Differences-in-Differences with Spatial Spillover

Kyle Butts

Univ. of Colorado, Boulder

February 1, 2021

ABSTRACT

Empirical work often uses treatment variables defined by geographic boundaries. When researchers ignore the common problem that the effects of treatment cross over borders, classical difference-in-differences estimation produces biased estimates for the average treatment effect. In this paper, I decompose this bias in two parts. First, the control group no longer identifies the counterfactual trend because their outcomes are affected by treatment. Second, changes in treated units' outcomes reflect the effect of their own treatment status and the effect from the treatment status of "close" units. I use Monte Carlo simulations to demonstrate when the biases are particularly large. Further, I show that a common solution used in the literature of removing 'contaminated' control units only prevents the first source of bias. Lastly, I propose improved estimation strategies that can remove both sources of bias. I highlight the importance of considering spillovers by revisiting analysis of the Tennessee Valley Authority.

1 — Introduction

Empirical work in economics often considers treatment assigned by geographic boundaries such as cities, counties, and states. Potentially, the effect of these treatments do not stay within these boundaries whether it be from people crossing borders to access treatment (imperfect compliance) or from general equilibrium effects influencing neighboring areas. Despite the problem of spatial spillovers being common across many settings, Berg and Streitz (2019) document that little empirical analysis takes the problem seriously. In a survey of eight top economics and finance journals in 2017, they find that only 21 articles out of 108 that run difference-in-differences estimation discuss spillovers.[1] In this paper, I show that failing to account for spatial spillovers can lead to poorly identified estimates of treatment effects.

In the causal inference literature, researchers assume that a unit's potential outcome does not depend on the treatment status of any other unit – what Rubin (1980) calls the Stable Unit Treatment Value Assumption (SUTVA). This article considers a common violation of SUTVA in a process I label 'spatial spillovers' where the treatment assignment of 'close' units' affects outcomes of nearby units.[2] Typically researchers are interested in the 'direct effect' on the treated units.[3] In the presence of spatial spillovers, I show using a potential outcomes framework that the difference-in-differences estimate identifies the direct effect of treatment (the estimand of interest) plus two additional bias terms resulting from the spillovers.

The intuition for the two bias terms resulting from spillovers is as follows. First, untreated units that are 'close' to treated units experience effects of treatment and therefore these 'control' units fail to identify the counterfactual trend. When estimating by difference-in-differences, the spillover onto the 'close' control units is averaged into the

---

[1] The articles surveyed are American Economic Review, Econometrica, the Journal of Political Economy, the Quarterly Journal of Economics, the Review of Economic Studies, the Journal of Finance, the Journal of Financial Economics, and the Review of Financial Studies.

[2] Close can refer to many different things, e.g. geographic distance, economic distance such as supply chains, node distance in a graph, or social relationships in schools or cities.

[3] This is sometimes called the 'partial equilibrium' effect or the 'local effect'.

untreated units' change in outcomes. In this case, the spillover is subtracted from the estimated treatment effect and biases the estimate in the opposite sign of the spillover effect. For example, consider a factory opening in a county. Agglomeration economies suggest that neighboring counties would also benefit from knowledge spillovers and improved access to supply chain (Duranton and Puga, 2003). The treatment effect estimate is negatively biased because the change in outcome in neighboring counties is higher than it would be absent treatment. Researchers would therefore underestimate the positive benefits of treatment.[4]

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

These biases can be written as functions of the potential outcomes in my framework and can be used by researchers to sign or bound the bias using insight from their economic context. The magnitude of the bias depends on two factors: the size of the spillovers and the number of units affected by spillovers. If spillovers are quite large in magnitude or spread far over distance, then the bias will generally be large. The results of this paper enable researchers to estimate bounds for it by making assumptions about the average size of the spillover effect and the number of units affected. I use Monte Carlo simulations to quantify the magnitude of bias in various spillover settings.

In the survey by Berg and Streitz (2019), only eight of the 21 papers that discuss spillovers control for this potential source of bias in their primary empirical specification to

---

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

prevent bias in the estimation of treatment effects. The other thirteen papers usually first estimate the direct effect without controlling for spillovers and then seperately estimate the spillover effects. This paper shows that not controlling for spillovers produces biased estimates of the direct effect.

Of the eight papers that include spillovers directly in their estimation strategy, six simply drop the control units they suspect to be affected by spillovers. Since this is a common strategy used by researchers, I use Monte Carlo simulations to show that this method is effective in the case where spillovers *only* occur onto control units. In the presence of spillover onto also treated units, the estimates for the direct effect will still be biased. Dropping control unit observations, however, makes estimates less precise, so parameterizing the spillover function may be a better strategy especially in the case where a high proportion of control units would need to be removed.

Removing both sources of bias requires a researcher to take a stance and parameterize a functional form for the spillovers. The particular functional form depends on the context of the application. However, there are a few practical considerations. First, the researcher should argue whether spillovers on control units and on treated units exist. Second, they should argue what the maximum distance these spillovers can occur. Last, they should consider if spillovers are additive in the number of nearby units that are treated.

In section 3, I use Monte-Carlo simulations to emphasize when misspecificied spillovers still remove most of the bias and when they fail to do so. The first result of these set of simulations is that a set of indicators for being within distance bins (coloquially called 'Donuts') from treated units will absorb all the bias so long as the maximum distance is larger than the maximum distance of spillovers.

Second, if the researcher wants to estimate spillovers, 'Donuts' are the most robust to misspecification of spillovers. However, in estimation of spillovers it is important to get right the question of whether spillovers are additive or non-additive. If the spillovers are non-additive, then 'Donuts' model spillovers well. If they are additive, then a set of variables for the number of treated units between a distance bin from the control unit captures spillovers well.

1.1. Contribution

The literature considering spillovers primarily has focused on "within-group" spillovers where units are in distinct groups and outcomes depend on the treatment status within the observation's group only.[5] However, spatial settings such as counties do not feature distinct groups which allow for the researcher to separate areas into distinct non-overlapping groups (a county is connected to all other counties by some chain of neighbors). My paper considers a more general form of interference. Identification in this framework requires the researcher to either identify control units that are free from spillovers (far-away from treatment) or to parameterize how control units are affected by treatment.

Vazquez-Bare (2019) models within-group spillovers such that their exposure mapping is only a function of unit i's group's treatment vector. He develops a potential outcomes framework which assumes that potential outcomes are a function of both own treatment-status D_i and a function of the vector of treatment assignments $h(\vec{D}, i)$. The function $h(\vec{D}, i)$ is referred to as an exposure mapping and adding this to the potential outcomes framework allows me to explicitly model SUTVA violations. Using this framework, Vazquez-Bare (2019) shows that a difference in means comparison between treated and control units is a biased estimated for the treatment effect.

My paper uses this potential outcome framework in the case where treatment spillover is not limited to within-group spillovers and considers estimation of treatment effects by difference-in-differences.

