Difference-in-Differences with Spatial Spillovers[*]

Kyle Butts[†]

November 4, 2021

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 including in settings with staggered treatment timing. To highlight the importance of spillover effects, I revisit analyses of three place-based interventions.

[†]University of Colorado Boulder, Economics Department (kyle.butts@colorado.edu)

1 — Introduction

Empirical work in economics often considers settings where a policy is targeted to groups of units by geographic boundaries, but the effect of treatment spillovers onto 'nearby' units.[1] For example, individuals in control areas can travel to the treated areas and receive treatment (e.g. a hospital opening serves nearby residents), or shocks to a labor market can affect nearby areas (e.g. a new factory opening increases service sector spending in the commuting zone). For other treated units, forces of agglomeration can increase effect size as treatment concentration increases (e.g. knowledge-sharing between firms) or forces of congestion/competition can decrease the effects of treatment (e.g. bidding up of wages causes the employment effects of factory openings to be mitigated).

In this paper, I introduce a potential outcomes framework that directly models spillover effects. While other papers have derived identification results using specific parameterizations, my paper is the first to model spillovers in a general form. Using this potential outcome framework, I define different possible treatment/spillover effects of interest and the relative advantage and interpretation of each. First, there is the 'switching effect' which is the relevant parameter for local policymakers who want to answer "What is the effect of switching my treatment status, holding fixed all other units' treatment?". It is difficult to estimate the switching effect without parametric assumptions. The reason for this is that the magnitude of the switching effect potentially depends on the treatment status of a unit's neighbors. For example, the construction of low-cost medical facilities will have larger treatment effects when neighboring areas do not have facilities. Switching effects therefore are defined separately for each level of exposure. Estimation of a switching effect requires knowledge of which units experience the same level of exposure. While the switching effect is not identified nonparametrically in general, there is one effect that is identified under mild assumptions. This effect is the 'direct effect' of treatment which represents the switching effect for units with zero exposure to other treatments. That is, what are the effects of turning on a unit's treatment if this were the *only* unit to be treated in their area.

---

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.

Second, there is the 'total effect' which is the relevant parameter for national policymakers who want to answer "What is the average effect of implementing the entire treatment regime?". This treatment effect summarizes the effects of going from the world with no treatment to the entire treatment regime. This effect comes from two sources: the direct effect and the average spillover effect from other treated units. Unlike the switching effect, the total effect is nonparametrically identified under a mild assumption that spillovers are 'local'. The 'local' assumption requires that spillovers are only experienced within a certain distance from treatment.

This paper then defines conditions under which each treatment effect is nonparametrically identified. I find that the switching effect is identified only with knowledge of the amount of 'exposure' that a unit experiences which requires *either* parametrization of spillovers *or* certain constraints on the spillover effects. The direct and total effects considered can be identified nonparametrically under mild assumptions about spillovers, namely that they are "local" in nature. The local assumption holds in the case where spillover effects are only experienced by units within a certain distance of treated units. This would hold in the case where individuals can travel to access treatment. However, this would fail in the case of large general equilibrium shocks that cause units far away to be affected by treatment (e.g. San Francisco being affected by shocks to the New York economy).

Next, I assess the commonly used difference-in-differences estimate. This paper shows that the difference-in-differences estimator is a biased estimate for the total effect and direct effect. Both the total effect and the direct effect are biased from spillovers on control units. This bias arises because untreated units that are 'close' to treated units experience effects of treatment and therefore these 'control' units fail to identify the counterfactual trend. Difference-in-differences estimates average the spillover onto the 'close' control units into the untreated units' change in outcomes. Hence, the spillover is subtracted from the estimated treatment effect and biases the estimate in the opposite sign of the spillover effect. For example, if nearby control units also benefit from treatment, then researchers would underestimate the treatment effect of the treated. The direct effect is also biased because changes in treated units' outcomes reflect the effect of their own treatment status and the effect from the treatment status of 'close' units. In difference-in-differences, the spillover

onto other treated units is averaged into the treated units' change in outcomes. Therefore the bias on the direct effect estimate is the same sign as the sign of the spillover.

Under the assumption that spillovers are local, I show a common adjustment of including an indicator for being affected by spillovers in the two-way fixed-effects model will remove *all bias* from the treatment effect estimate so long as the indicator captures all units affected by spillovers. The bias term in the difference-in-differences estimate for the total effect is the *average* spillover effects onto control units. An indicator variable that contains *all* control units affected by spillover can therefore remove this bias. When estimating the direct effect, a second indicator variable containing *all* treated units affected by spillovers can remove the second source of bias. Importantly, my method does not require researchers to make any assumptions about how spillover effects propagate across space. The only assumption required is selecting the maximum distance from treated units that spillovers can occur.[2]

The modified difference-in-differences framework with an added spillover indicator provides a poor estimate of spillover effects. The estimate averages spillover effects over all units with the indicator equal to 1 which can include units that are unaffected by spillovers. Therefore the estimate will be attenuated to zero. The above method can be improved by creating a set of distance bins from the treated units (e.g. being 0-20 miles, 20-40 miles, 40-60 miles from the treated units) and interacting them with a treatment indicator.[3] Since a set of 'rings' is collinear with a single indicator covering the bins, this estimator consistently estimates the treatment effect of interest. This estimator has the added benefit that each indicator will estimate the average spillover effect on treated/control units within that distance range which provides a more complete picture of which units are affected by treatment and at what magnitude.

To show the importance of considering spillover effects and the utility of my estimators, I revisit analyses of place-based policies in urban economics in Section 4. I revisit the analysis of the Tennessee Valley Authority by Kline and Moretti (2014). The Tennessee Valley

2. Since treatment effects typically decay over distance, this is not a strong assumption. Even if the maximum distance is not large enough, most of the bias will still be removed. On the other hand, making the maximum distance too large will increase the variance in the estimates as there are fewer remaining control units to identify the parallel counterfactual unit.

3. The choice of the distance bins depends on the economic context and in particular the source of the spillovers. In some cases, the bins can be determined in a data-driven way (Butts, 2021b).

Authority was a large-scale New Deal program that lowered the cost of power for indus-trial firms (Kitchens, 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 infrastructure and agglomeration economies (Severnini, 2014). I show that es-timation by difference-in-differences fails to account for these spillovers and therefore the authors obtain biased estimates of the total effect of the Tennessee Valley Authority.

