use this framework to point to limitations in commonly used ad-hoc approaches to controlling for spillovers. Then, I propose a new estimation strategy that is robust to more forms of spillovers. In particular, I find that specifications with an indicator for being "close to" treated units interacted with treatment status will remove all bias in the direct effect estimate so long as all units affected by spillovers are contained in the indicator. More, I show that a set of concentric 'rings' are the best at capturing the spatial structure of spillovers. When estimating spillover effects themselves, I find that it is important to correctly identify is spillovers are additive or non-additive in the number of nearby treated units.

Then, I show that these approaches can change estimates significantly in the context of estimating the effects of place-based policies. Since place-based policies change the nature of agglomeration in the local and surrounding area –that is cause spillovers–local effects of these policies can be misestimated without controlling for spillovers. I also show the importance of considering spillovers in weighing the pros and cons of various identificatin strategies. Identification strategies based on geographic continuity of unobservables can magnify the bias from spillovers as they restrict the compairson group to observations experiencing the largest spillover effects.

Finally, I show how researchers can include spillovers in settings with variation in treatment timing following Callaway and Sant'Anna (2020). I then use this method to show how considering spillovers in event-study framework provide additional insights for a researcher.

References

Angelucci, M., and V. Di Maro. 2016. "Programme evaluation and spillover effects." *Journal of Development Effectiveness* 8 (1): 22–43. 10.1080/19439342.2015.1033441.

Angrist, Joshua D. 2014. "The perils of peer effects." *Labour Economics* 30 98–108. 10.1016/j. labeco.2014.05.008.

Bailey, Martha J., and Andrew Goodman-Bacon. 2015. "The War on Poverty's Experiment in Public Medicine: Community Health Centers and the Mortality of Older Americans." *American Economic Review* 105 (3): 1067–1104. 10.1257/aer.20120070.

Basker, Emek. 2005. "Job Creation or Destruction? Labor Market Effects of Wal-Mart." *The Review of Economics and Statistics* 87 (1): 174–183.

Baum-Snow, Nathaniel, and Fernando Ferreira. 2015. *Causal Inference in Urban and Regional Economics*. Volume 5. 3–68, Elsevier, . 10.1016/B978-0-444-59517-1.00001-5.

Berg, Tobias, and Daniel Streitz. 2019. "Handling Spillover Effects in Empirical Research."

Berkes, Enrico, and Peter Nencka. 2020. "Knowledge Access: The Effects of Carnegie Libraries on Innovation."

Busso, Matias, Jesse Gregory, and Patrick Kline. 2013. "Assessing the Incidence and Efficiency of a Prominent Place Based Policy." *American Economic Review* 103 (2): 897–947. 10.1257/aer.103. 2.897.

Callaway, Brantly, and Pedro H.C. Sant'Anna. 2020. "Difference-in-Differences with multiple time periods." *Journal of Econometrics* S0304407620303948. 10.1016/j.jeconom.2020.12.001.

de Chaisemartin, Clément, and Xavier D'Haultfoeuille. 2019. *Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects*. (w25904): w25904. 10.3386/w25904.

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

Clarke, Damian. 2019. "A Convenient Omitted Variable Bias Formula for Treatment Effect Models." *Economics Letters* 174 84–88. 10.1016/j.econlet.2018.10.035.

Conley, T.G. 1999. "GMM estimation with cross sectional dependence." *Journal of Econometrics* 92 (1): 1–45. https://doi.org/10.1016/S0304-4076(98)00084-0.

Cuberes, David, Klaus Desmet, and Jordan Rappaport. 2021. "Urban Growth Shadows." *Journal of Urban Economics* 103334. 10.1016/j.jue.2021.103334.

Delgado, Michael S., and Raymond J.G.M. Florax. 2015. "Difference-in-differences techniques for spatial data: Local autocorrelation and spatial interaction." *Economics Letters* 137 123–126. 10.1016/j.econlet.2015.10.035.

Duranton, Giles, and Diego Puga. 2003. *Micro-Foundations of Urban Agglomeration Economies.* (w9931): w9931. 10.3386/w9931.