Sävje, Aronow, and Hudgens (2019) consider a similar potential outcome framework as Vazquez-Bare (2019), but they instead include spillovers as a part of their definition

---

[5] For example, researchers consider peer effects within schools. In this case, each school constitutes a group and it is assumed that outcomes do not depend on other schools' treatment status, i.e. SUTVA is satisfied between-groups. Estimation of spillover effects in this case is done in a 'partial intereference' framework that compares "control" units in a partially treated group with control units in completely untreated groups. Angelucci and Di Maro (2016) provides a summary for estimation of spillover effects in within-group settings. 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.

of "treatment effect". That is, a treated unit's treatment effect is defined as the direct effect *and* the general equilibrium spillover effects. Then, the authors describe how the average treatment effect (ATE) estimated by difference-in-differences averages over both individual heterogeneity and, central to their contribution, over heterogeneity in the exposure mapping, $h(\vec{D}, i)$. My paper develops a strategy to allow researchers to seperately identify treatment effects and spillovers.[6]

There are a few papers that discuss estimation of treatment effects with spatial spillovers using the difference-in-differences framework. Clarke (2017) finds an explicit form for bias when estimating a difference-in-differences model. However, their model only allows spillovers onto control units and assumes spillovers are constant regardless on the number of nearby treated units. Berg and Streitz (2019) and Verbitsky-Savitz and Raudenbush (2012) identifies biases in estimation of the direct effect for a specific potential outcome function where $h(\vec{D}, i)$ is the proportion of treated contiguous counties. Finally, Delgado and Florax (2015) consider the case where the exposure mapping is defined by a spatial weights matrix $W\vec{D}$ that enters linearly into the potential outcomes. This does not allow for the coefficient to change for control and treated units.[7]

I contribute to the econometric literature in two ways. First, my paper derives an explicit form for this bias in terms of general potential outcomes which allows my results to capture many different forms of spillovers. For the above papers, if I assume the particular functional forms for potential outcomes, I arrive at the same bias equation as theirs if they have one derived explicitly. Second, my paper also advances the literature by considering estimation of direct effects and spillover effects in event-study framework which allows for the common occurence of staggered treatment adoption.

Last, I contribute to a literature in urban economics on local place-based policies. Using structural urban models, Kline and Moretti (2014b) show that the 'local' or partial equilibrium effects of a place-based policy are offset in part by indirect or general equi-

---

[6] These estimates can be combined afterwards to estimate their definition of average treatment effect. The further advantage is that 'net' treatment effects can be estimated at different levels of exposure which is more relevant for future policy decisions.

[7] Although, one could change the weighting matrix to make ad-hoc assumption of the signs and relative magnitudes of spillovers onto treated and control units, but this allows for less data-driven estimation.

librium effects. I contribute to estimating aggregate effects of place-based policies by showing that typical estimation strategies can underestimate the local effect by ignoring spillovers. As an example, I revisit the analysis of the Tennessee Valley Authority by Kline and Moretti (2014a). The Tennessee Valley Authority was a large push to bring electrification to the Tennessee Valley with the goal of improving the manufacturing sector there. The scale of the program was large and the pro-manufacturing benefits likely spread further than the Authority's boundary due to the electrification infrastucture and agglomeration economies. I show that estimation by difference-in-differences fails to account for these spillovers and for different outcome variable it underestimates and overestimates the local effect, depending on the sign of the spillovers.

The rest of the paper is structured as follows. Section 2 presents the potential outcomes framework, defines the estimand of interest, and shows the resulting bias from estimating a classical difference-in-differences model. Section 3 presents Monte Carlo simulations to illllustrate the bias result and to evaluate currently used solutions in the literature. Section 4 presents an application for evaluating the Tennessee Valley Authority program. Section 5 discusses briefly how to incorporate spillovers into event-study estimation following the insights from Callaway and Sant'Anna (2020).

2 — Potential Outcomes Framework

Following the canonical difference-in-difference framework, there is a time t_0 where treatment turns on and remains on afterwards. This framework is extended to staggered treatment adoption in section 5. Potential outcomes for unit i at time t are a function of own treatment-status D_i and, departing from the standard potential outcomes framework, of a function of the entire vector of treatment assignments $h(\vec{D}, i)$ where $\vec{D} \in \{0, 1\}^n$ denotes the n-dimensional vector of all unit treatments. The potential outcomes are denoted $Y_{i,t}(D_i, h(\vec{D}, i))$. The function $h(\vec{D}, i)$ is referred to as an 'exposure mapping' and is a non-negative scalar- or vector- valued function. The exposure mapping measures the intensity at which unit i is affected by spatial spillovers. When unit i is sufficiently 'far' away, it has no exposure to spatial spillovers and $h(\vec{D}, i) = \vec{0}$. The exposure mapping

formalizes the SUTVA violation in that potential outcomes are affected by other unit's treatment assignment as summarized by $h(\vec{D}, i)$. To help better understand the exposure mapping function, I give three examples of $h(\vec{D}, i)$ that are commonly used in the literature.

Example 1. *First, $h(\vec{D}, i)$ can be a 0/1 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 geographic distance measure which tells the distance unit i is from unit j. In this case*

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

This exposure mapping is useful when it is assumed that spillovers do not decay over distance until $\bar{d}$ and the number of neighboring units treated do not impact the intensity of spillovers. 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. Access to a neighboring town library is binary in nature, so it does not matter whether you can access 1 or more nearby libraries, so the binary exposure mapping is a good approximation to spillovers in this context.

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

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

This exposure mapping is no longer binary, so the intensity of spillovers depend on the number of nearby units treated. In the context of large store openings, this exposure mapping captures the additive nature of agglomeration economies. As more nearby counties receive new stores, the agglomeration forces increase (Basker, 2005).

Example 3. *Last, is the spatial decay function where exposure decreases with distance. In this case, spillover intensity is the sum across all treated observations' decay term, i.e.*

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

---

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

This exposure mapping allows for the intensity of spillovers to depend on distance to treatment and also is additive in the number of nearby units treated.[9] *The speed of spillover decay over distance depends on the parameter α that can be calibrated by the researcher or estimated using non-linear least squares. In the literature on R&D investment, Keller (2002) uses a modified version of this exposure mapping where D_j is country j's R&D expenditure.*

After chosing an exposure mapping that makes sense in the economic context, a researcher must also specify the functional form of the potential outcomes. Typically, the spillover function enters into the regression linearly and potentially the coefficient is allowed to differ by treatment status.

2.1. Spatial Spillovers

With the potential outcomes defined, I now formalize what is meant by 'spatial spillovers'. I define 'spillover onto control units' as:

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

The spillover measures the difference in non-treated potential outcomes between being exposed at intensity $h(\vec{D}, i)$ and not being exposed. Then, the average spillover effect onto control units averages over potential heterogeneity in the effect size of spillovers and over heterogeneity in exposure intensity $h(\vec{D}, i)$:

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

To emphasize, the average spillover effect onto conrol units averages over each control unit's exposure mapping. For example, assume the potential outcomes is additively linear in D_i and $h(\vec{D}, i)$ and that the coefficient $\beta_{\text{spill,control}}$ measures the effect of $h(\vec{D}, i)$ on outcome Y among control units. Then an individual control unit's spillover effect is $\beta_{\text{spill,control}} h(\vec{D}, i)$. The average spillover effect onto control unit would therefore be $\tau_{\text{spill,control}} = \beta_{\text{spill,control}} * \mathbb{E}\left[h(\vec{D}, i)\right]$, i.e. the average over all control units exposure mapping.

---

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

Similarly, we define the average spillover effect onto also treated units as:

$$\tau_{\text{spill,treated}} \equiv \mathbb{E}\left[Y_{it}(1, h(\vec{D}, i)) - Y_{it}(1, 0) \mid D_i = 1\right].$$

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

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

which measures the effect of being treated in the absence of exposure to spillovers. This effect does not include the general equilibrium effects of multiple treated areas near one-another.

My definition of the direct effect of treatment differs from Sävje, Aronow, and Hudgens (2019) where they define the average treatment effect as

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

where the expectation is over individuals and their exposures. I prefer the former because it allows for seperate identification of the direct effect of treatment and the spillover effects themselves. For example, a county deciding whether to implement a policy will want to consider the treatment effect plus the spillover at their current level of exposure. The estimand proposed by Sävje, Aronow, and Hudgens (2019) returns the average of spillovers across levels of exposure and can not be used to predict spillovers at different levels.[10]

2.2. Bias in Difference-in-Differences Estimation

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

$$y_{it} = \tau D_{it} + \mu_i + \mu_t + \epsilon_{it}. \tag{4}$$

---

[10] Estimation of the direct effect and spillover effect can be combined to estimate the ATE as defined by Sävje, Aronow, and Hudgens (2019).