Following this empirical application, I discuss how my framework fits into a larger dis-cussion 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. States apply for Empowerment Zone status for ares that meet a list of requirements but due to limited funding, not all proposed areas were accepted. 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 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 units being just outside will have parallel counterfactual trends. Al-though, 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 5, I extend estimation of the direct effect and spillover effects of treat-ment into the TWFE/event-study framework by extending the work of Gardner (2021). This allows for a very common setting in which treatment turns on for different units in different periods. Their proposed two-stage estimator first estimates group and time fixed-effects us-ing untreated/not-yet-treated observations. Since some control/not-yet-treated units can be affected by spillovers, these units must be removed to consistently estimate the group and time fixed-effects. Then these estimated unit and time fixed-effects are subtracted from the observed outcome *in the full sample*. The resulting residuals are then regressed on the

treatment and spillover variables to estimate treatment and spillover effects. In the online appendix, 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.

1.1. Relation to the Literature

I contribute to the literature that focuses on estimation of treatment effects with spatial spillovers using the difference-in-differences framework in two ways. (Clarke, 2017; Berg and Streitz, 2019; Verbitsky-Savitz and Raudenbush, 2012; Delgado and Florax, 2015). My paper is the first to consider identification in terms of general potential outcomes. Other papers derive results for specific spillover mappings and if I assume the particular functional forms for potential outcomes, I arrive at the same bias equation as theirs. The literature has focused on estimating the total effect, so I contribute by discussing alternative treatment effect estimates and identification results. Second, my paper also advances the literature by considering estimation of direct effects and spillover effects in the event-study framework which allows for the common occurrence of staggered treatment adoption.

This paper relates to a broad literature on spillover effect estimation in randomized experiments.[4] There are two main strands of this literature. First, there is a large literature on the 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.[5] 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 total effect (treatment effect on the treated) and spillover effects (treatment effect on the untreated in the treated clusters). However, in many settings distinct clusters are not plausible, e.g. counties near the border of a state can affect each other. My proposed

---

4.   See Sobel (2006) and Hudgens and Halloran (2008) for early work. See Hu et al. (2021), Sävje et al. (2021), and Vazquez-Bare (2019) and references therein for recent work on this.

5.   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.

5

estimator generalizes this method by using control units that are unaffected by treatment without requiring a notion of clusters.

There is also a nascent literature exploring estimation of direct and spillover effects which does not require a completely untreated cluster in experimental settings (Sävje et al. (2021), Vazquez-Bare (2019), and Hu et al. (2021)). Those papers' identification results rely on *design-based* assumptions around the treatment-assignment mechanism, while this paper relies on *model-based* assumptions, including a modified parallel-trends assumption, for identification in non-experimental settings. I contribute to this literature by thinking about identification of different treatment effects in non-experimental settings.

The rest of the paper is structured as follows. Section 2 presents the potential outcomes framework, defines the estimand of interest, discusses identification of effects, and shows the resulting bias from estimating a classical difference-in-differences model. Section 3 focuses on the estimation of the direct effect, the total effect, and spillover effects. Section 4 presents an application for evaluating the Tennessee Valley Authority program. Last, Section 5 discusses briefly how to incorporate spillovers into event-study estimation following the insights from Gardner (2021).

2 — Potential Outcomes Framework

This section will present a potential outcome framework that formalizes spatial spillovers. The goal of this exercise is to formalize treatment effects that are no longer uniquely defined in settings with spillovers. Following the canonical difference-in-differences framework, I observe a panel of units for periods $t \in \{0, 1\}$ and treatment turns on between periods for some of the units.[6] Let D_i be an indicator equal to one if unit i is treated and therefore treatment status is defined as $D_{it} = D_i * \mathbf{1}_{t=1}$.

Assumption 1 (Random Sampling). $\{Y_{i0}, Y_{i1}, D_i\}_{i=1}^{n}$ is an independent and identically distributed sample of panel data.

Assumption (1) allows me to consider observations as coming from a larger population

---

6.   This can be extended to multiple periods by replacing Y_{i0} and Y_{i1} throughout with pre- and post- sample averages of Y_{it}. This framework is extended to staggered treatment adoption in Section 5.

(sometimes called a superpopulation). Later, when discussing inference, I will relax this assumption to account for spatial correlation between observations, but identification does not rely on independence. As an alternative, inference could come from a design-based perspective following Athey and Imbens (2021).

Potential outcomes, denoted $Y_{it}(D_i, h_i(\vec{D}))$, are a function of unit i's own treatment-status D_i and, departing from the standard potential outcomes framework, of a function of the entire vector of treatment assignments $h_i(\vec{D})$ where $\vec{D} \in \{0,1\}^n$ denotes the n-dimensional vector of all unit treatments. The non-negative scalar- or vector-valued function $h_i(\vec{D})$ is referred to as an 'exposure mapping'. This setup follows Vazquez-Bare (2019) who studies identification of spillover effects under an experimental setting.[7]

The exposure mapping formalizes the SUTVA violation by summarizing the intensity at which unit i is affected by spillovers. Without any restrictions on the exposure mapping, nonparametric identification of treatment effects will be impossible. The main assumption for identification, as will be formalized in section 2.3, is that when unit i is sufficiently 'far' away, it has no exposure to spatial spillovers and $h_i(\vec{D}) = \vec{0}$. In the world of no spillovers, that is $h_i(\vec{D}) = \vec{0}$ for all individuals, potential outcomes would return to only depending on own-treatment status.

The identification results of this paper do not require parameterization of the exposure mapping. However, to help better understand the exposure mapping, I give three examples that are commonly used in the literature.

Example 1. $h_i(\vec{D})$ can be an indicator variable that equals one only if there is a treated unit within $\bar{d}$ miles of unit i.[8] Let $d(i,j)$ be a distance measure that is the distance between unit i and j. In this case

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

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

---

7. Note that throughout, expectations and inference are conditional on the vector of exposure mappings. In design-based settings, the distribution of exposure mappings is known so expectations can be taken without conditioning on the vector of exposure mappings. See Sävje et al. (2021) for discussion estimation and inference when the treatment design is known.

8. Similarly, a dummy for counties that share contiguous borders is commonly used.

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 do 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, $h_i(\vec{D})$ can be a function that equals the number of units treated within distance $\bar{d}$, i.e.

$$h_i(\vec{D}) = \sum_{j \neq i} D_j * \mathbf{1}_{d(i,j) < \bar{d}}. \tag{2}$$

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 wholesale 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, $h_i(\vec{D})$ is a spatial decay function where exposure decreases with distance. In this case, spillover intensity is the sum across all treated observations' decay terms, i.e.

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

This exposure mapping allows for the intensity of spillovers to depend on the distance to treatment and also is additive in the number of nearby units treated.[9] 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 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.

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

Often, researchers will try to parameterize potential outcomes by reasoning through the nature of the spillovers in a way similar to the above examples. Estimation in this case will rely on the functional form assumptions made by researchers. This paper will take an alternative approach to identification and estimation which will allow for estimation of treatment effects without requiring direct knowledge of the functional form of the exposure mapping/potential outcomes.

