# Design-Based Inference for Spatial Experiments with Interference

Peter M. Aronow, Cyrus Samii and Ye Wang*

July 14, 2019

**Abstract**

We consider design-based causal inference in settings where randomized treatments have effects that bleed out into space in complex ways that overlap and in violation of the standard "no interference" assumption for many causal inference methods. We define a spatial "marginalized individualistic response," which characterizes how, on average, units of observation that are a specified distance from an intervention point are affected by treatments at that point, averaging over effects emanating from other intervention points. We establish conditions for non-parametric identification, non-parametric consistency, and consistent recovery of structural effects. We propose methods for both sample-theoretic and permutation-based inference.

Version 0.2

*Peter M. Aronow is Associate Professor, Departments of Political Science and Biostatistics, Yale University, 77 Prospect St., New Haven, CT 06520 (Email: peter.aronow@yale.edu). Cyrus Samii (contact author) is Associate Professor, Department of Politics, New York University, 19 West 4th St., New York, NY 10012 (Email: cds2083@nyu.edu). Ye Wang is PhD student, Department of Politics, New York University, 19 West 4th St., New York, NY 10012 (Email: yw1576@nyu.edu).

# 1 Introduction

Consider an experiment where treatments are assigned to specific points in a geographic space. Then, we observe how the treatments affect outcomes over various points in this geography. An example, with which we work below, is a forest conservation experiment. Monitoring stations are set up at points in and around a forest area, and we use satellite data to track deforestation outcomes throughout the forest. These treatments may generate effects that bleed out in complex ways. In the forest conservation experiment, the effect of the stations may be to displace deforestation. This would cause deforestation to decrease near a station, but then to increase further away. The displacement effect of one station may interact with the effect from another station, making the overall effects complex. We can imagine similarly complex spatial effects from other types of treatments. Vaccination campaigns can have spatial effects on disease prevalence through herd immunity. Product campaigns can have spatial effects on consumer behavior by word-of-mouth, contagion, and equilibrium responses of firms. Such spatial effects present instances of "interference," whereby effects are a product not of a given unit's treatment status, but rather of the overall distribution of treatments (Cox, 1958, p. 19).

This paper develops design-based methods for analyzing such spatial experiments, accounting for interference. Design-based methods derive causal and statistical inferential properties off of the experimental design, which is typically under the control of the analyst and therefore known to the analyst. We focus on a causal quantity that we call the "marginalized individualistic response" (MIR). This quantity measures how, on average, outcomes at a given distance from an intervention point are affected by activating a treatment at that intervention point, taking into account ambient effects emanating from treatments at other intervention points. These ambient effects depend on the overall level of treatment saturation. We show that the MIR is non-parametrically identified as a simple contrast

under randomization. This is true even under interference. Reliable statistical inference follows from more stringent conditions on potential outcomes. We also show that there is a direct mapping from the MIR to the types of effects that are assumed by parametric models of spatial effects. That is, when effects emanating from different intervention points are additive, the MIR recovers the average of these additive effects. If, however, spatial effects are not simply additive, but rather exhibit complex interactions, the MIR still yields an interpretable quantity.

Our analysis is related to a few streams of current methodological research. First, our approach draws most directly on recent design-based analyses of causal effects under interference. This includes the idea of using permutation for inference when interference is present, as in Rosenbaum (2007), Tchetgen-Tchetgen and VanderWeele (2010), Bowers et al. (2013), and Aronow and Samii (2017), as well as the idea of defining estimands that marginalize over the randomization distribution, as in Hudgens and Halloran (2008), Manski (2012), and Savje et al. (2018). Second, our analysis is related to recent work on "bipartite causal inference" by Zigler and Papadogeorgou (2018), who define different types of direct and indirect effect estimands in a setting that is otherwise similar to ours—that is, treatments assigned to points in space that then have effects that emanate outward into space.

We begin by developing the formal inferential setting and main theoretical results, using a simple toy example to illustrate concepts. We then develop a number of extensions and refinements. We then turn to an application based on a forest conservation experiment.