The estimator $\hat{\tau}$ is a biased estimate for τ_{direct} in the presence of spillovers. To show this, I first present the equivalent to the parallel counterfactual trends assumption in the context of the new potential outcome framework.

Assumption 1 (Parallel Counterfactual Trends)**.**

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

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.

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

Theorem 1 (Bias from Difference-in-Differences Estimation)**.**
If Assumption 1 holds, the expectation of the estimate $\hat{\tau}$ from (4) is

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

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

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

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

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

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

2.3. Bounding the Bias

The previous section finds an explicit form of the bias when estimating (4). The explicit form of bias can be used by researchers to provide bounds on the bias in a manner simmilar to Rambachan and Roth (2020). To provide bounds on the bias, researchers must identify the set $\Delta = [\underline{\Delta}, \overline{\Delta}]$ such that it contains the bias

$$\tau_{\text{spill,treated}} - \tau_{\text{spill, control}} \in \Delta$$

Choosing this set requires considerations of the economic context. It is guided by choices of parameterization of the exposure mapping and potential outcomes. To exemplify this bias, I will illustrate how a researcher would bound the bias in a simple example. I will consider the example of the effect of new libraries on school outcomes. Following the motivation above, I assume that spillovers occur for any county that is within $\bar{d}$ miles of a treated county, i.e.

$$h(\vec{D}, i) \equiv \text{Near}_{it} = \max_{i \neq j} D_j * 1[d(i, j) < \bar{d}]. \tag{5}$$

This specification is a simple one in that the spillover is not additive in the number of nearby treated counties and does not decay over distance before reaching $\bar{d}$. Since treated counties have access to their own library, I will assume that no spillovers on to treated units exist. The potential outcome is assumed to be additive in both treatment and in the exposure mapping:

$$y_{it} = \beta_{\text{direct}} D_{it} + \beta_{\text{spill,control}} (1 - D_{it}) \text{Near}_{it} + \mu_t + \mu_i + \varepsilon_{it}, \tag{6}$$

where μ_t and μ_i are unit and time fixed effects respectively. In this case, the spillover on an individual control unit is given by $\tau_{\text{spill,control}}$.

If a researcher estimates (4), then the bias will be

$$-\tau_{\text{spill,control}} = -\beta_{\text{spill,control}} \frac{\sum_{i:D_{it}=0} \text{Near}_{it}}{N_C},$$

where N_C are the number of control units respectively. Simply put, the bias will be negative one times the proportion of control units that receive spillover effects time the size of the effect on control units. Then the bound can be created by bounding $\beta_{\text{spill,control}}$ and the proportion of control units that receive spillovers.

11

With a chosen Δ, the partially identified set is given by $[\hat{\tau} + \underline{\Delta}, \hat{\tau} + \overline{\Delta}]$. Inference on this set can be done using Fixed-Length Confidence Intervals as in Rambachan and Roth (2020). The confidence interval can be formed using standard normal critical values and spatially adjusted standard errors following Conley (1999) since we are assuming a specific spatial relationship between outcomes:

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

2.4. Improving Estimation by Parameterizing Spillovers

The previous section shows how to estimate the magnitude of potential bias in the estimation of (4). However, parameterization of the potential outcomes and exposure mapping allows for unbiased estimates of τ_{direct} by controlling for the exposure mapping directly. This is recommended for two reasons.

First, controlling for exposure mapping allows for unbiased estimates of the direct effect of treatment if the potential outcomes are correctly specified. In the case where potential outcomes are 'somewhat' incorrectly specified, this can remove some of the bias in estimation. Simulations below highlight that since spillovers are likely to decay over distance, controlling for immediate neighbors removes a large portion of the average spillover effects even under misspecification.

Second, the spillover effects themselves are potentially relevant for policy makers. If treatment for a given area is positive and there are positive benefits to neighboring regions, the estimated benefits will double-understate the true effect of treatment. First, it will not include the neighboring units' benefit in the calculation. Second, since the treatment effect will subtract the neighboring units' benefit, the neighboring units' benefit will be *double undercounted*. Similarly, a policy maker trying to determine the net benefits for a treated location must consider the direct effect and the spillover effect onto treated units. This is because the size of benefits created by the treatment depend on the number of treated units nearby.

There are a few practical considerations researchers should make when parametrizing the spillovers and the answers to the questions will depend ultimately on the economic context being studied. First, researchers should consider whether control units and/or

12

treated units experience effects from the treatment status of other units. For example, in the context of library construction, treated units likely do not experience spillovers while close control units do. In the context of factory openings, both control and treated units likely experience spillovers.

Second, a researcher must decide how far the spillovers can reasonably be experience and whether the effect decays over distance. Things such as driving distance or access via public transportation can be helpful in deciding which units are 'close enough' to experience spillovers. There are two commonly used techniques to model treatment effect decay.[11] First, as seen above is to use an exponential decay function with regards to distance. The other option is to use a set of 'cocentric donut' indicators which are easy to interpret. An example of this is shown in the empirical application in Section 4.

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

3 — Monte Carlo Simulations

3.1.   *Spillover on Controls Only*

I now turn to a series of Monte Carlo Simulations to highlight the importance for controlling for spillovers in simulations. I use the set of counties in the contiguous United States and the data generating process used is defined by the exposure mapping given in (5) with cutoff distance $\bar{d} = 40$ miles and the data generating process given in (6). A unit of observation is a US county and the periods are $t \in \{1, \ldots, 20\}$ with treatment turning on

---

[11] If the spillovers is due to crossing borders into the treated area, then a reasonable decision about whether the effect decays is the frequency in which people use the good. For example, libraries whose benefit is only experienced from frequent usage would suggest that the effect decays over distance. On the other hand, abortion clinics opening are used infrequently and therefore the effect likely doesn't decay over distance within a certain cutoff.

after period $t = 10$. The unit and time fixed effects are generated by $\mu_t \sim N(0.2t, 0.1^2)$ and $\mu_i \sim N(6, 2^2)$ respectively, and the error term is $\varepsilon \sim N(0, 2^2)$. Last, the size of treatment and spillovers are as follows: $\beta_{\text{direct}} = 2$, $\beta_{\text{spill, control}} = 1$ and $\beta_{\text{spill, treat}} = 0$.