Goodman-Bacon, Andrew. 2018. *Difference-in-Differences with Variation in Treatment Timing.* (w25018): w25018. 10.3386/w25018.

Halloran, M. Elizabeth, and Claudio J. Struchiner. 1995. "Causal Inference in Infectious Diseases:." *Epidemiology* 6 (2): 142–151. 10.1097/00001648-199503000-00010.

Keller, Wolfgang. 2002. "Geographic Localization of International Technology Diffusion." *The American Economic Review* 92 (1): 52.

Kitchens, Carl. 2014. "The Role of Publicly Provided Electricity in Economic Development: The Experience of the Tennessee Valley Authority, 1929–1955." *The Journal of Economic History* 74 (2): 389–419. 10.1017/S0022050714000308.

Kline, Patrick. 2011. "Oaxaca-Blinder as a Reweighting Estimator." *American Economic Review* 101 (3): 532–537. 10.1257/aer.101.3.532.

Kline, Patrick, and Enrico Moretti. 2014a. "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): 275–331. 10.1093/qje/qjt034.

Kline, Patrick, and Enrico Moretti. 2014b. "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs." *Annual Review of Economics* 6 (1): 629–662. 10.1146/annurev-economics-080213-041024.

Miguel, Edward, and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72 (1): 159–217. 10.1111/j.1468-0262.2004.00481.x.

Neumark, David, and Jed Kolko. 2010. "Do enterprise zones create jobs? Evidence from California's enterprise zone program." *Journal of Urban Economics* 68 (1): 1–19. 10.1016/j.jue.2010.01.002.

Neumark, David, and Timothy Young. 2019. "Enterprise zones, poverty, and labor market outcomes: Resolving conflicting evidence." *Regional Science and Urban Economics* 78 103462. 10.1016/j.regsciurbeco.2019.103462.

Severnini, Edson R. 2014. "The Power of Hydroelectric Dams: Agglomeration Spillovers." 70.

Sobel, Michael E. 2006. "What Do Randomized Studies of Housing Mobility Demonstrate?: Causal Inference in the Face of Interference." *Journal of the American Statistical Association* 101 (476): 1398–1407. 10.1198/016214506000000636.

Sun, Liyang, and Sarah Abraham. 2020. "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects." 53.

Sävje, Fredrik, Peter M. Aronow, and Michael G. Hudgens. 2019. "Average treatment effects in the presence of unknown interference." *arXiv:1711.06399 [math, stat]*, http://arxiv.org/abs/1711.06399, arXiv: 1711.06399.

Vazquez-Bare, Gonzalo. 2019. "Identification and Estimation of Spillover Effects in Randomized Experiments." *arXiv:1711.02745 [econ]*, http://arxiv.org/abs/1711.02745, arXiv: 1711.02745.

Verbitsky-Savitz, Natalya, and Stephen W. Raudenbush. 2012. "Causal Inference Under Interference in Spatial Settings: A Case Study Evaluating Community Policing Program in Chicago." *Epidemiologic Methods* 1 (1): . 10.1515/2161-962X.1020.

A — Proofs

Proof of Theorem 1

$$\mathbb{E}\left[\hat{\tau}\right] = \underbrace{\mathbb{E}\left[\bar{Y}_{i,Post} - \bar{Y}_{i,Pre} \mid D_i = 1\right] - \mathbb{E}\left[\bar{Y}_{i,Post} - \bar{Y}_{i,Pre} \mid D_i = 0\right]}_{\text{Difference-in-Differences}}$$

$$= \mathbb{E}\left[\bar{Y}_{i,Post}(1, h(\vec{D}, i)) - \bar{Y}_{i,Pre}(0, \vec{0}) \mid D_i = 1\right] - \mathbb{E}\left[\bar{Y}_{i,Post}(0, h(\vec{D}, i)) - \bar{Y}_{i,Pre}(0, \vec{0}) \mid D_i = 0\right]$$

