

# Differences-in-Differences with Spatial Spillover

Kyle Butts

Univ. of Colorado, Boulder

November 27, 2020

## **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, differences-in-differences produces biased estimates for the average treatment effect. In this paper, I identify two sources of bias. 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 one source of bias. Lastly, I propose improved estimation strategies that can remove both sources of bias.

## 1 — Introduction

Empirical work in economics often considers treatment assigned by geographic boundaries such as cities, counties, and states. The effect of these treatments do not typically stay within these boundaries whether it be from people crossing borders or from general equilibrium mechanisms affecting neighboring areas. In the causal inference literature, the Stable Unit Treatment Value Assumption (SUTVA) is often assumed which says that treatment effects do not depend on the treatment status of any other unit (Rubin, 1980). This article considers a common violation of SUTVA in a process I label ‘spatial spillovers’ where ‘close’ units treatment assignment affects outcomes of nearby units.<sup>1</sup>

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.<sup>2</sup> Of those 21, only eight include spillovers to prevent bias in the estimation of treatment effects.

In the presence of these spatial spillovers, I identify two sources of bias that result when estimating treatment effects by standard differences-in-differences methodologies. 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 differences-in-differences, the spillover onto the ‘close’ control units is averaged into the untreated 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, if a factory opening in a given county benefits both that county and neighboring counties, the treatment effect estimate is negatively biased because the change in outcome in neighboring counties is higher than it would be absent treatment.

Second, changes in treated units’ outcomes reflect the effect of their own treatment

---

<sup>1</sup> Close can refer to many different things, e.g. geographic distance, node distance in graph theory, or social relationships in schools or cities.

<sup>2</sup> 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.

status and the effect from the treatment status of “close” units. Therefore the spillover is added to the treated units’ change in outcomes. Therefore the bias of treatment effect is the same sign as the sign of the spillover. For example a factory opening in two neighboring counties might cause the benefit of each individual factory to decrease. The estimated treatment effect will be negatively biased because there is a negative spillover onto also treated units.

These biases are explicit functions of potential outcomes and can be used intuitively by researchers to sign the bias. The magnitude of the bias depends on two factors: the size of the spillover 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. This paper provides an explicit form for this bias and enables researchers to give bounds for it under assumptions about the size of the spillover effect and on number of units affected. I use Monte Carlo simulations to quantify the magnitude of bias in various spillover settings.

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.<sup>3</sup> Of those 21, only eight include spillovers in their empirical specification to prevent bias in the estimation of treatment effects.

@TODO: Paragraph on practical guidance on including spillover measures

### *1.1. Contribution*

There are two different models of SUTVA violations in the literature. Within-group spillovers are when units are in distinct groups and outcomes depend on the treatment status of your group only. For example, SUTVA is violated if a person’s likelihood of getting sick depends on the vaccine rate of his community. In this case, the community is the group within which spillovers can occur. Between-group spillovers occur when groups are overlapping which allows for more general forms of spillover. For example, a counties’

---

<sup>3</sup> 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.

economic outcomes depend on the economic outcomes of nearby counties (even across state borders). @TODO: Add citations for papers that find this.

One set of literature discusses ‘partial interference’ where spillovers are assumed to only occur within a group and not across groups.<sup>4</sup> This setting compares “control” units in a partially treated group with control units in completely untreated groups to estimate spillover effects.<sup>5</sup> However, the setting in my paper does not feature distinct groups which allow for the researcher to separate areas into distinct non-overlapping groups.

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  $h_i(\vec{D})$  (this is referred to as an exposure mapping). Adding an exposure mapping to the potential outcomes framework allows to explicitly model SUTVA violations. Using this framework, he shows that a difference in means comparison between treated and control units is a biased estimator 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 differences-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 of “treatment effect”. Then, the authors describe how the ATE estimated by diff-in-diff averages over both individual heterogeneity and, central to their contribution, over heterogeneity in the exposure mapping,  $h_i(\vec{D})$ . This paper develops a strategy to allow researchers to separately identify treatment effects and spillovers.

The paper closest to mine is Delgado and Florax (2015) where they consider spillover only on to control units and identify a bias that can come from estimating a diff-in-diff model without explicitly controlling for spillovers. Using Monte-Carlo simulations, they find that there is an omitted variable bias problem by not including a measure of spillovers.

---

<sup>4</sup> Halloran and Struchiner (1995) considers community-vaccine rates in epidemiology; Sobel (2006) considers interference in the Moving to Opportunity Program; and Angrist (2014) studies the context of school peer effects.

<sup>5</sup> Angelucci and Di Maro (2016) provides a summary for estimation of within-group treatment effects.

This paper derives an explicit form for this bias in terms of potential outcomes and also includes the presence of spillovers on to also treated units.