For the first simulation, I restrict spillovers only to occur among control units, i.e. $\beta_{\text{spill, treat}} = 0$, and estimate the two-way fixed effects model, (4). This data generating process where spillovers only occur onto control units matches what has been typically assumed in the literature. I assign treatment among counties randomly with various unconditional probabilities between 3 percent and 50 percent. The data-generating process is therefore

$$y_{it} = \mu_t + \mu_i + \beta_{\text{direct}} D_{it} + \beta_{\text{spill, control}} (1 - D_{it}) \text{Near}_{it} + \varepsilon_{it} \tag{7}$$

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

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

There is a trade-off between the bias and the variance of the estimator when using this methodolgy. As the treatment probability increases, the number of control units

Figure 1 — Bias of $\hat{\tau}$ with Spillovers Effects *only* on Control Units
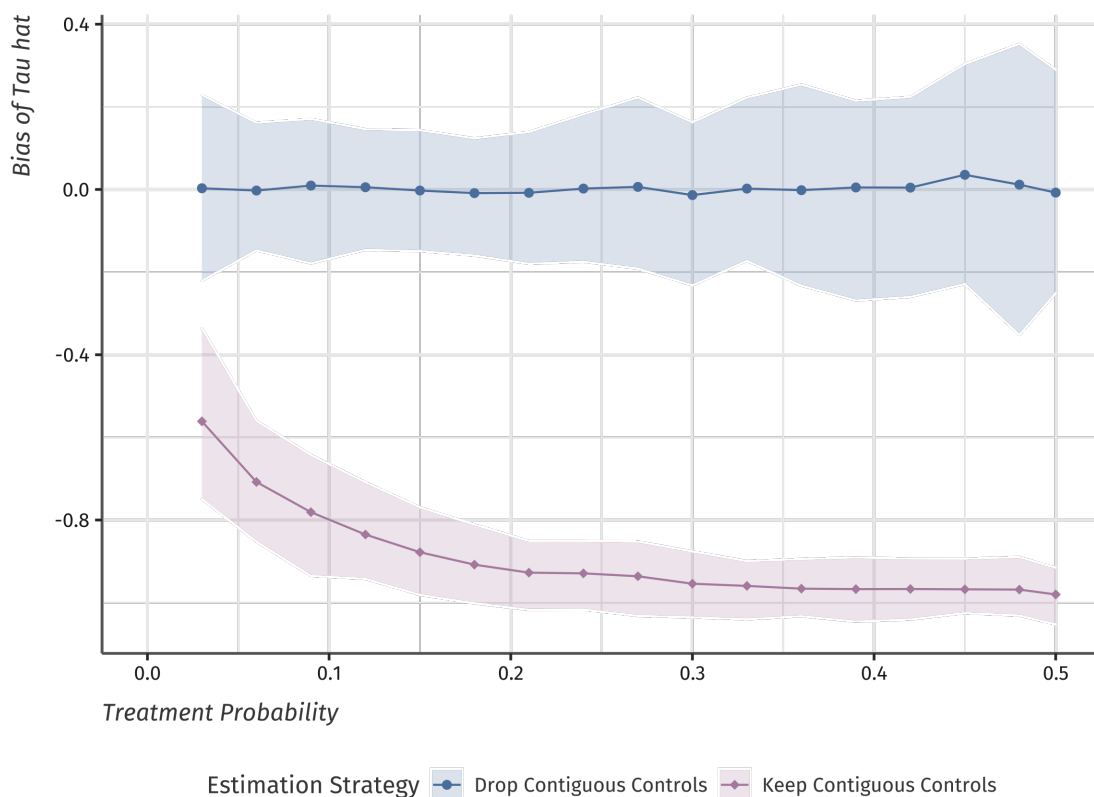
Notes: This figure plots the bias of $\hat{\tau}$ found from estimating (4) for data generated as described in the text and Equation 7. Each point corresponds to the average bias for the given treatment probability and the band is the 95 percent empirical confidence interval over 1000 simulations. The line with diamond markers estimates with all control units. The line with circle markers removes control units that share a border with a treated county.

removed increases as well. This naturally yields a more variable estimator as seen in the wider 95 percent empirical confidence intervals in Figure 1. This trade-off can be avoided altogether by parameterizing the spillovers and including them in estimation directly.

3.2. Spillovers onto Control and Treated Units

For the following simulation, I will add in spillovers onto also treated units with $\beta_{\text{spill, treat}} = -0.5$. The DGP is as follows

$$y_{it} = \mu_t + \mu_i + 2D_{it} + (1 - D_{it})\text{Near}_{it} - 0.5D_{it}\text{Near}_{it} + \varepsilon_{it} \tag{8}$$

The proportion of treated units affected by the spillover depends on the spatial autocorrelation of the treatment assignment. In empirical applications, it is often the case that treatment assignment is concentrated in specific areas which yields higher proportion of treated units receiving spillover effects.

Figure 2 — Example of Kriging Field
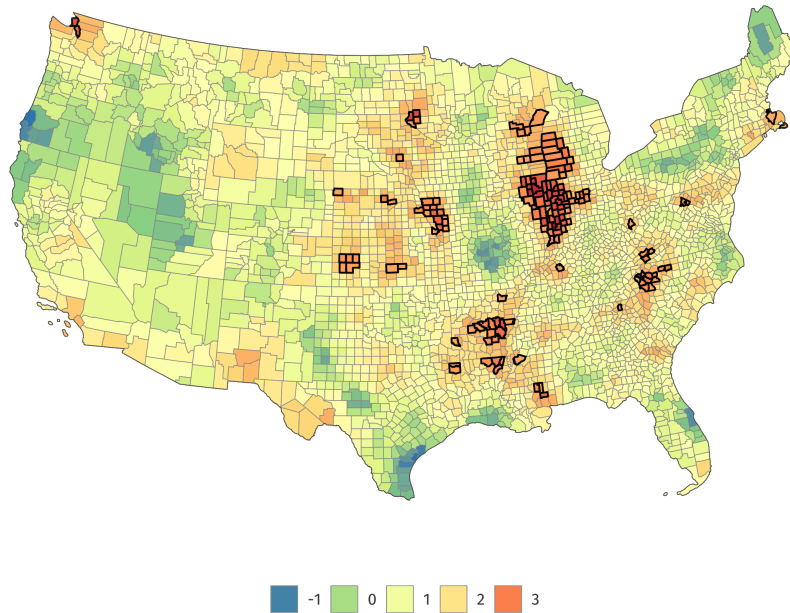
Notes: This figure plots an example of a kriging simulation using the 'gstat' package in R using an exponential variogram. The outlined counties are in the top ten percent of values for this simulation.

In order to model spatial autocorrelation of treatment, I turn to a method from geosciences called 'kriging'. Kriging generates a Gaussian field for a large grid of points across the entire United States where the spatial autocorrelation of the points is described by a set

Figure 3 — Bias of $\hat{\tau}$ at Different Levels of Spatial Autocorrelation
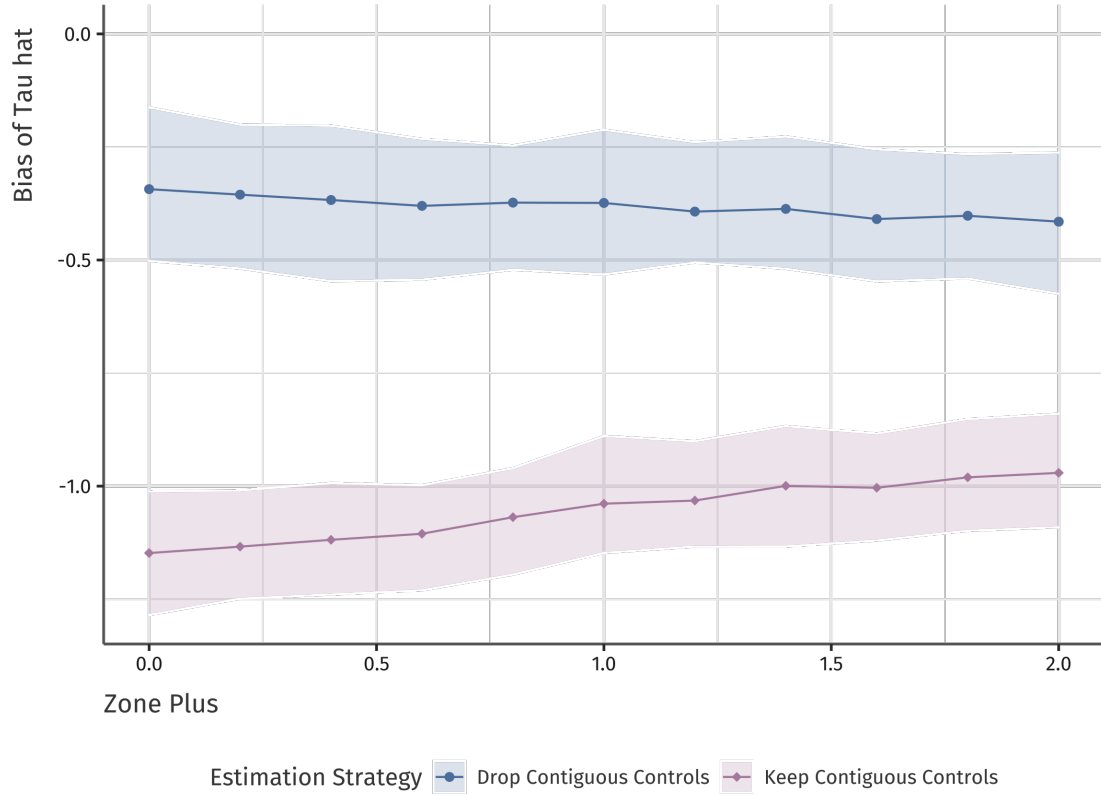
Notes: This figure plots the bias of $\hat{\tau}$ found from estimating (4) for data generated as described in the text and Equation 8. Treatment probability is assigned according to the method described by the text and (9). Each point corresponds to the average bias for the given Zone Plus and the band is the 95 percent empirical confidence interval over 1000 simulations. The line with diamond markers estimates with all control units. The line with circle markers removes control units that share a border with a treated county.