$$= \mathbb{E}\left[\bar{Y}_{i,Post}(1, h(\vec{D}, i)) - \bar{Y}_{i,Pre}(0, \vec{0}) \mid D_i = 1\right]$$
$$\quad - \mathbb{E}\left[\bar{Y}_{i,Post}(0, h(\vec{D}, i)) + \bar{Y}_{i,Post}(0, \vec{0}) - \bar{Y}_{i,Post}(0, \vec{0}) - \bar{Y}_{i,Pre}(0, \vec{0}) \mid D_i = 0\right]$$

$$= \mathbb{E}\left[\bar{Y}_{i,Post}(1, h(\vec{D}, i)) - \bar{Y}_{i,Pre}(0, \vec{0}) \mid D_i = 1\right] - \mathbb{E}\left[\bar{Y}_{i,Post}(0, \vec{0}) - \bar{Y}_{i,Pre}(0, \vec{0}) \mid D_i = 0\right]$$
$$\quad - \mathbb{E}\left[\bar{Y}_{i,Post}(0, h(\vec{D}, i)) - \bar{Y}_{i,Post}(0, \vec{0}) \mid D_i = 0\right]$$

$$= \mathbb{E}\left[\bar{Y}_{i,Post}(1, h(\vec{D}, i)) - \bar{Y}_{i,Pre}(0, \vec{0}) \mid D_i = 1\right]$$
$$\quad - \mathbb{E}\left[\bar{Y}_{i,Post}(0, \vec{0}) - \bar{Y}_{i,Pre}(0, \vec{0}) \mid D_i = 1\right]$$
$$\quad - \mathbb{E}\left[\bar{Y}_{i,Post}(0, h(\vec{D}, i)) - \bar{Y}_{i,Post}(0, \vec{0}) \mid D_i = 0\right]$$

$$= \mathbb{E}\left[\bar{Y}_{i,Post}(1, h(\vec{D}, i)) - \bar{Y}_{i,Pre}(0, \vec{0}) - \bar{Y}_{i,Post}(0, \vec{0}) + \bar{Y}_{i,Pre}(0, \vec{0}) \mid D_i = 1\right]$$
$$\quad - \mathbb{E}\left[\bar{Y}_{i,Post}(0, h(\vec{D}, i)) - \bar{Y}_{i,Post}(0, \vec{0}) \mid D_i = 0\right]$$

$$= \mathbb{E}\left[\bar{Y}_{i,Post}(1, h(\vec{D}, i)) - \bar{Y}_{i,Post}(0, \vec{0}) \mid D_i = 1\right] - \mathbb{E}\left[\bar{Y}_{i,Post}(0, h(\vec{D}, i)) - \bar{Y}_{i,Post}(0, \vec{0}) \mid D_i = 0\right]$$

$$= \mathbb{E}\left[\bar{Y}_{i,Post}(1, h(\vec{D}, i)) + \bar{Y}_{i,Post}(1, \vec{0}) - \bar{Y}_{i,Post}(1, \vec{0}) - \bar{Y}_{i,Post}(0, \vec{0}) \mid D_i = 1\right]$$
$$\quad - \mathbb{E}\left[\bar{Y}_{i,Post}(0, h(\vec{D}, i)) - \bar{Y}_{i,Post}(0, \vec{0}) \mid D_i = 0\right]$$

$$= \mathbb{E}\left[\bar{Y}_{i,Post}(1, \vec{0}) - \bar{Y}_{i,Post}(0, \vec{0}) \mid D_i = 1\right] + \mathbb{E}\left[\bar{Y}_{i,Post}(1, h(\vec{D}, i)) - \bar{Y}_{i,Post}(1, \vec{0}) \mid D_i = 1\right]$$
$$\quad - \mathbb{E}\left[\bar{Y}_{i,Post}(0, h(\vec{D}, i)) - \bar{Y}_{i,Post}(0, \vec{0}) \mid D_i = 0\right]$$

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

B — Details on Dropping Control Unit Approach

I use the set of counties in the contiguous United States and the data generating process used is defined by the exposure mapping given in (1) with cutoff distance $\bar{d} = 40$ miles and the data generating process given in (??). A unit of observation is a US county and the periods are $t \in \{1, \ldots, 20\}$ with treatment turning on after period $t = 10$. The unit and time fixed effects are generated by $\mu_t \sim N(0.2t, 0.1^2)$ and $\mu_i \sim N(6, 2^2)$ respectively, and the error term is $\varepsilon \sim N(0, 2^2)$. Last, the size of treatment and spillovers are as follows: $\beta_{\text{direct}} = 2$, $\beta_{\text{spill, control}} = 1$ and $\beta_{\text{spill, treat}} = 0$.