Similarly, Clarke (2017) finds an explicit form for bias when estimating a diff-in-diff model with spillover onto control units. However, their model only allows spillovers onto control units and assumes spillovers are constant regardless on the number of nearby treated units. Lastly, Berg and Streit (2019) and Verbitsky-Savitz and Raudenbush (2012) find results for a specific potential outcome function where  $h_i(\vec{D})$  is the proportion of treated contiguous counties. They allow this spillover to enter the potential outcome additively and the coefficient is allowed to differ by own treatment status.

In my framework, if I assume the particular functional forms for potential outcomes of Clarke (2017), Berg and Streit (2019), and Verbitsky-Savitz and Raudenbush (2012), I arrive at the same bias equation as theirs.

The paper is structured as follows. ...

Similarly, 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. Lastly, Berg and Streit (2019) and Verbitsky-Savitz and Raudenbush (2012) find results for a specific potential outcome function where  $h_i(\vec{D})$  is the proportion of treated contiguous counties. They allow this spillover to enter the potential outcome additively and the coefficient is allowed to differ by own treatment status. In my framework, if I assume the particular functional forms for potential outcomes of Clarke (2017), Berg and Streit (2019), and Verbitsky-Savitz and Raudenbush (2012), I arrive at the same bias equation as theirs.

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 illustrate the bias result and to evaluate currently used solutions in the literature. Section 4 presents an application for evaluating the Tennessee Valley Authority program.

## 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. As in Vazquez-Bare (2019), potential outcomes for unit  $i$  at time  $t$ , denoted  $Y_{i,t}(D_i, h_i(\vec{D}))$ , are a function of own treatment-status  $D_i$  and, departing from the canonical framework, of a function of the entire vector of treatment assignments  $h_i(\vec{D})$  where  $\vec{D} \in \{0, 1\}^n$ . The function  $h_i(\vec{D})$  is referred to as an ‘exposure mapping’ and is a non-negative scalar.<sup>6</sup> 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  $h_i(\vec{D}) = 0$ . The exposure mapping formalizes the SUTVA violation in that potential outcomes are affected by other unit’s treatment assignment and can be summarized by  $h_i(\vec{D})$ .

To help better understand an exposure mapping, I give three examples of  $h_i(\vec{D})$ . Let  $d(i, j)$  be a geographic distance measure which tells the distance unit  $i$  is from unit  $j$ . First,  $h_i(\vec{D})$  can be a 0/1 variable that equals one only if there is a treated unit within  $x$ -miles of unit  $i$  (also commonly used is counties that share borders). In this case

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

This exposure mapping is useful when it is assumed that spillovers do not decay over distance and the number of neighboring units treated do not impact the intensity of spillovers. For example, this 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 likely travel to a nearby library. Access to a neighboring town library likely does not depend on whether you can access 1 or more nearby libraries, so the binary value is a good approximation to the level of spillovers.

Second,  $h_i(\vec{D})$  can be a function that equals the proportion of the  $k$ -nearest neighbors that are treated, i.e.

$$h_i(\vec{D}) = 1/k \sum_{j \in k(i)} D_j,$$

---

<sup>6</sup> The derivation of bias does not require this assumption.

where  $k(i)$  is the index of unit  $i$ 's  $k$ -nearest neighbors. Last, is the spatial decay function which depends on the decay parameter  $\alpha$ . In this case, spillover intensity is the sum across all treated observations' decay term, i.e.

$$h_i(\vec{D}) = \sum_{j \neq i} D_j \exp^{-\alpha d_i(j)}$$

The exposure mapping is part of the parameterization of spillovers, but the functional form of the potential outcome still needs to be specified.

### 2.1. Spatial Spillovers

With a new definition of potential outcomes, I now formalize what is meant by “spatial spillovers”. I define ‘spillover onto control units’ as:

$$Y_i(0, h_i(\vec{D})) - Y_i(0, 0).$$

Then, the average spillover onto control units averages over potential heterogeneity in spillovers and over heterogeneity in exposure intensity  $h_i(\vec{D})$ :

$$\tau_{\text{spill,control}} = \mathbb{E}_i \left[ Y_i(0, h_i(\vec{D})) - Y_i(0, 0) \mid D_i = 0 \right].$$

To emphasize, the average spillover effect depends on the average of exposure mapping. For example, assuming the potential outcomes is additively linear in  $D_i$  and  $h_i(\vec{D})$ , then an individual control unit's spillover effect is  $\beta_{\text{spill,control}} h_i(\vec{D})$ . The average spillover onto control unit would therefore be  $\tau_{\text{spill,control}} = \beta_{\text{spill,control}} * \mathbb{E}_i \left[ h_i(\vec{D}) \right]$ , i.e. the average over all control units exposure mapping.