of parameters. One such field is depicted in Figure 2. For a given field, I find the counties in the top ten percent of values and give these counties a value of $\text{Zone}_i = 1$. Then I assign treatment with a probability that differs based on Zone_i. To avoid changes in bias due to a difference in the number of treated units, the unconditional probability is equal to 5 percent in all simulations. To increase the spatial autocorrelation of treatment, I increase the probability of treatment for units with $\text{Zone}_i = 1$ and hence lower the probability for

units with $\text{Zone}_i = 0.$ [12]

The results for the bias of $\hat{\tau}$ are in Figure 3. The line with diamond markers is run on the full sample. As the level of spatial autocorrelation increases, the bias surprisingly decreases. The reason for this is that as treatment becomes more concentrated in the 'Zones', the number of control units receiving spillovers decreases while the number of treated units receiving spillovers increases. In this particular data-generating process, the effect of fewer control units on the bias is larger than the bias from more treated units.

The line with circle markers in Figure 3 repeats the exercise of dropping control units that share a border with treated units. In the case where spillovers occur on treated units, all bias is not removed when dropping control units. That is because we only remove $\tau_{\text{spill,control}}$ but $\tau_{\text{spill,treated}}$ remains. To remove the second source of bias, it is necessary to acount directly for spillovers in the estimation strategy.

3.3. Misspecification of Spillovers

Now we turn to simulations where we include potentially incorrectly specified exposure mappings and see their performance in a set of different data generating processes. In the general form, I generate data with the same fixed effects and error term as before, but for many different exposure mappings $h(\vec{D}, i)$

$$y_{it} = \mu_t + \mu_i + 2D_{it} + \beta_{\text{spill, control}}(1 - D_{it})h(\vec{D}, i) + \varepsilon_{it} \tag{10}$$

In order to keep the magnitude of bias constant across specifications, I normalize the average spillover magnitude to be equal across specifications.[13] For simplicity, I also remove spillovers onto treated units, but the results of which specifications perform best

---

[12] To change the spatial autocorrelation of treatment, I use a variable 'Zone Plus' to adjust the relative probability of receiving treatment:

$$P(D_i \mid \text{Zone}_i) = (.05 + \text{Zone Plus} * \text{Zone}_i) \times \underbrace{\frac{.05}{.05 * .9 + (.05 + \text{Zone Plus}) * .1}}_{\text{Normalize } P(D_i) = 0.05} \tag{9}$$

The second term normalizes probabilities so the unconditional probability stays at five percent. In our simulations, Zone Plus ranges from 0 which is the case where $P(D_i|\text{Zone}_i) = P(D_i) = 0.05$ to 2 which has $P(D_i \mid \text{Zone}_i) = 0.01 + 0.41 * \text{Zone}_i$.

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

in the simulation are qualitatively the same. For each true data-generating process, I estimate (10) with the correct spillover and misspecified alternatives $\tilde{h}(\vec{D}, i)$. The goal of this exercise is to see which specifications perform well under broad classes of spillovers.

The set of spillover specifications include 'Within x mi.' which equals 1 if the counties' center of population is within x miles to a treated county's center of population, and 'Within x mi. (Additive)' is the number of treated county's within x 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 Equation (3) with $\alpha = 0.02$.

The last set of spillover specifications is commonly used in the literature as a semi-parametric estimator and are referred to as 'Donuts' or 'Rings'. These consist of a set of indicators for falling within distance bins from the nearest treated unit (e.g. indicators for being between 0-10, 10-20, and 20-30 miles to closest treated unit). This approach is worth discussing in more detail because it will turn out to be the most consistent approach. These donuts are able to trace out all of the non-additive spillovers so long as the distance bins are numerous enough. For example, the donuts indicators add up to all the 'Within x mi.' specifications if there are enough rings. The donuts can also trace out the decaying spillovers with the coefficients on each indicator decreasing in an exponential way. Donuts that go further than the maximum spillover distance will be estimated near zero, so adding too many rings is not a large concern.[14] The last specification is an additive version of 'Donuts' which is the number of treated units within each distance bin.

The results of the simulations are produced in Table 1. For each entry in the table, the column label corresponds to the true exposure mapping used in the data-generating process and the row label corresponds to the exposure mapping, $\tilde{h}(\vec{D}, i)$, used in estimation. The corresponding cell gives the mean bias from estimating Equation (10) with the $\tilde{h}(\vec{D}, i)$.

The first thing to note is that all specifications remove a large portion of the bias relative to traditional two-way fixed effect estimate and in particular, correctly specifying the data-generating process removes almost all the bias. The second result is that both the 'Donuts' and the 'Within' specification do the best amongst all the misspecified spillovers

---

[14] To quote Homer Simpson, *"Donuts. Is there anything they can't do?"*

Table 1 — Bias from Misspecification of Spillovers

Specification	Within 40mi	Within 80mi	Within 40mi. (Additive)	Within 80mi. (Additive)	Decay	Decay (Additive)
				Data-Generating Process		
TWFE (No Spillovers)	0.258	0.258	0.258	0.258	0.258	0.258
Within 40mi.	−0.005	0.213	−0.005	0.176	0.159	0.143
Within 80mi.	−0.009	−0.009	−0.009	−0.009	−0.009	−0.009
Within 100mi.	−0.006	−0.006	−0.006	−0.006	−0.006	−0.006
Within 40mi. (Additive)	0.043	0.221	−0.006	0.177	0.174	0.143
Within 80mi. (Additive)	0.034	0.134	−0.012	−0.009	0.099	−0.010
Decay	−0.159	0.070	−0.174	0.014	−0.009	−0.033
Decay (Additive)	−0.023	0.148	−0.084	0.019	0.088	−0.008
Donuts (0-20, 20-30, 30-40)	−0.005	0.213	−0.005	0.176	0.159	0.143
Donuts (0-20, 20-30, 30-40, 40-60, 60-80)	−0.009	−0.009	−0.009	−0.009	−0.009	−0.009
Donuts (0-20, 20-30, 30-40, 40-60, 60-80) (Additive)	0.036	0.134	−0.008	−0.008	0.100	−0.009

Notes. Each cell corresponds to the mean bias from 1000 simulations of the data generating process specified by the column header and under the specification given by Equation (10) by the row label.

in terms of removing bias, so long as they are wide enough to capture all the spillovers. This is due to the fact that indicator variables will correctly identify the average spillovers onto control units if they cover all the affected control units. If researchers are primarily interested in identifying the direct effect, then either 'Donuts' or 'Within' specifications with a large cutoff distance are preferred as they remove almost all bias in all settings.