In particular, I will redefine the standard difference-in-differences identifying assumptions in the context of the new potential outcome framework. First, I assume that units do not adjust their actions in period 0 from knowledge of future treatment:

Assumption 2 (No Anticipation). $Y_{i0}(D, h) = Y_{i0}(0, 0)$ for all values of D and h.

Second, I formalize the parallel counterfactual trends assumption as follows:

Assumption 3 (Parallel Counterfactual Trends). Counterfactual trends do not depend on D_i:

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

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.

2.1. Treatment and Spillover Effects

In setups without spillovers, the treatment effect for an individual unit is well defined as $\tau_i \equiv Y_{i1}(1) - Y_{i1}(0)$, i.e. the effect of turning on treatment for a unit, and the above two assumptions would lead to the identification of the average effect by the traditional difference-in-differences estimator. In the presence of spillovers, multiple treatment effects can be defined in this setting. This subsection will define them and discuss their relative advantages/disadvantages.

The natural analogue to the above treatment effect which I will label the 'switching effect' is:

$$\tau_{i,switch}(h_i(\vec{D})) \equiv Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(0, h_i(\vec{D})).$$

This is the effect of changing only unit i's treatment status while keeping their exposure to spillovers constant. This treatment effect is policy-relevant as it summarizes what would happen if a 'local' policymakers decide to enact the policy for unit i (implicity keeping h constant).[10] It is important to note that the switching effect can depend on the level of exposure. For instance, consider the construction of libraries in towns. The effect of library construction in a town that is far away from any other library (exposure of $\vec{0}$) can be different than a town that is very close to a neighboring town's library (large exposure), hence the dependence of $\tau_{i,switch}$ on $h_i(\vec{D})$. Since the size of the switching effect depends on a unit's exposure level, we will consider an average switching effect at each value of exposure,

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

There is a particular switching effect at zero exposure that I call the 'direct effect' of treatment,

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

which measures the average effect of being treated *among unexposed treated units*. Specifying this particular switching effect will prove useful when discussing identification as it requires fewer assumptions than switching effects at other exposure levels. As discussed above, this is not necessarily a policy-relevant parameter for every unit if the switching effect has different magnitudes at different levels of exposure.

The second policy-relevant parameter is the 'total effect of treatment':

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

As opposed to the switching effect, which keeps exposure constant, the total effect looks at turning on treatment *and* going from zero exposure to $h_i(\vec{D})$ simultaneously. This can be thought of as going from the world with a *complete* absence of treatment, $\vec{0}$ to the current treatment vector $\vec{D}$. This treatment effect is helpful for 'national' policymakers to evaluate

10. This is what Sävje et al. (2021) call the 'assignment-conditional unit-level treatment effect'.

what *were* the effects of the entire vector of *enacted policies*.

To better understand how the above treatment effects are connected, it will be useful to formalize a unit's 'spillover effect' as:

$$\tau_{i,\text{spill}}(D_i, h_i(\vec{D})) \equiv Y_{i1}(D_i, h_i(\vec{D})) - Y_{i1}(D_i, \vec{0}).$$

The spillover effect measures the difference in potential outcomes between being exposed at intensity $h_i(\vec{D})$ 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. That is $\tau_{i,\text{spill}}(1, h_i(\vec{D}))$ does not necessarily equal $\tau_{i,\text{spill}}(0, h_i(\vec{D}))$. We can then average over treated/control units to form average spillover effects:

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

where the expectation is over individuals and hence over different exposures. This will give us the average spillover effect among *all* treated/control units regardless if they experience any spillover effects. It is often also of interest to estimate the average spillover effects on subsets of units (e.g. all units experiencing non-zero exposure). Estimation of spillover effects will be discussed in Section 3.2.

With spillover effects formally defined, the difference between the proposed estimands become more clear by the following algebraic identity:

$$\overbrace{Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(0, h_i(\vec{D}))}^{\text{Switching Effect}} =$$

$$\underbrace{\underbrace{Y_{i1}(1,0) - Y_{i1}(0,0)}_{\text{Direct Effect}} + \underbrace{Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(1,0)}_{\text{Spillover on Treated}}}_{\text{Total Effect}} - \underbrace{(Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0,0))}_{\text{Spillover on Control}}$$

From this identity, we see that whenever $\tau_{i,spill}(D_i, h(\vec{D}_i))$ does not depend on treatment status D_i, then we have that the switching effect is constant and equal to the direct effect. In this situation, the simpler identification assumptions behind the direct effect will identify the switching effect at all exposure levels. For example, if treatment is a factory opening

11

and the spillover effect on employment is due to increased spending on services, then it could be reasonable to assume the effect does not depend on treatment status and hence the switching effect is constant and equal to the direct effect.

However, the spillover effects can sometimes differ depending on treatment status. Continuing the above example, perhaps control units could experience negative employment shocks from a nearby country opening a factory if workers relocate/commute to the treated areas resulting in a negative spillover effect on control units. Other treated units could, for instance, experience positive spillover effects due to agglomeration economies. When it is the case that spillover effects depend on treatment status, the switching effect can be thought of as two separate effects. First, the unit is affected by changing its treatment status as summarized by the direct effect. Second, the unit changes from experiencing spillover effects as a control unit to experiencing spillover effects as a treated unit.

2.2. What Does Difference-in-Differences Estimate?

With treatment effects properly defined, I will first derive what the standard difference-in-differences estimand identifies under parallel trends assumption.

Proposition 1 (Decomposition of Difference-in-Differences Estimand).

If assumptions (1), (2) and Parallel Counterfactual Trends hold, the population difference-in-differences estimand can be decomposed as follows:

$$\underbrace{\mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 1\right] - \mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 0\right]}_{\text{Difference-in-Differences}} \tag{4}$$

$$= \tau_{\text{direct}} + \tau_{\text{spill}}(1) - \tau_{\text{spill}}(0) \tag{5}$$

$$= \tau_{\text{total}} - \tau_{\text{spill}}(0) \tag{6}$$

The proof is given in Appendix A, but the intuition behind the additional terms is 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, $\tau_{\text{spill}}(1)$. 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, $\tau_{\text{spill}}(0)$.

Researchers typically estimate the canonical two-way fixed-effects model,

$$y_{it} = \tau D_{it} + \mu_i + \lambda_t + \epsilon_{it}. \tag{7}$$

The estimator $\hat{\tau}$, which is well understood to be the sample analog of (4), is, therefore, a biased estimate for τ_{direct} and τ_{total} in the presence of spillovers. In particular, if a researcher estimates a difference-in-differences model and reports $\hat{\tau}$ as an estimate of the total effect, their estimate would be biased in the presence of spillover effects on control units. If spillover effects are of the same sign as the total effect, then estimates would be attenuated towards zero. On the other hand, if spillovers are the opposite sign, then results would increase in magnitude.