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

$$\tau_{\text{spill,treated}} = \mathbb{E}_i \left[ Y_i(1, h_i(\vec{D})) - Y_i(1, 0) \mid D_i = 1 \right].$$

It is important to clarify what I am assuming is the estimand of interest researchers would like to estimate when using differences-in-differences. I assume that the term ‘Average Treatment Effect’ is trying to measure what I will call the ‘direct effect of treatment’:

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

which measures the effect of being treated in the absence of exposure to spillovers. This differs from Sävje, Aronow, and Hudgens (2019) where they define the Average Treatment Effect as

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

where the expectation is over individuals and their exposures. I prefer the former because it allows for separate identification of the direct effect of treatment and the spillover effects themselves.<sup>7</sup>

## 2.2. Bias in Differences-in-Difference Estimation

In this section, I identify the two sources of bias in differences-in-differences estimation. For exposition, I will refer to two periods,  $t = 0$  and  $t = 1$ , but this can be replaced by averages across  $Y$  in the pre-period and the post-period respectively. First, I present the equivalent to the parallel counterfactual trends assumption in the context of the new potential outcome framework.

### Assumption 1 (Parallel Counterfactual Trends)

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

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 typical Parallel Counterfactual Trends because typically SUTVA assumes that every unit has zero exposure.

When researchers run the canonical differences-in-differences regression

$$y_{it} = \alpha + \tau D_{it} + \epsilon_{it},$$

$\hat{\tau}$  will be a biased estimate for  $\tau_{\text{direct}}$ . Given that Assumption 1 holds, the estimate can be decomposed as the direct effect and the two sources of spillover bias.

---

<sup>7</sup> The spillover effects themselves might be of interest to the researcher, so clearly separating them is beneficial in estimation. Later they can be combined to estimate the 'net effects' that Sävje, Aronow, and Hudgens (2019) estimate.



**Theorem 1** *If Assumption 1 holds,*

$$\begin{aligned}
\mathbb{E}[\hat{\tau}] &= \underbrace{\mathbb{E}[Y_{i1} - Y_{i0} \mid D_i = 1] - \mathbb{E}[Y_{i1} - Y_{i0} \mid D_i = 0]}_{\text{Difference-in-Differences}} \\
&= \mathbb{E}\left[Y_{i1}(1, \vec{0}) - Y_{i1}(0, \vec{0}) \mid D_i = 1\right] + \mathbb{E}\left[Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(1, \vec{0}) \mid D_i = 1\right] \\
&\quad - \mathbb{E}\left[Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0, \vec{0}) \mid D_i = 0\right] \\
&= \tau_{\text{direct}} + \tau_{\text{spill, treated}} - \tau_{\text{spill, control}}
\end{aligned}$$

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 onto the treated units. Second, the change in outcomes among control units combines the parallel counterfactual trend with the average spillover onto control units. Since  $\hat{\tau}$  is found by subtracting this second difference, we subtract the average spillover onto the control.

### 2.3. Bounding the Bias

The previous section finds an explicit form of the bias when estimating Equation 3. The explicit form of bias can be used by researchers to provide bounds on the bias in a manner similar to Rambachan and Roth (2020). To provide bounds on the bias, researchers must identify the set  $\Delta = [\underline{\Delta}, \bar{\Delta}]$  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 show a simple example that allows a researcher to bound the bias.

#### 2.3.1 Example Parameterization of Bias

To exemplify the above bias, I will model a particular form of spillovers by parameterizing the exposure mapping and the potential outcomes equation. For the exposure mapping, I

will assume that spillovers occur for any county that is within 40 miles of a treated county, i.e.

$$\text{Near}_{it} \equiv h_i(\vec{D}) = \max_{i \neq j} D_j * 1[d(i, j) < 40]. \quad (1)$$

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  $\bar{d}$ . The potential outcome is assumed to be additive in both treatment and in the exposure mapping, but the coefficient on  $\text{Near}_{it}$  is allowed to differ between treated and non-treated units:

$$y_{it} = \mu_t + \mu_i + \beta_{\text{direct}} D_{it} + \beta_{\text{spill,control}} (1 - D_{it}) \text{Near}_{it} + \beta_{\text{spill,treated}} D_{it} \text{Near}_{it} + \varepsilon_{it}, \quad (2)$$

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

If a researcher estimates

$$y_{it} = \alpha + \mu_t + \mu_i + \tau D_{it} + \varepsilon_{it}, \quad (3)$$

then the bias will be  $\tau_{\text{spill,treated}} - \tau_{\text{spill,control}}$  where the two bias terms are

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

and

$$\tau_{\text{spill,treated}} = \beta_{\text{spill,treated}} \frac{\sum_{i:D_{it}=1} \text{Near}_{it}}{N_T},$$

where  $N_C$  and  $N_T$  are the number of control and treated units respectively. Simply put, the bias will be the product of the proportion of treated units that receive spillover effects and the size of the effect on treated units minus the proportion of control units that receive spillover effects time the size of the effect on control units.