However, we will now turn to seeing how these misspecifications are able to capture the spillovers themselves, and we will see that 'Donuts' are superior to 'Within' indicators in their ability to trace out spillovers, so they should be the default used by researchers unless there is theoretical motivation to use any specific parametrization for exposure mapping. For the next table, for each control county, I will predict their spillover value 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 (to make numbers comparable across data-generating processes),

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

Table 2 — Percent of Spillovers Predicted by Specification

| Specification | Data-Generating Process | | | | | |
	Within 40mi	Within 80mi	Within 40mi. (Additive)	Within 80mi. (Additive)	Decay	Decay (Additive)
TWFE (No Spillovers)	0.0%	0.0%	0.0%	0.0%	0.0%	0.0%
Within 40mi.	99.4%	25.9%	85.6%	38.8%	59.5%	56.1%
Within 80mi.	39.8%	96.2%	34.3%	71.7%	85.6%	68.0%
Within 100mi.	33.2%	80.3%	28.6%	59.8%	71.5%	56.7%
Within 40mi. (Additive)	85.3%	21.2%	99.5%	40.6%	52.0%	60.7%
Within 80mi. (Additive)	45.8%	61.8%	47.2%	98.4%	71.0%	93.6%
Decay	60.1%	82.5%	52.7%	75.8%	97.5%	82.2%
Decay (Additive)	60.7%	56.9%	63.8%	93.5%	79.0%	98.7%
Donuts (0-20, 20-30, 30-40)	98.4%	23.7%	85.9%	37.5%	58.9%	56.2%
Donuts (0-20, 20-30, 30-40, 40-60, 60-80)	96.6%	91.7%	84.2%	72.7%	91.9%	78.4%
Donuts (0-20, 20-30, 30-40, 40-60, 60-80) (Additive)	83.5%	57.4%	97.6%	95.0%	73.5%	94.9%

Notes. Each cell corresponds to the mean square prediction error of the spillover effect for each control unit normalized by the total variance of spillover effects from 1000 simulations of the data generating process specified by the column header and under the specification given by Equation (10) by the row label.

This will produce a percentage of spillovers that are explained by the specification, so numbers closer to 1 represent better modelling of the spillovers.

Results are presented in Table 2. For each specification, correctly specifying the spillovers produces the best estimates for spillovers as expected. Among misspecified exposure mapping, 'Donuts' that are large enough to caputre all the spillovers perform better than all other specifications in all non-additive data-generating processes. When the spillovers are additive, the predicted spillovers from the non-additive donut specifications do not as accurately measure these spillovers as they did before.

For additive data-generating processes, 'Within (Additive)' and 'Decay (Additive)' don't do a great job at predicting spillovers for the other data-generating process. On the other hand, the additive version of donuts sucessfully predicts spillovers in both additive data-generating processes. Therefore, the most important aspect of correctly estimating spillovers is to consider whether they are additive in the number of treated units or not. Researchers should use a version of donuts either way.

4 — Application in Place-Based Policy Analysis

To illustrate the importance of accounting for spatial spillovers in the estimation of treatment effects, I revisit the analysis of the Tennessee Valley Authority (TVA) in Kline and Moretti (2014a). The TVA program was a large-scale federal investment started in 1934 that focused on modernizing the Tennessee Valley economy. The program focused on electrification through dams and transportation canals in order to improve the manufacturing economy. By the end of WWII, the TVA became the largest single power supplier in the country.[15] With over $20 Billion (in 2000 dollars) spent and hundreds of dollars transferred per person in the Authority, the impacts are very likely to cross far past the authority's borders.

Kline and Moretti (2014a) focus on a set of outcome variables, but I will focus on three in the analysis. The first two are (the log of) agricultural and manufacturing employment and the last is (the log of) median household income. Since the TVA focused on manufacturing employment, the authors predict that employment will grow in manufacturing and shrink in agricultural. Since manufacturing jobs pay more on average, the median household income will therefore grow.

In Kline and Moretti (2014a), they discuss the nature of spillovers that can occur. For agriculture employment, the authors claim that improved wages in the Authority will draw agriculture workers out of nearby counties. Hence they predict a negative spillover. For manufacturing, the sign is ambiguous. There could be positive spillovers if electrification brought cheap power to the neighboring areas. Also, agglomeration effects increase manufacturing employment in nearby counties.[16] 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.

The analysis in Kline and Moretti (2014a) begins by comparing long-differences in

---

[15] More details on the program are found in Kline and Moretti (2014a).

[16] Duranton and Puga (2003) give theoretical motivations for different sources of agglomeration. In this context, agglomerations due to 'sourcing' is most likely. When the Authority has a large increase in manufacturing, they develop resource supply chains that in turn make inputs cheaper for neighboring counties.

county-level outcomes from 1940 to 2000 between treated counties in the Authority and control counties outside. The primary difference-in-differences specification is

$$y_{c,t} = \mu_c + \mu_t + \tau \text{TVA}_c \times \text{Post}_t + X_{c,1940} \times \text{Post}_t \gamma + \epsilon_{c,t}, \tag{11}$$

where c denotes county, TVA_c is an indicator variable for being in the Authority, Post_t equals one when $t = 2010$ and equals zero when $t = 1940$, and y is a set of outcome variables.[17] Pre-treatment control variables, $X_{c,1940}$, are interacted with Post_t to allow for places to be on different long-term trends.[18]

Kline and Moretti (2014a) estimate (11) to identify the 'local effect' of the TVA – what I am calling the 'direct effect'. However, their point estimates compare, in part, changes in outcomes in TVA counties with changes in outcomes for neighboring counties that likely were impacted by the large-scale program. Their identification strategy depends *explicitly* on a parallel counterfactual trends assumption after controlling for covariates and *implicitly* on the SUTVA assumption.

As described above and estimated below, the program likely induced manufacturing and agricultural employment effects in nearby counties. By comparing counties inside the Authroity to counties on the other side of the border, the authors likely underestimate the negative effect on agricultural employment while the bias in the manufacturing effect and median family income is theoretically ambiguous. The authors do recognize the problem of these comparisons and remove counties that share borders with the authority's, but due to the scale of the program, the spillovers are likely to extend further than this and appear to do so in the empirical analysis below.