This data generating process where spillovers only occur onto control units matches what has been typically assumed in the literature. I assign treatment among counties randomly with various unconditional probabilities between 3 percent and 50 percent. The data-generating process is therefore

$$y_{i,t} = \mu_t + \mu_i + \beta_{\text{direct}} D_{i,t} + \beta_{\text{spill, control}}(1 - D_{i,t}) \text{Near}_{i,t} + \varepsilon_{i,t} \tag{10}$$

The size of the bias from estimating the two-fixed effects model, (4), at different treatment probabilities are presented in Figure B.1a as the line with diamond markers. Each point represents a set of 10,000 simulation and displays the mean bias as well as a 95 percent empirical confidence interval. As displayed in the figure, even for a low treatment probability of three percent, the bias is quite large with a 95 percent empirical confidence interval between -0.28 and -0.75. As treatment frequency increases, the bias increases as well but at a slower rate due to fewer additional control units receiving spillover units. The slow increase in bias is in part due to the assumption that spillovers are not additive in the number of nearby units treated.

A common solution in the literature is to remove control units from the estimated sample that are most likely to be affected by spillovers. To do this, I remove contiguous control counties which is a close, but misspecificied, measure of $h(\vec{D}, i)$. The results are shown in Figure B.1a by the line with circle markers. Even though the exposure mapping is misspecified, contiguous counties approximates the true exposure mapping well enough such that the bias stays centered constantly around zero as most control units experiencing spillovers are removed. However, if the distance cutoff $\bar{d}$ were larger, more control units would remain in the sample that experience non-zero exposures. In this case, the bias would fall between the two lines.