2.3. Identification of Treatment Effects

The following subsection will talk about the identification of treatment effects and present corresponding estimators for each effect.

Switching Effect

It is difficult to identify the switching effect in general as identifying the unobserved counterfactual outcome $Y_{i1}(0, h_i(\vec{D}))$ requires knowledge of which control units have the same level of exposure. The reason for this can be seen by rewriting this in terms of a difference-in-differences style estimand:

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

Writing it in this way highlights the two problems in estimating the second term. First, this requires knowledge of the unobserved exposure mapping to be correctly specified to identify control units with $h_i(\vec{D}) = h$. More, this estimand requires a more demanding parallel

trees assumption.[11]

Assumption 4 (Switching Effect Parallel Trends). For a given exposure, $\vec{h}$, counterfactual trends do not depend on D_i:

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

This assumption requires parallel trends among treated and control units at exposure $\vec{h}$. This assumption is implied by Parallel Counterfactual Trends and an additional assumption that spillover effects are equal between control and treated units that experience exposure $\vec{h}$. Under this assumption, the switching effect can be identified.

Proposition 2 (Identification of the Switching Effect). For a given exposure, $\vec{h}$, if assumption (1), (2), and Switching Effect Parallel Trends holds, then

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

The switching effect can be estimated by a difference-in-differences estimator on the subsample of units with exposure $\vec{h}$. However, this approach relies on correct specification of the exposure mapping which is a difficult assumption to maintain in practice. Researchers may be tempted to introduce the parameterized exposure mapping linearly to "compare individuals at the same level $\vec{h}$". However, this assumes additive separability between the treatment dummy and the exposure mapping. Additive separability would imply the switching effect does not depend on exposure level, i.e. the switching effect is constant across $\vec{h}$.

Direct and Total Effects

To identify the direct and total effects, additional assumptions on the nature of spillover effects are needed. In particular, I will formalize the idea that spillovers are 'local' in that

---

11. This problem is similar to those highlighted by Callaway et al. (2021) for difference-in-differences with continuous treatment.

units are only affected by treatment if they are near treatment. For example, if treatment has to be accessed in person, then it is natural that further away places are not affected by treatment. However, this assumption may fail under general equilibrium shocks that do not necessarily decay over distance.

Assumption 5 (Spillovers Are Local). There exists a distance $\bar{d}$ such that

(i) For all units i,

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

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

Part (i) of assumption (5) requires that spillovers are 'local' in that units are no longer exposed to spillovers after some maximum distance $\bar{d}$. Part (ii) of Assumption (5) requires that there exists control and treated units with no exposure. This assumption is far less strict than the identifying assumption for the switching effect in that we only need to identify units with non-zero exposure and don't need to parameterize the exposure mapping any further. For a given $\bar{d}$ that satisfies Assumption (5), let S_i be an indicator equal to one if $\min_{j:\ D_j=1} d(i,j) \leq \bar{d}$, i.e. unit i is within $\bar{d}$ miles of the closest treated unit.

The intuition behind identifying the direct effect separately from the spillover effects requires that we use control units experiencing no spillover effects to identify the counterfactual trend and use treated units experiencing no spillover effects to identify the direct effect. Since we are leveraging subgroups of treated and control units, we need a parallel trends assumption for each subgroup:

Assumption 6 (Direct/Total Effect Parallel Trends). Counterfactual trends do not depend on D_i and S_i:

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

It is worth remarking on Direct/Total Effect Parallel Trends in terms of applied work. Often, individuals use only a subsample of observations that are close to treated observations in estimating their difference-in-differences model. The idea is that close units are more likely to be on similar counterfactual trends than units that are further away. The identification result and its corresponding estimator are formed by using further away units (i.e. units

with $S_i = 0$). In this case, if parallel trends only hold locally, then focusing on further away units can fix the spillover bias but increase bias from non-parallel trends.

Note that for identification of the total effect, you only need that the counterfactual trend among *all* treated units is the same as the counterfactual trend among control units with $S_i = 0$ which is implied by Direct/Total Effect Parallel Trends. Identification of the direct effect will be average the individual treatment effects *only* for treated units with $S_i = 0$. So identification of the average treatment effect among *all* treated units requires a subtle assumption to ensure treatment effect homogeneity.

Assumption 7 (Direct Effect Does Not Depend on S_i). For $s \in \{0, 1\}$,

$$\mathbb{E}\left[Y_{i1}(1, \vec{0}) - Y_{i1}(0, \vec{0}) \mid D_i = 1, S_i = s\right] = \mathbb{E}\left[Y_{i1}(1, \vec{0}) - Y_{i1}(0, \vec{0}) \mid D_i = 1\right]$$

With these assumptions, we have the following identification result for the direct and total effects:

Proposition 3 (Identification of Direct and Total Effects). If assumption (1), (2), Spillovers Are Local, and Direct/Total Effect Parallel Trends) hold, then (i)

$$\mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 1\right] - \mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 0, S_i = 0\right] = \tau_{\text{total}} \tag{8}$$

and (ii)

$$\begin{aligned}
\mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 1, S_i = 0\right] &- \mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 0, S_i = 0\right] \\
&= \mathbb{E}\left[Y_{i1}(1, \vec{0}) - Y_{i1}(0, \vec{0}) \mid D_i = 1, S_i = 0\right] \tag{9} \\
&= \tau_{\text{direct}},
\end{aligned}$$

where the last equality holds only under assumption (7).

The proof is given in Appendix A, but the intuition is simple. Conditioning on $S_i = 0$ will use only observations with $h_i(\vec{D}) = 0$ which gives rise to the identification of the direct and total effect. This proposition shows that a researcher does not need to know how the spillovers occur over space to have an unbiased estimate for treatment effects. All that

is required to identify the direct or total effects is being able to classify which units are considered 'close' to treatment.[12]

3 — Estimation of Effects

3.1. Estimation of Direct and Total Effects

To estimate the direct effect and the total effect, we could replace the terms in proposition 3 with their sample analogs. However, it will prove advantageous to modify the two-way fixed-effects model to estimate the direct or the total effect because it will allow the estimation of spillover effects as well. To estimate the total effect, the modified model is given by:

$$y_{it} = \tau D_{it} + \gamma_0 (1 - D_{it}) S_i + \mu_i + \lambda_t + \epsilon_{it}. \tag{10}$$

To estimate the direct effect, the modified model is given by:

$$y_{it} = \tau D_{it} + \gamma_0 (1 - D_{it}) S_i + \gamma_1 D_{it} S_i + \mu_i + \lambda_t + \epsilon_{it}. \tag{11}$$

From standard regression analysis the estimates of $\hat{\tau}$ from equations (11) and (10) are sample analogues of (9) and (8).