Then researchers wanting to create the bias set  $\Delta$ , will assume minimum and maximum values of  $\beta_{\text{spill,control}}$ ,  $\beta_{\text{spill,treated}}$ , and a maximum proportion of exposed treated and control units. With these assumptions, the set  $\Delta$  can be approximated and the partially identified set is given by  $[\hat{\tau} + \underline{\Delta}, \hat{\tau} + \bar{\Delta}]$ . Inference on this set can be done in the manner described in Rambachan and Roth (2020).

In the example of library construction, a researcher would assume that people would not drive further than 40 miles to access the library. Further, a researcher can argue that treatment effect does not depend on whether other counties also have access to libraries. Therefore, only  $\beta_{\text{spill,control}}$  needs to be bounded for the bias to be bounded.

#### 2.4. Improving Estimation by Parameterizing Spillovers

The previous section shows how to estimate the magnitude of potential bias in the estimation of Equation 3. However, parameterization of the potential outcomes and exposure mapping allows for unbiased estimates of  $\tau_{\text{direct}}$  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.

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 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, the net benefits for a treated location depend on the direct effect and the spillover effect onto treated units. Therefore whether the benefits depend on the number of treated units nearby.

There are a few practical considerations researchers should make when parametrizing the spillovers.

In a set of Monte Carlo simulations, I check the sensitivity of estimates to misspecification of spillovers.

### 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. The data generating process used is defined by the exposure mapping given in Equation 1 and the data generating process given in Equation 2. A unit of observation is a US county and the periods are  $t \in \{1, \dots, 20\}$  with treatment turning on in period  $t = 10$ . The intercept is  $\alpha = -2$ , the fixed effects are  $\mu_t \sim N(t * 0.2, 0.1^2)$  and  $\mu_i \sim N(6, 2^2)$ , and the error term is  $\epsilon \sim N(0, 2^2)$ . 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.  $\tau_{\text{spill, treat}} = 0$  and estimate the typical two-way fixed effects model. This data generating process where spillovers only occur onto control units matches what has been typically assumed in the literature. I assign treatment among U.S. counties randomly with various probabilities between 3 percent and 50 percent. The DGP is therefore

$$y_{it} = -2 + \mu_t + \mu_i + 2D_{it} - (1 - D_{it})\text{Near}_{it} + \epsilon_{it} \quad (4)$$

The size of the bias from estimating Equation 3 at different treatment probabilities are presented in Figure 1 as the line with diamond markers. 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.

A common solution in the literature is to remove control units from the estimated sample that are most likely to be affected by spillovers. I do this in Figure 1 and the results are shown by the line with circle markers. Even though, I remove contiguous controls which is not the correct measure of  $h_i(\vec{D})$ , it approximates it well enough such that the bias stays centered constantly around zero as most control units experiencing spillovers are removed. As the portion of control units receiving spillovers that are kept in the sample increases, 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 methodology. As the treatment probability increases, the number of control units removed increases as well. This naturally yields a more variable estimator as seen in the wider 95 percent empirical confidence intervals.

### 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

$$y_{it} = -2 + \mu_t + \mu_i + 2D_{it} - (1 - D_{it})\text{Near}_{it} - 0.5D_{it}\text{Near}_{it} + \varepsilon_{it} \quad (5)$$

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.

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 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. The unconditional probability is equal to 5 percent in all simulations. To do this, I use a variable ‘Zone Plus’ to adjust the relative probability of receiving treatment:

$$P(D_i | \text{Zone}_i) = (.1 + \text{Zone Plus} * \text{Zone}_i) \frac{.05}{.1 * .9 + (.1 + \text{Zone Plus}) * .1} \quad (6)$$

The second term normalizes probabilities so the unconditional probability stays at 5 percent. In our simulations, Zone Plus ranges from 0 which is the case where  $p = 10$  percent for all counties to 2 which has  $P(D_i | \text{Zone}_i) = 0.166 + 0.35 * \text{Zone}_i$

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 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. The unconditional probability is equal to 5 percent in all simulations. To do this, I use a variable ‘Zone Plus’ to adjust the relative probability of receiving treatment:

$$P(D_i | \text{Zone}_i) = (.1 + \text{Zone Plus} * \text{Zone}_i) \times \frac{.05}{.1 * .9 + (.1 + \text{Zone Plus}) * .1} \quad (7)$$

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  $p = \text{five percent}$  for all counties to 2 which has  $P(D_i | \text{Zone}_i) = 0.0166 + 0.35 * \text{Zone}_i$

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 actually 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 DGP, 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  $\tau_{\text{spill},\text{control}}$  but  $\tau_{\text{spill},\text{treated}}$  remains. To remove the second source of bias, it is necessary to account directly for spillovers in the estimation strategy.

### 3.3. Misspecification of Spillovers

## 4 — Application in Urban Economics

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 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.<sup>8</sup> 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 (2014) 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.