The authors use an Oaxaca-Blinder estimator on the first differences and the results of Kline (2011) show that this estimator is a weighted difference-in-differences. The

---

[17] Outcome variables are all in logs and include population, average manufacturing wage, agricultural employment, manufacturing employment, value of farm production, median family income, and median housing value. The text writes the equivalent first-differences regression:

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

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

Figure 4 — TVA Effective Sample and Spillover Variables
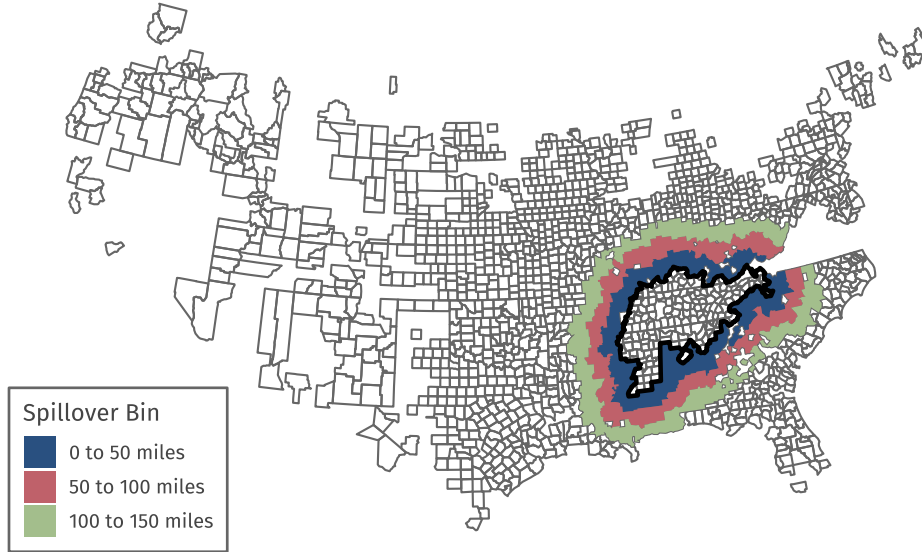
Notes: The above figure plots all the counties used in the estimation. Counties that fall within the distance intervals $\{(0, 50], (50, 100], (100, 150]\}$ measured in miles are colored by their respective bin.

weights can be seen in Figure III of their text, but generally counties close to the TVA are upweighted. However, in Table 3 below, I will show their estimator does not differ much from the standard difference-in-differences results since the weights are not that different from uniform weights.

In order to more likely satisfy the paralell trends assumption, the run a logistic regression to predict being in the TVA based on their set of control variables $X_{c,pre}$ and keep only observations in the top 75% of predicted probability. The counties used in the sample are presented in Figure 4.

I extend their analysis to include spatial spillovers in the difference-in-differences specification. To parametrize the exposure mapping, I include a set of indicator variables for being in distance bins away from the Authority. I do this for two reasons. Specifically, I use the following intervals Dist = $\{(0, 50], (50, 100], (100, 150]\}$ measured in miles and define Between(d) as an indicator for being within the interval $d \in$ Dist away from the

Authority. Figure 4 displays the three spillover variables by filling in each distance bin in a different color. As discussed above, I suspect the spillover effects to change with distance to the Authority, hence I allow the effect to be seperately identified in bins. I keep the number of bins small to improve precision of spillover estimates.

The specification with spillovers is given as follows:

$$y_{i,2000} - y_{i,1940} = \alpha + \text{TVA}_i \tau + \sum_{d \in \text{Dist}} \text{Between}(d)\delta_d + X_{i,1940}\beta + (\varepsilon_{i,2000} - \varepsilon_{i,1940}) \qquad (12)$$

The coefficients δ_d measure the spillover effect onto control units at different distances from the Authority. From the above simulations, then $\hat{\tau}$ will be an unbiased estimate for what the authors call the 'local' effect so long as spillovers don't occur past 150 miles from the TVA.

The results of the analysis are presented in Table 3. The first column lists which dependent variable (measured in logs) was used in the row. The following columns contain point estimates for τ and δ_s in different specifications and can be interpreted as decadel growth rates in outcomes. The column labeled 'Kline & Moretti (2011)' use an Oaxaca-Blinder estimator while the column labeled difference-in-differences uses an ordinary least squares estimator for (11). The two columns are very close in magnitude to each other. They find a decline in agricultural employment of about 5.6% per decade, an increase in manufacturing employment of about 6.1% per decade, and an increase in median family income of about 2.2% per decade.

Turning to the last four columns that estimate (12), column (3) contains a point estimate for τ and and columns (4)-(6) contain point estimates of the spillover effects δ_d. For agricultural employment, our 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 3.1%. Between 50 and 100 miles the point estimate is -1.1% and between 100 and 150 miles, the point estimate is -2.5%. This is likely due to the fact that higher paying manufacturing jobs within the Authority drew migrants from nearby counties. Because the spillovers onto the control counties is negative, the original difference-in-differences estimator was positively biased. The new point estimate indicates a decline of agricultural employment of about 6.8% per decade compared to 5.1% in the standard difference-in-difference specification.

25

Table 3 — Effects of Tennessee Valley Authority on Decadel Growth

	Kline and Moretti (2014a)	Diff-in-Diff	Diff-in-Diff with Spillovers			
				TVA between	TVA between	TVA between
	TVA	TVA	TVA	0-50 mi.	50-100 mi.	100-150 mi.
Dependent Var.	(1)	(2)	(3)	(4)	(5)	(6)
Population	0.0052	0.0053	−0.0040	−0.0212	−0.0077	−0.0053
	(0.0099)	(0.0090)	(0.0089)	(0.0179)	(0.0125)	(0.0123)
Average manufacturing wage	0.0064	0.0070**	0.0075*	0.0022	−0.0037	0.0023
	(0.0048)	(0.0030)	(0.0045)	(0.0047)	(0.0040)	(0.0047)
Agricultural employment	−0.0561***	−0.0514***	−0.0678***	−0.0310**	−0.0112	−0.0252***
	(0.0100)	(0.0087)	(0.0102)	(0.0123)	(0.0094)	(0.0084)
Manufacturing employment	0.0606***	0.0560***	0.0461**	−0.0104	−0.0128	−0.0248*
	(0.0170)	(0.0187)	(0.0210)	(0.0205)	(0.0257)	(0.0147)
Value of farm production	0.0024	−0.0060	−0.0184	−0.0188	−0.0101	−0.0266
	(0.0245)	(0.0269)	(0.0313)	(0.0252)	(0.0227)	(0.0176)
Median family income	0.0219***	0.0191***	0.0196**	0.0056	−0.0062	−0.0011
	(0.0061)	(0.0067)	(0.0077)	(0.0079)	(0.0060)	(0.0032)
Median housing value	−0.0020	−0.0055	−0.0029	0.0039	0.0068	0.0006
	(0.0107)	(0.0091)	(0.0091)	(0.0163)	(0.0132)	(0.0063)

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 clustered at the state-level. The column labeled Kline and Moretti (2014a) replicates the Oaxaca-Blinder estimator found in column (3) of Table III. The column labeled 'Diff-in-Diff' estimates (11) by OLS. The final four columns labeled 'Diff-in-Diff with Spillovers' are estimates from (12).

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

For manufacturing, our point estimates for spillovers in column (4)-(6) are not statistically significant, but the signs on all three spillover estimates suggest that neighboring counties had a potentially negative spillover effect. Since there is little to no spillovers present, the new point estimate in column (3) of 4.6% is not too far from the standard difference-in-differences estimate of 5.6%. Similarly, there is not a significant spillover on median family income. The point estimate therefore stays the same between the standard and spillover-adjusted difference-in-differences estimate of about 1.9%.

These results show that including spillovers in the estimation of direct treatment effects is potentially important and in the case of agricultural employment can lead to *significant* differences in treatment effect estimates. Analysis of place-based policies that do not account for the fact that treatment effects can spillover beyond the borders of treated areas can potentially be biased. The above results suggest this can over or underestimate

the size of the effects depending on the sign of the spillovers.

5 — Event Study with Spillovers

The intuition from how spillovers cause biases in estimates of the direct effect of treatment extend into the setting where there is staggered adoption in treatment. Two-way fixed effects estimates can be viewed as a weighted sum of 2×2 difference-in-differences estimates and therefore the bias terms will be identically weighted of the bias terms from the 2×2 estimates, assuming the parallel counterfactual trends assumption (1).[19]

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

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

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

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

Assumption 2 (Parallel Counterfactual Trends on a "Never-Treated" Group)**.**

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

[19] Various forms for these weights are described in Goodman-Bacon (2018), Callaway and Sant'Anna (2020), Sun and Abraham (2020), and Chaisemartin and D'Haultfoeuille (2019). I won't recharacterize the weights in this article and guide interested readers to the source articles themselves.