Since $\hat{\tau}$ is a linear combination of sample averages, inference can be done using large-n asymptotics. However, it is important to account for two sources of correlation in the error structure. First, errors should be clustered by unit i to allow for correlation across periods. Second, since our model is predicated on the fact that nearby places affect one another, we should account for such spatial correlation by allowing for spatial correlations following Conley (1999). More recent work by Ferman (2020) show that for large-n asymptotics to be used, the structure of spatial correlation must be limited in that errors are assumed to be uncorrelated after a certain cutoff distance. A natural candidate for this cutoff would be $\bar{d}$ used for the creation of S_i.

A natural question is why wouldn't researchers make the distance large to guarantee they

---

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.

have an unbiased estimate of τ_{direct} or τ_{total}? Equation (8) and (9) should make this problem clear. Since these estimates rely on averages among units with $S_i = 0$, as $\bar{d}$ increases, the number of units with $S_i = 0$ decreases which yields 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.[13] 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. In separate work, Butts (2021b) discusses data-driven methods to determine how far S_i should be under a more strict parallel trends assumption.

3.2. Estimation of Spillover Effects

The estimates of γ_D in the modified TWFE models (10) and (11) do not estimate the corresponding average spillover effects $\tau_{\text{spill}}(D)$ for $D = 0, 1$. This is because $\hat{\gamma}_D$ average spillover effects across only units with $S_i = 1$:

$$\mathbb{E}\left[\hat{\gamma}_0\right] = \mathbb{E}\left[\tau_{i,spill}(0) \mid S_i = 1, D_i = 0\right], \text{ and}$$

$$\mathbb{E}\left[\hat{\gamma}_1\right] = \mathbb{E}\left[\tau_{i,spill}(1) \mid S_i = 1, D_i = 1\right].$$

These estimates average spillover effects for all units with $S_i = 1$. Since some units with $S_i = 1$ do not have positive exposure, this is not the average spillover effect for units with non-zero exposure either. 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 some units experience substantial spillover effects.

It is possible to more accurately estimate spillover effects by using a set of concentric ring indicators (e.g 0-20 miles, 20-40 miles, and 40-60 miles from the nearest treated unit) instead of a single large ring indicator to more flexibly estimate the spillover effects. For example, for the total effect, equation (10) can be modified as

$$y_{it} = \tau D_{it} + \sum_{i=1}^{n}(1 - D_{it})\text{Ring}_i + \mu_i + \lambda_t + \epsilon_{it}, \tag{12}$$

---

13. Although, spillover effects typically will grow weaker over distance and if distant units are mistakenly treated as if they have zero exposure, bias should be small.

Figure 1 — Comparison of Single vs. Multiple Rings Estimation of Spillover Effects
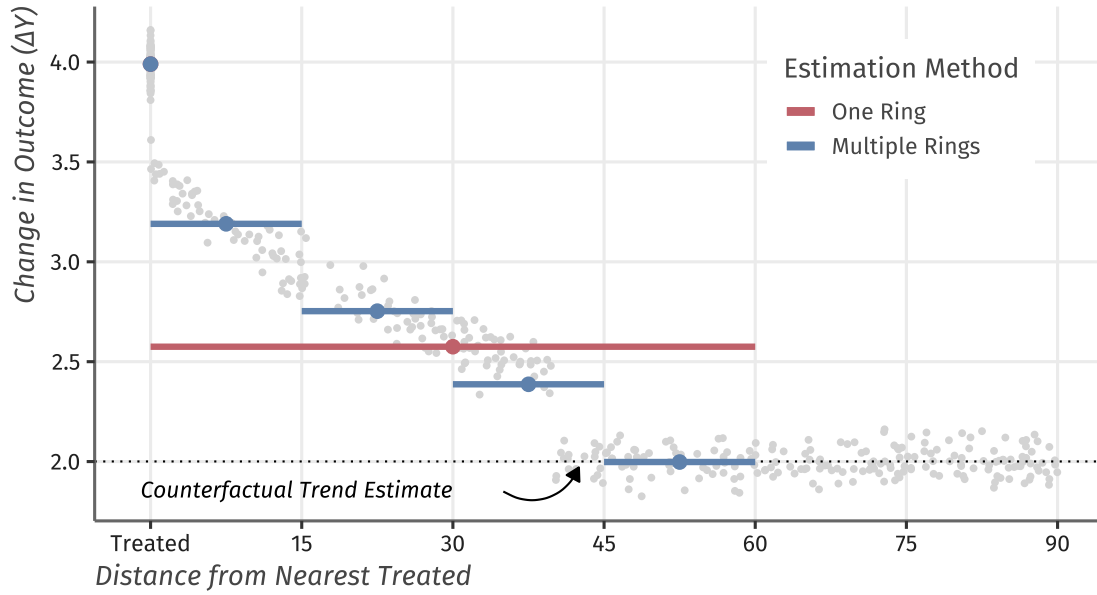
Notes: The figure compares estimates of equation (10) with S_i a ring of $(0, 60]$ miles from the nearest treated unit and of equation (12) with rings for $(0, 15], (15, 30], (30, 45], (45, 60]$ miles for a simulated set of data.

where Ring$_i$ is a given ring. Similarly, equation (11) can be modified to include multiple rings. Since the set of rings is collinear with a single large ring, the direct and total effects estimate will continue to be identified. The estimates for each ring will represent the average of $\tau_{i,spill}$ for all observations within a given distance ring. Having estimates of average spillover effects at different distances allows a researcher to better understand the nature of the spillovers.

A simple example of how rings improve the estimation of spillover effects can be seen in Figure 1. The figure displays observations of $Y_{i,1} - Y_{i,0}$ on the y-axis and the distance to the nearest treated unit on the x-axis. 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 the estimate of a single ring in two ways. Multiple rings are better able to capture the decaying nature of spillover effects over distance. Furthermore, the ring that is included past the true maximum distance where spillover effects occur (45-60 miles) has a point estimate near zero. The rings method is 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 $\bar{d}$ satisfying assumption (5)). However, the precision of estimates would decrease as the number of rings increases and therefore there is a bias-variance trade-off in the number and width of rings.[14]

One disadvantage of using a set of indicators 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 the 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, separate estimation of the direct effect and spillover effects would be the preferred method.

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 (2014). The TVA program was a large-scale federal investment started in 1934 that focused on the 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 facto-

---

14. 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. Butts (2021b) provides a data-driven method to selecting rings under a more strict version of the parallel trends assumption.

ries.[15] 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 (2014) analyze 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 (2014) 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

$$y_{c,t} - y_{c,1940} = \alpha + \text{TVA}_c \tau + X_{c,1940}\gamma + (\varepsilon_{c,t} - \varepsilon_{c,1940}), \tag{13}$$

where c denotes county, TVA_c is an indicator variable for being in the Authority, and y is a set of outcome variables (in logs).[16] Pre-treatment control variables, $X_{c,1940}$, are interacted with Post_t to allow for places to be on different long-term trends.[17] To improve the likelihood of the parallel 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 (2014) estimate (13) 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.