Figure B.1 — Effectiveness of Removing 'Contaminated' Control Units

(a) Spillovers Effects *only* on Control

(b) Spillovers Effects on Control and Treated
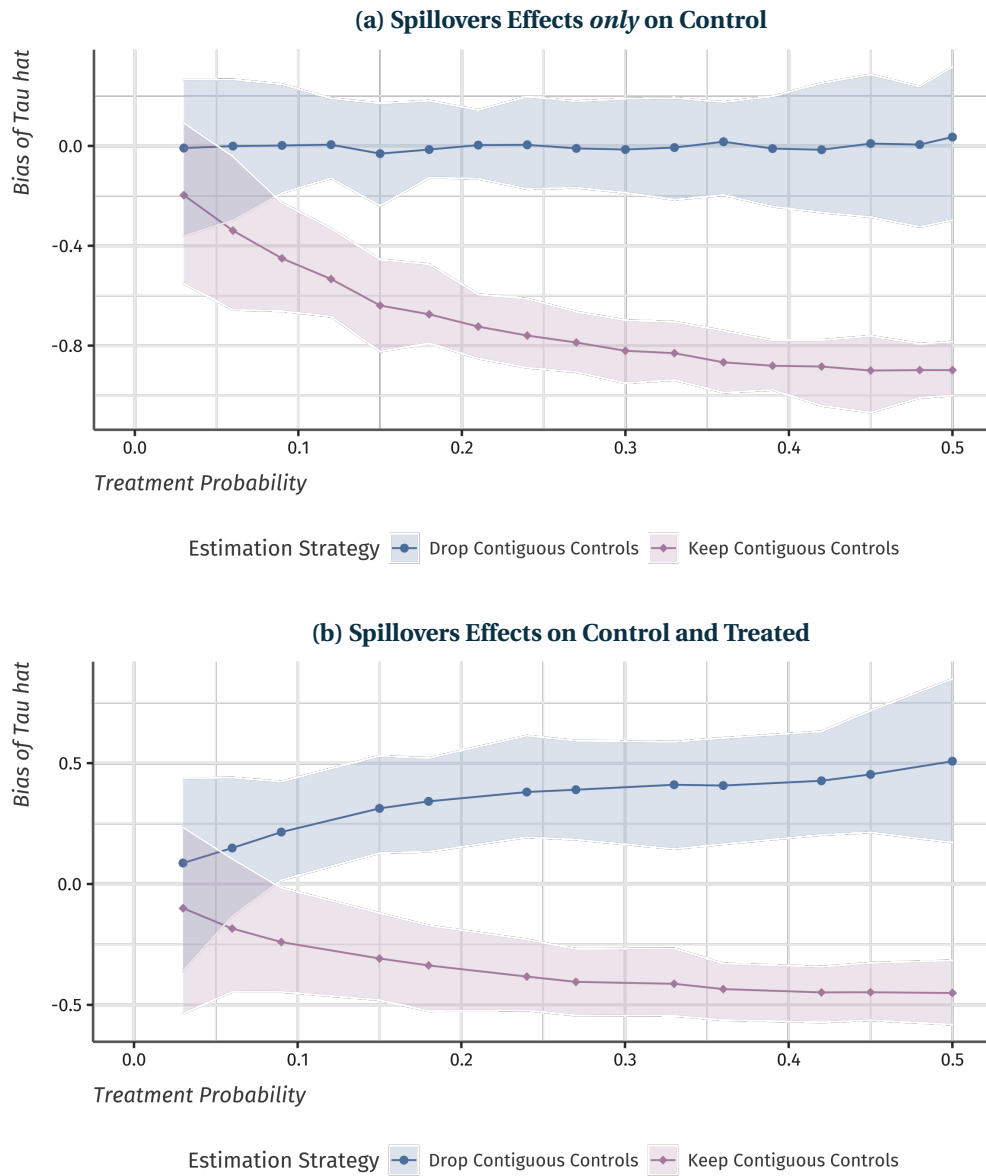
Notes: This figure plots the bias of $\hat{\tau}$ found from estimating (4) for data generated as described in the text by equations (10) for Panel A and (11) for Panel B. Each point corresponds to the average bias for the given treatment probability and the band is the 95 percent empirical confidence interval over 1000 simulations. The line with diamond markers estimates with all control units. The line with circle markers removes control units that share a border with a treated county.

There is a trade-off between the bias and the variance of the estimator when using this methodolgy. As the treatment probability increases, the number of control units removed increases as well. This naturally yields a more variable estimator as seen in the wider 95 percent empirical confidence intervals in Figure B.1a. This trade-off can be avoided altogether by parameterizing the spillovers and including them in estimation directly.

There is a second problem with the method of removing control units in the presence of spillover effects onto treated units. Since removing control units does not control for this second source of spillover effects, the estimates will be biased. In the second simulation, I add in spillover effects into the above data-generating process and set $\beta_{\text{spill, treat}} = 0.5$:

$$y_{i,t} = \mu_t + \mu_i + \beta_{\text{direct}}D_{i,t} + \beta_{\text{spill, control}}(1 - D_{i,t})\text{Near}_{i,t} + \beta_{\text{spill, treat}}D_{i,t}\text{Near}_{i,t} + \varepsilon_{i,t} \qquad (11)$$

The results are displayed in Figure B.1b. Again, two-way fixed effect estimation results in a biased estimate as shown by the line with diamond markers. The magnitude is smaller than in Panel (a) because the positive spillovers onto treated units cancel out with the positive spillovers onto control units. However, removing the control units near the treated units results in a biased estimate due to the average spillover onto also treated units, as seen by the line with circle markets in Panel (b). This problem is particularly concerning as a researcher may assume that there is positive spillovers on control units that is negatively biasing their estimate and when removing those units the estimate would increase in magnitude as expected. Researchers could therefore potentially assume they have removed all bias from their estimates even though the second form of bias remains.