Assumption 3 (Parallel Counterfactual Trends on "Not-Yet-Treated" Groups).

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

5.1. Estimating Direct Effects

The group-time average direct effect is the building block of analysis in the framework proposed by C&S. Estimation of the group-time average direct treatment effect, $ATT_{\text{direct}}(g,t)$ depends on which parallel trends assumption a researcher use. If a researcher is using Assumption 2, then estimation will be done using a supsample containing only control units and units in group g. The difference-in-difference estimate uses period t outcomes as the post-period and some period $g - \delta$ as the pre-period where δ is the number of periods before initial treatment. In the context of spillovers, the canoncial difference-in-differences estimation will have to be adjusted using a parameterized exposure mapping to prevent the bias detailed in the 2×2 setting. As shown in Section 3, this can be done using a set of donut indicators that is wide enough to capture all of the spillovers.

If a researcher is using Assumption 3, then estimation will be done with a subsample containing treated units from group g and all units that are in groups $g' > t$. Again, estimation of the group-time direct effect will have to include controls for spillovers.

Then, as detailed in C&S, these ATT can be aggregated using different weights depending on the pararmeter of interest (see Table 1 of Callaway and Sant'Anna (2020)). These aggregations take the form of

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

These weights include the option to make event study estimates, or group-level average treatment effects (across time). Estimation and different aggregation of the direct effects can be done using the 'did' package produced by CS where the measure of spillovers are included as covariates.

5.2. Estimating Spillover Effects

Estimation of the spillover effects is potentially much more difficult as estimation of continuous variables in the presence of staggered treatment is still a work in progress and

past the scope of this article. However, for indicator variables, estimation can be done in much the same way as the direct effect.

Since a set of donut indicators captures non-additive spillovers generally, I will focus on estimation of these donut indicators. For a given ring indicator, you would define groups by the period that a unit's ring indicator turns on. However, treated units can always move to more inner rings as treatment turns on over time, so control units should only enter estimation for the most inner indicator they enter.

Then, estimation can be done assuming that all control units are on the same counterfactual trends in the absence of spillovers. For a given ring, the 'treated' units consist of units that have that ring be the most inner spillovers they experience and the 'control' units consist of all control units whose spillover status remains constant between the pre-period, $g - \delta$ and t. As above, estimation of the group-time average spillover effect should include the other spillover indicators to avoid biases from other spillover effects.

Therefore, for each spillover donut, estimation and aggregation can be done in much the same way as for the direct effect. However, estimation and aggregation can not be done using the 'did' program as subsamples have to be constructed more specifically for each group-time average spillover effect.

6 — Conclusion

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

Further, I show that parameterizing these spillovers can remove most of the biases even in the presence of some misspecification. In particular, I find that specifications with a set of 'Donuts' are the best at capturing the spatial structure of spillovers. When estimating spillover effects themselves, I find that it is important to correctly identify is

29

spillovers are additive or non-additive in the number of nearby treated units.

Then, I show that these approaches can change estimates significantly in the context of estimating the effects of place-based policies. Since place-based policies change the nature of agglomeration in the local and surrounding area –that is cause spillovers–local effects of these policies can be misestimated without controlling for spillovers.

Finally, I show how researchers can include spillovers in settings with variation in treatment timing following Callaway and Sant'Anna (2020).

References

Angelucci, M. and V. Di Maro (Jan. 2016). "Programme evaluation and spillover effects". In: *Journal of Development Effectiveness* 8.1, pp. 22–43. ISSN: 1943-9342, 1943-9407. DOI: 10.1080/19439342.2015.1033441.

Angrist, Joshua D. (Oct. 2014). "The perils of peer effects". In: *Labour Economics* 30, pp. 98–108. ISSN: 09275371. DOI: 10.1016/j.labeco.2014.05.008.

Basker, Emek (Feb. 2005). "Job Creation or Destruction? Labor Market Effects of Wal-Mart". In: *The Review of Economics and Statistics* 87.1, pp. 174–183.

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

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

Callaway, Brantly and Pedro H.C. Sant'Anna (Dec. 2020). "Difference-in-Differences with multiple time periods". In: *Journal of Econometrics*, S0304407620303948. ISSN: 03044076. DOI: 10.1016/j.jeconom.2020.12.001.

Chaisemartin, Clément de and Xavier D'Haultfoeuille (May 2019). *Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects.* w25904, w25904. DOI: 10.3386/w25904. URL: http://www.nber.org/papers/w25904.pdf.

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

Conley, T.G. (1999). "GMM estimation with cross sectional dependence". In: *Journal of Econometrics* 92.1. Citation Key: CONLEY19991, pp. 1–45. ISSN: 0304-4076. DOI: https://doi.org/10.1016/S0304-4076(98)00084-0.

Delgado, Michael S. and Raymond J.G.M. Florax (Dec. 2015). "Difference-in-differences techniques for spatial data: Local autocorrelation and spatial interaction". In: *Economics Letters* 137, pp. 123–126. ISSN: 01651765. DOI: 10.1016/j.econlet.2015.10.035.

Duranton, Giles and Diego Puga (Sept. 2003). *Micro-Foundations of Urban Agglomeration Economies*. w9931, w9931. DOI: 10.3386/w9931. URL: http://www.nber.org/papers/w9931.pdf.

Goodman-Bacon, Andrew (Sept. 2018). *Difference-in-Differences with Variation in Treatment Timing*. w25018, w25018. DOI: 10.3386/w25018. URL: http://www.nber.org/papers/w25018.pdf.

Halloran, M. Elizabeth and Claudio J. Struchiner (Mar. 1995). "Causal Inference in Infectious Diseases:" in: *Epidemiology* 6.2, pp. 142–151. ISSN: 1044-3983. DOI: 10.1097/00001648-199503000-00010.

Keller, Wolfgang (2002). "Geographic Localization of International Technology Diffusion". In: *The American Economic Review* 92.1, p. 52.

Kline, Patrick (May 2011). "Oaxaca-Blinder as a Reweighting Estimator". In: *American Economic Review* 101.3, pp. 532–537. ISSN: 0002-8282. DOI: 10.1257/aer.101.3.532.

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

— (Aug. 2014b). "People, Places, and Public Policy: Some Simple Welfare Economics of Local Economic Development Programs". In: *Annual Review of Economics* 6.1, pp. 629–662. ISSN: 1941-1383, 1941-1391. DOI: 10.1146/annurev-economics-080213-041024.

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

Rambachan, Ashesh and Jonathan Roth (2020). *An Honest Approach to Parallel Trends*. Working Paper.

Rubin, Donald B. (Sept. 1980). "Randomization Analysis of Experimental Data: The Fisher Randomization Test Comment". In: *Journal of the American Statistical Association* 75.371, p. 591. ISSN: 01621459. DOI: 10.2307/2287653.

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

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

Sun, Liyang and Sarah Abraham (2020). "Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects". In: p. 53.

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

Verbitsky-Savitz, Natalya and Stephen W. Raudenbush (Jan. 2012). "Causal Inference Under Interference in Spatial Settings: A Case Study Evaluating Community Policing Program in Chicago". In: *Epidemiologic Methods* 1.1. ISSN: 2161-962X. DOI: 10.1515/2161-962X.1020. URL: https://www.degruyter.com/view/j/em.2012.1.issue-1/2161-962X.1020/2161-962X.1020.xml.

A — Proofs

Proof of Theorem 1

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

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

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

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

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

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

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

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

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

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

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

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

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