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,

15. More details on the program are found in Kline and Moretti (2014). The effects on wholeseale electricity are discussed in Kitchens (2014).

16. The two-period 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. Their estimator does not differ much from the standard difference-in-differences results since the weights are not that different from uniform weights.

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

21

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.

the sign is ambiguous. There could be positive spillovers if electrification brought cheap power and agglomeration economies to the neighboring areas. 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. 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 Authority 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, but due to the scale of the program, the spillovers are likely to extend further than this. Therefore, 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 specification. To parametrize the exposure mapping, I use a set of rings as described in

Table 1 — Effects of Tennessee Valley Authority on Decadal Growth

Dependent Var.	Diff-in-Diff	Diff-in-Diff with Spillovers				
	TVA (1)	TVA (2)	TVA between 0-50 mi. (3)	TVA between 50-100 mi. (4)	TVA between 100-150 mi. (5)	TVA between 150-200 mi. (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)	(0.0281)	(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)	(0.0403)	(0.0297)
Manufacturing employment	0.0894***	0.0993**	0.0228	0.0225	−0.0055	−0.0066
	(0.0324)	(0.0432)	(0.0510)	(0.0561)	(0.0347)	(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 (13) by OLS and is similar to the estimate reported in Kline and Moretti (2014). The final four columns labeled 'Diff-in-Diff with Spillovers' are estimates from (14).

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

Section 3.2. Specifically, the specification with spillovers is given as follows:

$$y_{i,t} - y_{i,1940} = \alpha + \mathsf{TVA}_i\tau + \sum_{d \in \mathsf{Dist}} \mathsf{Between}(d)\delta_d + X_{i,1940}\beta + (\varepsilon_{i,t} - \varepsilon_{i,1940}), \quad (14)$$

where $\mathsf{Dist} = \{(0, 50], (50, 100], (100, 150], (150, 200]\}$ measured in miles and define $\mathsf{Between}(d)$ as an indicator for being within the interval $d \in$ Dist away from the Authority and $t \in \{1960, 2000\}$. Figure 2 displays the four spillover variables by filling in each distance bin in a different color. The coefficients δ_d estimate the average spillover effect onto control units for each of these distance bins. From proposition 3, $\hat{\tau}$ will be an unbiased estimate for the 'total' effect so long as spillovers do not occur past 200 miles from the TVA.

The results of the long run analysis from 1940 to 2000 are presented in Panel A of Table 1. Each row contains the results for a different outcome variable (measured in logs). Columns (1)-(5) contain point estimates for τ and δ_d's in different specifications. The point estimates can be interpreted as decadal growth rates in outcomes. The column labeled difference-in-differences uses an ordinary least squares estimator for the specification without spillovers, equation (13). 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 (14), column (2) contains a point estimate for τ 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 because higher-paying manufacturing jobs within the Authority drew farm-worker migrants from nearby counties. Because the spillovers onto the control counties are 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 suggests 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 1 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 whole-

sale 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 spill 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.

4.1. Identification Strategies in Place-Based Policy Analysis

More generally, my framework provides important insights into identification strategies when analyzing 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 the problem (e.g. aggregating census tracts to metropolitan areas). However, this approach combines the treatment and spillover 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 into 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 other side of the border. Other times, researchers compare approved applicants to narrowly rejected applicants. Both strategies aim to balance unobservables between treated and control units. 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.

25

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.[18]

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 and 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 and 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 affects units on the other side of a border, then estimated treatment effects can be biased.

5 — Estimating Event Study with Spillovers

Now, I turn to the common setting where treatment turns on at different times for different units, often called staggered adoption. The intuition from how spillovers cause biases in estimates of the total/direct effect of treatment extends into the staggered adoption setting. In this setting, the literature has shown that two-way fixed-effect estimates can be viewed

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

as a weighted sum of 2×2 difference-in-differences estimates.[19] Therefore the bias terms will be an identically weighted sum of the bias term(s) from the 2×2 estimates, assuming that the parallel counterfactual trends assumption (3) holds. However, since weights on some of the 2×2 estimates can be negative, the sign of the spillover effects does not determine the sign of the weighted average of the spillover effects. This makes the bias from spillovers much more difficult to sign.

To estimate treatment and spillover effects in the presence of spillovers and staggered treatment timing, I will propose an estimation strategy that follows the 'imputation-based' approach proposed in concurrent work by Gardner (2021) and Borusyak et al. (2021) by incorporating spillovers directly into estimation similar to equations (8) and (9). I assume I observe an independent and identically distributed panel for units i and periods t. The imputation-based method relies on a model-based assumption for untreated and unexposed potential outcomes to formalize parallel trends:

Assumption 8 (Parallel Counterfactual Trends (Staggered)). For all units and all periods, the untreated and unexposed outcome is given by

$$Y_{it}(0, \vec{0}) = \mu_i + \lambda_t + \varepsilon_{it}, \tag{15}$$

where $\mathbb{E}\left[\varepsilon_{it}\right] = 0$, μ_i and λ_t are non-stochastic unit and time fixed-effects.

This formalizes parallel trend by imposing that there is a common time-trend in the absence of treatment, λ_t. Note that this does not impose any structure on the effects of treatment and exposure $(D_{it}, h_i(\vec{D}))$. I also assume that units do not have any anticipatory effects before treatment starts:

Assumption 9 (No Anticipation (Staggered)). For all observations with $D_{it} = 0$ and $h_i(\vec{D}) = 0$, $Y_{it}(D, \vec{h}) = Y_{it}(0,0)$ for all values of D and $\vec{h}$.

With a model for $Y_{it}(0, \vec{0})$, treatment effects for individual i and time t under treatment status D_{it} and exposure $h_i(\vec{D})$ are given by $\tau_{it}(D_{it}, h_i(\vec{D})) = Y_{it}(D_{it}, h_i(\vec{D})) - Y_{it}(0, \vec{0})$. Then,

---

19. 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.

the total effect and the direct effect are formed similar to above:

$$\tau_{\text{total}} \equiv \mathbb{E}\left[\tau_{it}(1, h_i(\vec{D})) \mid D_{it} = 1\right] \text{ and } \tau_{\text{direct}} \equiv \mathbb{E}\left[\tau_{it}(1, 0) \mid D_{it} = 1, h_i(\vec{D}) = \vec{0}\right].$$

However, it is also common in event-study analyses to allow heterogeneity in effects by the number of years that a unit experiences treatment. Let K_{it} denote the number of years since treatment turns on (-1 for the year prior, 0 for the initial year, and so on). Then we can define dummy variables $D_{it}^k \equiv D_{it} \mathbf{1}_{K_{it}=k}$ and estimate average total effects for each relative period:

$$\tau_{\text{total}}^k \equiv \mathbb{E}\left[\tau_{it}(1, h_i(\vec{D})) \mid D_{it}^k = 1\right],$$

and similarly for the direct effect.