There are two primary forces causing spillover in this context. First, the large-scale investment in dams and electricity transmission lines brought cheap power to the region. It became cheaper for nearby locations to get connected to low-cost power and therefore this likely caused a transition to manufacturing and away from farming in nearby counties. Second, Duranton and Puga (2003) highlight the agglomeration of input sourcing. When the Authority has a large increase in manufacturing, they develop resource supply chains that in turn make inputs cheaper for neighboring counties. These both suggest that spillovers will occur close to the spillover region and that the exposure will decay over distance.

The analysis in Kline and Moretti (2014) begins by comparing long-differences in county-level outcomes from 1940 to 2000 between treated counties in the Authority and control counties outside. The primary difference-in-differences specification is

$$y_{c,t} = \mu_c + \alpha \text{Post}_t + \tau \text{TVA}_c \times \text{Post}_t + X_{c,1940}\beta + X_{c,1940} \times \text{Post}_t \gamma + \epsilon_{c,t}, \quad (8)$$

where  $c$  denotes county,  $\text{TVA}_c$  is an indicator variable for being in the Authority,  $\text{Post}_t$  equals one when  $t = 2010$  and equals zero when  $t = 1940$ , and  $y$  is a set of outcome variables.<sup>9</sup> Pre-treatment control variables,  $X_{c,1940}$ , are interacted with  $\text{Post}_t$  to allow for

---

<sup>8</sup> More details on the program are found in Kline and Moretti (2014).

<sup>9</sup> The text writes the equivalent first-differences regression:

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

Outcome variables are all in logs and include population, average manufacturing wage, agricultural employ-

places to be on different long-term trends.<sup>10</sup>

Kline and Moretti (2014) estimate Equation 8 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 between TVA counties with neighboring counties that likely were impacted by the large-scale electrification program. As described above and estimated below, the program likely induced positive manufacturing and negative agricultural employment effects in nearby counties. Therefore, the authors likely underestimate the impact on employment by comparing counties inside the Authority to counties on the other side of the border. 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 weights can be seen in Figure III of their text, but generally counties close to the TVA are upweighted. Hence, their method magnifies the bias compared to a standard difference-in-differences estimator. However, in Table 1 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.

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  $\text{Dist} = \{(0, 100], (100, 250], (250, 500]\}$  measured in miles and define  $\text{Between}(d)$  as an indicator for being within the interval  $d \in \text{Dist}$  away from the Authority. As discussed above, I suspect the spillover to decrease as counties move further away, hence I allow the effect to be separately identified in bins. However, I

---

ment, manufacturing employment, value of farm production, median family income, and median housing value.

<sup>10</sup> See footnote 8 in Kline and Moretti (2014) for a full listing of control variables.

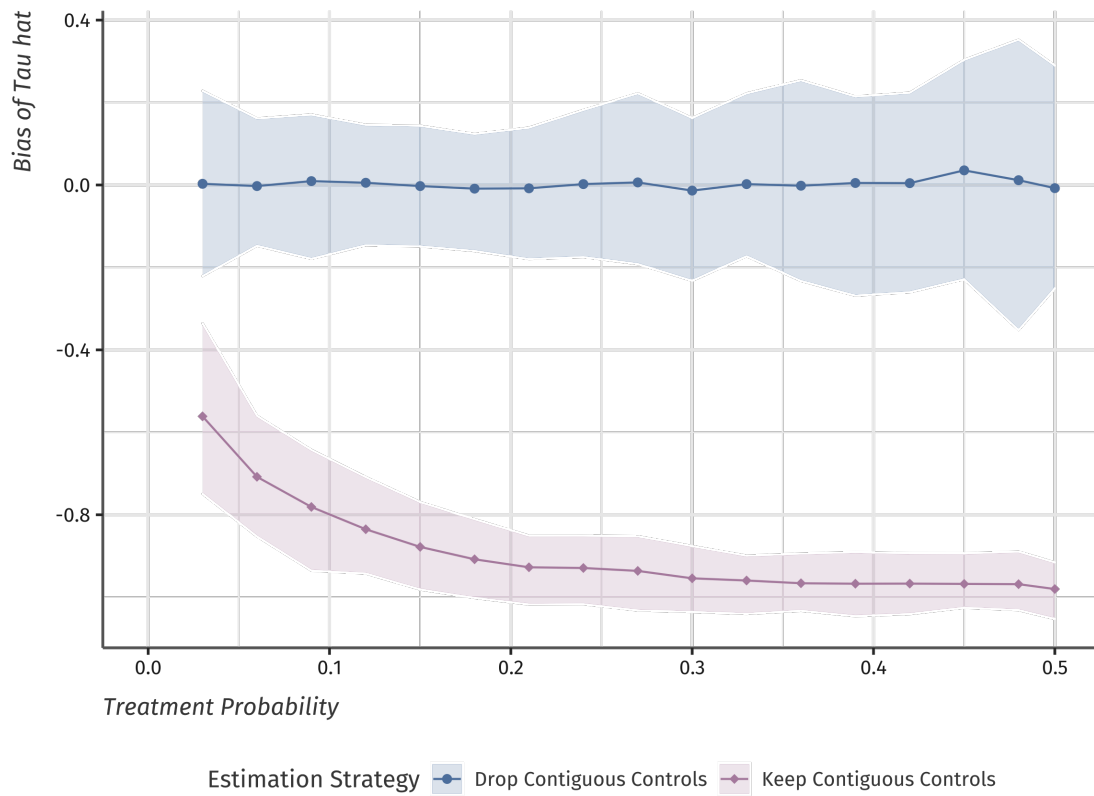


The specification with spillovers is given as follows:

$$y_{i,2000} - y_{i,1940} = \alpha + \text{TVA}_i \tau + \sum_{d \in \text{Dist}} \text{Between}(d) \delta_d + X_{i,1940} \beta + (\varepsilon_{i,2000} - \varepsilon_{i,1940}) \quad (9)$$

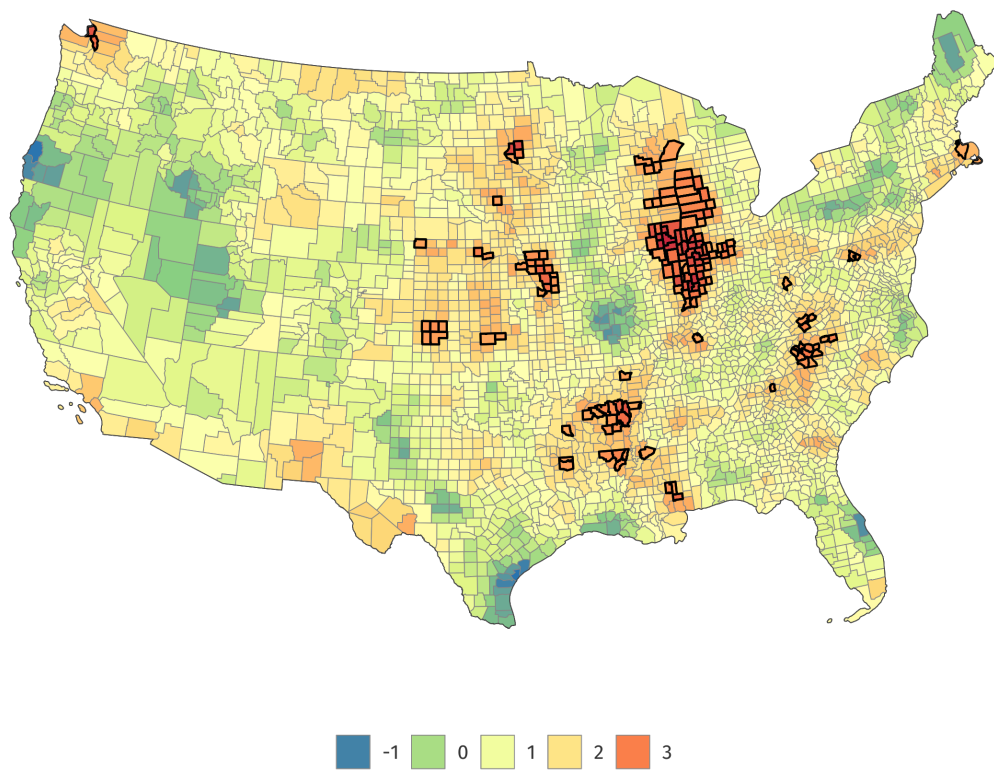
The coefficients  $\delta_d$  measure the spillover effect at different distances from the Authority.