Estimation of τ_{it} is the difference between the observed Y_{it} and the unobserved outcome $Y_{it}(0, \vec{0})$ which we will estimate using observations with $D_{it} = 0$ and $h(\vec{D}, 0) = \vec{0}$. Similar to the previous section, under assumption (5) we define S_{it} to be an indicator equal to one if in period t, unit i is within $\bar{d}$ distance from the nearest treated unit. Following the procedure laid out in Gardner (2021), I propose a modified version of the two-stage difference-in-differences estimator:

1. Estimate $Y_{it} = \mu_i + \lambda_t + u_{it}$ for observations with $D_{it} = 0$ and $S_{it} = 0$ to estimate the common trend λ_t and unit fixed-effects μ_i. For all observations, residualize $\tilde{Y}_{it} \equiv Y_{it} - \hat{\mu}_i - \hat{\lambda}_t$ which is our estimate for τ_{it}.

2. Regress $\tilde{Y}_{it}$ on treatment and spillover dummy variables (discussed in table (2)).

The first stage uses untreated and unexposed outcomes to estimate (15). Then, $\tilde{Y}_{it}$ will serve as an estimate for τ_{it}. The second stage estimates averages of τ_{it} for different groups depending on the parameters of interest. This is equivalent to regressing the residual $\tilde{Y}_{it}$ on dummy variable(s). Table 2 provides examples of treatment effects of interest and corresponding dummy variables to include.

As discussed in Gardner (2021), inference must account for the fact that the regressand in the second stage is estimated in the first stage. Gardner (2021) propose reframing the two-stage process as a two-stage GMM estimator and discusses how to perform valid inference

28

Table 2 — Second Stage Variables

Estimand	Included Variables
Total Effect	D_{it}
Total Effect (Event Study)	D_{it}^k dummies
Direct Effect	$D_{it}(1 - S_{it})$
Direct Effect (Event Study)	$D_{it}^k(1 - S_{it})$ dummies
Spillover Effect on Control	$S_{it}(1 - D_{it})$ or $\text{Ring}_{it}(1 - D_{it})$
Spillover Effect on Control (Event Study)	$S_{it}^k(1 - D_{it})$ or $\text{Ring}_{it}^k(1 - D_{it})$

in this setting. This procedure is implemented in the R/Stata package did2s (Butts, 2021a). In appendix section (B), I use this two-stage estimator to revisit the analysis of Bailey and Goodman-Bacon (2015).

6 — Discussion

This paper has considered the common environment where treatment is assigned via administrative boundary while the effects of treatment spread across these borders. I model this phenomenon in a potential outcomes framework and show why difference-in-differences estimation does not identify any treatment effect of interest. I then improve upon ad-hoc estimation strategies by proposing 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 will remove all bias in the treatment effect estimate so long as all units affected by spillovers are contained in the indicator under a modified parallel trends assumption. More, I show that a set of concentric 'rings' can remove all the bias and improve the estimation of spillover effects. A limitation of this research is in deciding how wide and how many rings to include in estimation. Concurrent work by Butts (2021b) discusses data-driven ring selection under a more stringent parallel trends assumption that does not readily apply in the context of large geographic units such as counties.

Then, I show that in settings with spatial spillovers, estimates can change significantly. In particular, place-based interventions 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 general equilibrium effects. I also show the importance of considering spillovers in weighing the pros and cons of various identification strategies. Identification strategies based on geographic continuity of unobservables can magnify the bias from spillovers as they restrict the comparison group to observations experiencing the largest spillover effects.

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.

Athey, Susan, and Guido W. Imbens. 2021. "Design-based analysis in Difference-In-Differences settings with staggered adoption." *Journal of Econometrics*. https://doi.org/ 10.1016/j.jeconom.2020.10.012.

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."

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2021. "Revisiting Event Study Designs: Robust and Efficient Estimation." 48.

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.

Butts, Kyle. 2021a. *did2s: Two-Stage Difference-in-Differences Following Gardner (2021).* https://github.com/kylebutts/did2s/.

Butts, Kyle. 2021b. "Difference-in-Differences with Geocoded Microdata."

Callaway, Brantly, Andrew Goodman-Bacon, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with a Continuous Treatment." *arXiv:2107.02637 [econ]*, http://arxiv.org/abs/2107.02637, arXiv: 2107.02637.

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.

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.

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.

Ferman, Bruno. 2020. "Inference in Differences-in-Differences: How Much Should We Trust in Independent Clusters?" *arXiv:1909.01782 [econ]*, http://arxiv.org/abs/1909.01782, arXiv: 1909.01782.

Gardner, John. 2021. "Two-Stage Difference-in-Differences." Technical report.

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.

Hu, Yuchen, Shuangning Li, and Stefan Wager. 2021. "Average Treatment Effects in the Presence of Interference." *arXiv:2104.03802 [econ, stat]*, http://arxiv.org/abs/2104.03802, arXiv: 2104.03802.

Hudgens, Michael G, and M. Elizabeth Halloran. 2008. "Toward Causal Inference With Interference." *Journal of the American Statistical Association* 103 (482): 832–842. 10.1198/016214508000000292.

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. 2014. "Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority." *The Quarterly Journal of Economics* 129 (1): 275–331. 10.1093/qje/qjt034.

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. 2021. "Average treatment effects in the presence of unknown interference." *The Annals of Statistics* 49 (2): . 10.1214/20-AOS1973.

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

A.1. Proof of Proposition 1

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

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

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

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

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

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

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

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

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

$$\equiv \tau_{\text{direct}} + \tau_{\text{spill}}(1) - \tau_{\text{spill}}(0) = \tau_{\text{total}} - \tau_{\text{spill}}(0)$$

For part (i):

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

where the first equality comes from the fact that $S_i = 0$, the second equality from parallel trends assumption (6), and the last by definition of the total effect.