**Figure 1 — Bias of  $\hat{\tau}$  with and without Control Units with Spillovers**



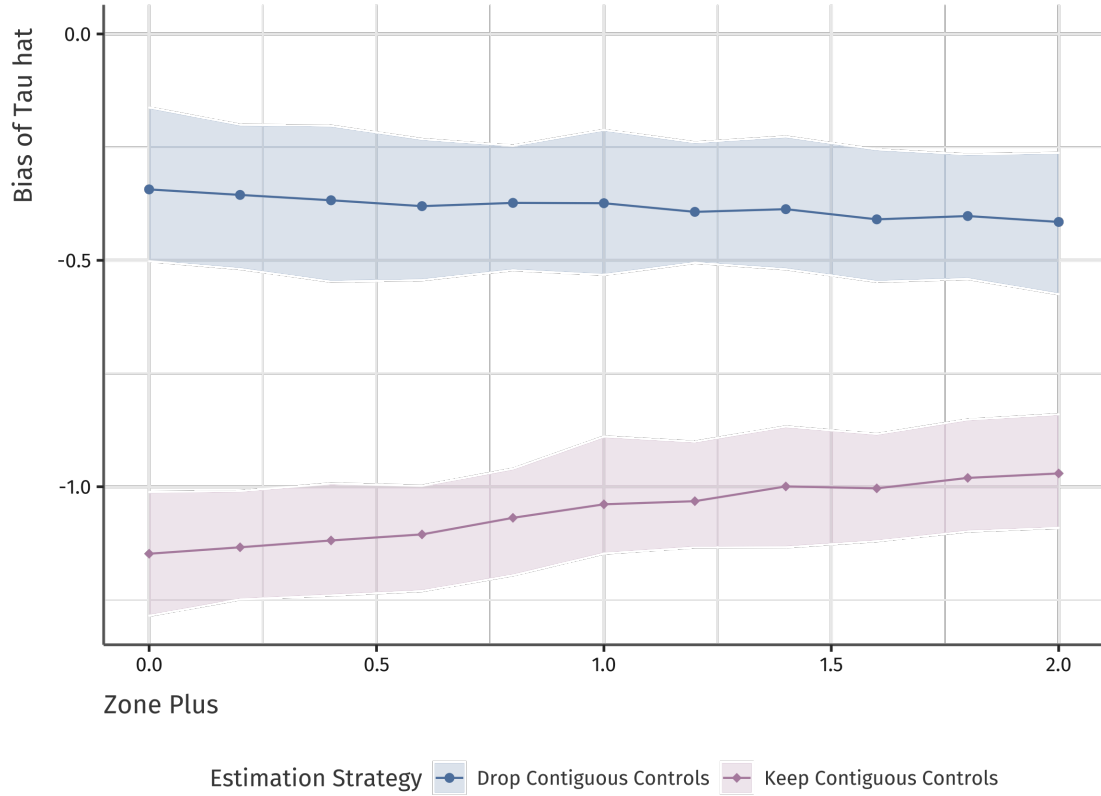
*Notes:* This figure plots the bias of  $\hat{\tau}$  found from estimating Equation 3 for data generated as described in the text and Equation 4. Each point corresponds to the average bias for the given treatment probability and the band is the 95 percent empirical confidence interval over 1000 simulations. 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.

**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.

**Figure 3 — Bias of  $\hat{\tau}$  at Different Levels of Spatial Autocorrelation**



*Notes:* This figure plots the bias of  $\hat{\tau}$  found from estimating Equation 3 for data generated as described in the text and Equation 5. Treatment probability is assigned according to the method described by the text and Equation 7. 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.

**Table 1 — Effects of Tennessee Valley Authority on Decadel Growth**

	Kline & Moretti (2011)	Diff-in-Diff	Diff-in-Diff with Spillovers			
	TVA	TVA	TVA	TVA between 0-100 mi.	TVA between 100-250 mi.	TVA between 250-500 mi.
<i>Dependent Var.</i>	(1)	(2)	(3)	(4)	(5)	(6)
Population	0.0073 (0.0112)	0.0075 (0.0109)	−0.0003 (0.0112)	−0.0196 (0.0184)	−0.0069 (0.0125)	−0.0043 (0.0128)
Average manufacturing wage	0.0059 (0.0047)	0.0067** (0.0029)	0.0066 (0.0043)	0.0019 (0.0050)	−0.0066 (0.0047)	0.0026 (0.0054)
Agricultural employment	−0.0645*** (0.0106)	−0.0584*** (0.0099)	−0.0760*** (0.0101)	−0.0351*** (0.0134)	−0.0179* (0.0097)	−0.0249*** (0.0090)
Manufacturing employment	0.0625*** (0.0164)	0.0588*** (0.0157)	0.0518*** (0.0179)	−0.0057 (0.0210)	−0.0099 (0.0238)	−0.0220 (0.0139)
Value of farm production	0.0025 (0.0225)	−0.0049 (0.0262)	−0.0161 (0.0291)	−0.0170 (0.0245)	−0.0144 (0.0251)	−0.0225 (0.0190)
Median family income	0.0242*** (0.0066)	0.0214*** (0.0073)	0.0217*** (0.0083)	0.0060 (0.0078)	−0.0051 (0.0062)	−0.0027 (0.0037)
Median housing value	0.0016 (0.0110)	−0.0015 (0.0094)	0.0020 (0.0093)	0.0052 (0.0166)	0.0101 (0.0140)	0.0001 (0.0071)

*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 (2014) replicates the Oaxaca-Blinder estimator found in column (3) of Table III. The column labeled ‘Diff-in-Diff’ estimates Equation 8 by OLS. The final four columns labeled ‘Diff-in-Diff with Spillovers’ are estimates from Equation 9.

\* $p < 0.1$ ; \*\* $p < 0.05$ ; \*\*\* $p < 0.01$ .

The results of the analysis are presented in Table 1. The first column lists which dependent variable (measured in logs) was used in the row. The following columns contain point estimates for  $\tau$  and  $\delta_s$  in different specifications and can be interpreted as decadal 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 Equation 8. The two columns are very close in magnitude to each other. They find a decline in agricultural employment of about 5.7% per decade, an increase in manufacturing employment of about 7% per decade, and an increase in median family income of about 2.6% per decade.

Turning to the last four columns that estimate Equation 9, the columns contain point estimates for  $\tau$  and  $\delta_d$ . For agricultural employment, our point estimates show there was a decline in agriculture employment in control units near to the authority. This is likely due to the fact that electrification of the region provided benefits beyond the border of the authority.<sup>11</sup> 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 9.6% per decade compared to 5.7%.

Similarly, manufacturing employment grew in nearby counties likely due to agglomeration economies and the improved regional transportation system. However, the size of the spillovers are smaller than in the agricultural employment regression, so the bias is smaller in this instance. The new point estimate is an increase of manufacturing employment of 8.54% per decade compared to an original estimate of 7%. Median family income similarly increased in nearby counties and therefore the estimate was negatively biased. The point estimate is an increase of family income of 3.5% per decade compared to 2.58%. These results show that including spillovers in the estimation of direct treatment effects is important and can lead to *significant* differences in treatment effect estimates.

---

<sup>11</sup> As noted in footnote 9 of Kline and Moretti (2014), the spillover could theoretically be of either sign. It could be positive if farmers chose to locate in the authority to receive the benefits instead of just over the border.

## 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](https://doi.org/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](https://doi.org/10.1016/j.labeco.2014.05.008).
- 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*.
- Clarke, Damian (2017). “Estimating Difference-in-Differences in the Presence of Spillovers”. In: *Munich Personal RePEc Archive*, p. 52.
- 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](https://doi.org/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](https://doi.org/10.3386/w9931). URL: <http://www.nber.org/papers/w9931.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](https://doi.org/10.1097/00001648-199503000-00010).
- 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](https://doi.org/10.1257/aer.101.3.532).
- Kline, Patrick and Enrico Moretti (Feb. 2014). “Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority”. In: *The Quarterly Journal of Economics* 129.1, pp. 275–331. DOI: [10.1093/qje/qjt034](https://doi.org/10.1093/qje/qjt034).

- 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](https://doi.org/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](https://doi.org/10.1198/016214506000000636).
- 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](https://doi.org/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

$$\begin{aligned}
\mathbb{E}_i [\hat{\tau}] &= \mathbb{E}_i [Y_{i1} - Y_{i0} \mid D_i = 1] - \mathbb{E}_i [Y_{i1} - Y_{i0} \mid D_i = 0] \\
&= \mathbb{E}_i \left[ Y_{i1}(1, h_i(\vec{D})) - Y_{i0}(0, \vec{0}) \mid D_i = 1 \right] - \mathbb{E}_i \left[ Y_{i1}(0, h_i(\vec{D})) - Y_{i0}(0, \vec{0}) \mid D_i = 0 \right] \\
&= \mathbb{E}_i \left[ Y_{i1}(1, h_i(\vec{D})) - Y_{i0}(0, \vec{0}) \mid D_i = 1 \right] - \mathbb{E}_i \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}_i \left[ Y_{i1}(1, h_i(\vec{D})) - Y_{i0}(0, \vec{0}) \mid D_i = 1 \right] - \mathbb{E}_i \left[ Y_{i1}(0, \vec{0}) - Y_{i0}(0, \vec{0}) \mid D_i = 0 \right] \\
&\quad - \mathbb{E}_i \left[ Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0, \vec{0}) \mid D_i = 0 \right] \\
&= \mathbb{E}_i \left[ Y_{i1}(1, h_i(\vec{D})) - Y_{i0}(0, \vec{0}) \mid D_i = 1 \right] - \mathbb{E}_i \left[ Y_{i1}(0, \vec{0}) - Y_{i0}(0, \vec{0}) \mid D_i = 1 \right] \\
&\quad - \mathbb{E}_i \left[ Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0, \vec{0}) \mid D_i = 0 \right] \\
&= \mathbb{E}_i \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] - \mathbb{E}_i \left[ Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0, \vec{0}) \mid D_i = 0 \right] \\
&= \mathbb{E}_i \left[ Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(0, \vec{0}) \mid D_i = 1 \right] - \mathbb{E}_i \left[ Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0, \vec{0}) \mid D_i = 0 \right] \\
&= \mathbb{E}_i \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] - \mathbb{E}_i \left[ Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0, \vec{0}) \mid D_i = 0 \right] \\
&= \mathbb{E}_i \left[ Y_{i1}(1, \vec{0}) - Y_{i1}(0, \vec{0}) \mid D_i = 1 \right] + \mathbb{E}_i \left[ Y_{i1}(1, h_i(\vec{D})) - Y_{i1}(1, \vec{0}) \mid D_i = 1 \right] \\
&\quad - \mathbb{E}_i \left[ Y_{i1}(0, h_i(\vec{D})) - Y_{i1}(0, \vec{0}) \mid D_i = 0 \right] \\
&= \tau_{\text{direct}} + \tau_{\text{spill, treated}} - \tau_{\text{spill, control}}
\end{aligned}$$