Similarly, for part (ii):

$$
\begin{aligned}
\mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 1, S_i = 0\right] &- \mathbb{E}\left[Y_{i1} - Y_{i0} \mid D_i = 0, S_i = 0\right] \\
&= \mathbb{E}\left[Y_{i1}(1,0) - Y_{i0}(0,0) \mid D_i = 1, S_i = 0\right] - \mathbb{E}\left[Y_{i1}(0,0) - Y_{i0}(0,0) \mid D_i = 0, S_i = 0\right] \\
&= \mathbb{E}\left[Y_{i1}(1,0) - Y_{i0}(0,0) \mid D_i = 1, S_i = 0\right] - \mathbb{E}\left[Y_{i1}(0,0) - Y_{i0}(0,0) \mid D_i = 1, S_i = 0\right] \\
&= \mathbb{E}\left[Y_{i1}(1,0) - Y_{i0}(0,0) - Y_{i1}(0,0) + Y_{i0}(0,0) \mid D_i = 1, S_i = 0\right] \\
&= \mathbb{E}\left[Y_{i1}(1,0) - Y_{i1}(0,0) \mid D_i = 1, S_i = 0\right] \\
&= \mathbb{E}\left[Y_{i1}(1,0) - Y_{i1}(0,0) \mid D_i = 1\right] \\
&\equiv \tau_{\text{direct}},
\end{aligned}
$$

where the first equality comes from the fact that $S_i = 0$, the second equality from parallel trends assumption (6) and the second to last equality follows from assumption (7).

B — 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

$$Y_{it} = \theta_i + X_{it}\beta + \sum_{k=-7}^{-2} \pi_y D_{it}^k + \sum_{k=0}^{15} \tau_y D_{it}^k + \varepsilon_{it}, \tag{B.1}$$

where D_i^k is the typical 'event study' indicator for being treated for k years, θ_i is county fixed-effects, and X_{it} contains a set of controls.[20] The coeffcents 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. 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.

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 emergencies 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 are 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 an indicator for being within 25 miles of a treated county as the spillover variable. The results are presented in Figure B1. The confidence intervals labeled with circles represent point estimates for the average spillover effect on

---

20. 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 B1 — Direct and Spillover Effects of Community Health Centers
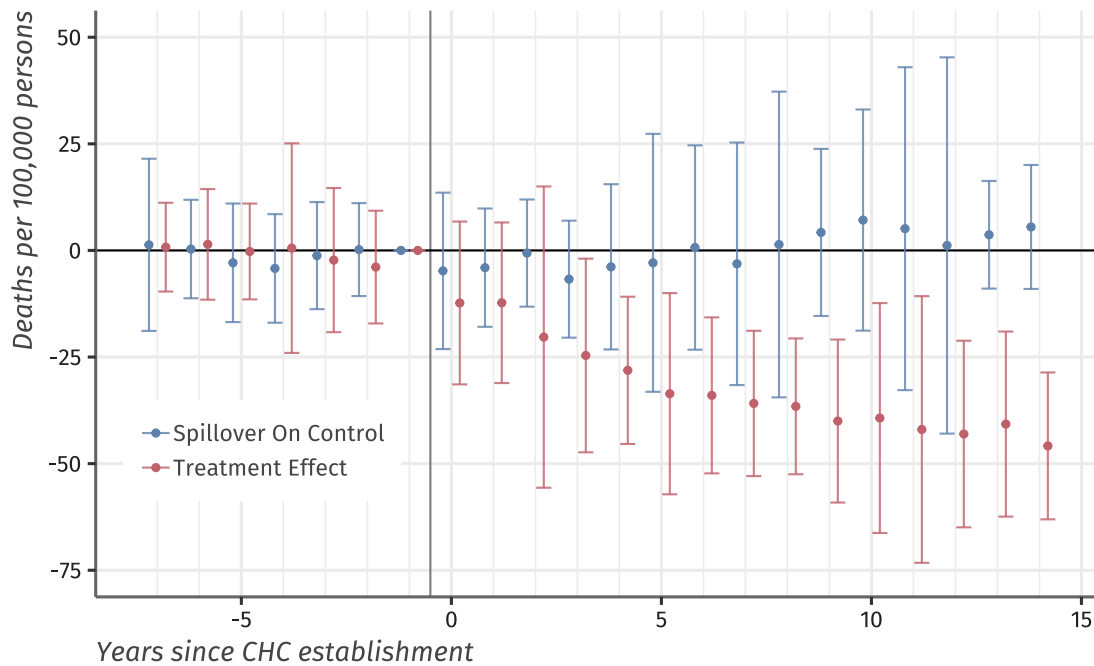
Notes: This figure plots event study estimates for the total effect and the spillover effect on control units within 25 miles of treatment at different periods relative to establishment year. The estimates are generated using the 'did2s' package (Butts, 2021a).

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 B1 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.

C — Monte Carlo Simulations

C.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 total 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_{it} for many different exposure mappings $h_i(\vec{D})$. I generate data for unit i is a US county at time $t \in \{1, \ldots, 20\}$ using the following data-generating process:

$$y_{it} = \lambda_t + \mu_i + 2D_{it} + \beta_{\text{spill, control}}(1 - D_{it})h_i(\vec{D}) + \varepsilon_{it}, \tag{C.1}$$

where $\lambda_t \sim N(0.2t, 0.1^2)$ and $\mu_i \sim N(6, 2^2)$ respectively and the error term is IID and distributed as $\varepsilon_{it} \sim N(0, 2^2)$. For simplicity, I also remove spillovers onto treated units, but the results of which specifications perform best in the simulation are the same. to keep the magnitude of bias constant across specifications, I normalize the average spillover magnitude to be equal across specifications.[21] For each true data-generating process, I estimate (C.1) 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 counties center of population. This corresponds to different choices of $\bar{D}$ and hence S_i in the text. '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

---

21. This results in a constant bias for TWFE estimates across data-generating processes.

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 (C.1) 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 $\sum_{i=1}^{n} \frac{(\hat{\tau}_i - \tau_0)^2}{n}$

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 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. Hence, if the direct effect is of principal concern, then the non-additive rings estimate is preferred.

The results of the simulations are produced in Table C1. 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 (C.1) 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 estimates 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 (see Proposition 3). These specifications do so without increasing the variance of the estimator relative to the true specification. 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.

C.2. Spillover Effect Estimation

Now, I turn to analyze 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),

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

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

Results of this exercise are presented in Table C2. For each specification, correctly specifying the spillovers produces the best estimates for spillovers as expected. Among misspecified exposure mapping, a 'Rings' specification that is large enough to capture 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

39

Table C2 — Percent of Spillovers Predicted by Specification

Specification	Within 40mi	Within 80mi	Within 40mi. (Additive)	Within 80mi. (Additive)	Decay	Decay (Additive)
			Data-Generating Process			
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%

Notes. Each cell corresponds to the results from 1000 simulations of the data generating process specified by the column header and estimated by equation (C.1) 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 Equation C.2.

the non-additive ring specifications do not as accurately measure these spillovers as they did before.[22] Non-additive specifcations perform somewhat poorly at estimating spillovers in the additive data-generating processes, 'Within (Additive)' and 'Decay (Additive)'. The 'Rings (Additive)' specification successfully predicts spillovers in both additive data-generating processes.

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