# 2 Setting

Suppose a set of $\mathcal{N}$ of intervention points indexed by $i = 1, ..., N$. These point reside on a two-dimensional field $\mathcal{X}$ of points indexed by $x = (x_1, x_2)$ (e.g., latitude and longitude). An experimental design assigns a binary treatment $Z_i = 0, 1$ to each intervention. The ordered

vector of experimental assignments variables is $\mathbf{Z} \equiv (Z_1, ..., Z_n)$, and the *ex post* realized assignment from the experiment is given by $\mathbf{z} \equiv (z_1, ..., z_n)$. The experimental design fixes the set of possible assignment vectors $\mathcal{Z}$ as well as a probability distribution over that set, $\mathbb{P}_{\mathbf{z}}$. Our analysis below considers the case of an experimental design based on Bernouilli randomization for each $Z_i$—i.e., (possibly weighted) coin flips to determine treatment status at each point.

Potential outcomes at any point $x \in \mathcal{X}$ are defined for each value of $\mathbf{z}$, $(Y_x(\mathbf{z}))_{\mathbf{z}\in\mathcal{Z}}$. Given a realized treatment assignment, $\mathbf{z}$, we observe the corresponding potential outcome at $x$:

$$Y_x = \sum_{\mathbf{z}\in\mathcal{Z}} Y_x(\mathbf{z}) I(\mathbf{Z} = \mathbf{z})$$

Data for points in $\mathcal{X}$ may come in various formats including raster data or data for a discrete set of points in $\mathcal{X}$. Let $\mathbf{Y}(\mathbf{z}) = (Y_x(\mathbf{z}))_{x\in\mathcal{X}}$ denote the full set of potential outcomes when $\mathbf{Z} = \mathbf{z}$ and $\mathbf{Y} = (Y_x(\mathbf{Z}))_{x\in\mathcal{X}}$ denote the full set of realized outcomes.

We map the full set of potential outcomes $\mathbf{Y}(\mathbf{z})$ for all points in the outcome space $\mathcal{X}$ back to the intervention points in $\mathcal{N}$ by defining the "circle average" function:

$$\mu_i(\mathbf{Y}(\mathbf{z}); d) = \frac{1}{|\{x : d_i(x) = d\}|} \int_{x\in\{x:d_i(x)=d\}} Y_x(\mathbf{z}) dx,$$

where $d_i(x) = ||x(i) - x||$ measures the distance between point $x$ and intervention point $i$, which is located at $x(i)$. Therefore, $\{x : d_i(x) = d\}$ is the circle with radius $d$ around intervention point $i$, and finally, $|\{x : d_i(x) = d\}|$ measures the circumference of this circle. Thus, $\mu_i(\mathbf{Y}; d)$ returns the average of points along the edge of the circle centered on intervention point $i$ and with radius $d$. The realized circle average for intervention point $i$ is thus,

$$\mu_i(\mathbf{Y}; d) = \sum_{\mathbf{z}\in\mathcal{Z}} \mu_i(\mathbf{Y}(\mathbf{z}); d) I(\mathbf{Z} = \mathbf{z}).$$

This representation allows us to see how an experiment is a process of sampling potential circle averages for intervention points, and therefore allows us to apply sample theoretic (in particular, Horvitz-Thompson-type) results in our analysis below

When observation points are discrete points, then the circle average function would use a coarsened distance operator,

$$d_i(x;\kappa) = \lfloor \|x(i) - x\| \times 10^{\kappa} + 0.5 \rfloor \times 10^{-\kappa}.$$

The function $d_i(x;\kappa)$ measures the coarsened distance between point $x$ and intervention point $i$, with coarsening parameter, $\kappa$. Below, we focus on the continuous case, and extension to the coarsened case would follow in a straightforward manner by substituting appropriately.

Figure 1 illustrates a toy example with $N = 4$ and raster outcome data. The figure shows a "null raster" for which none of the intervention points has been assigned to treatment and so $\mathbf{z} = (0, 0, 0, 0)$, in which case outcomes are $Y_x((0, 0, 0, 0))$ for all $x$ in the space. As we can see, outcomes are defined for any point $x$ in the space, although outcomes are constant within raster cells. This is a feature of raster data. Other types of data may exhibit finer levels of granularity — e.g., data produced from kriging interpolation that varies smoothly in space. We take these outcome data, and any coarsening or smoothing operations that they incorporate, as fixed. For our design-based inference, the only source of stochastic variation is from $\mathbf{Z}$.

As discussed above, spatial effects can exhibit considerably complexity. For the sake of illustration, suppose in our toy example that treatments tend to transmit effects that are non-monotonic in distance and that these effects simply accumulate in an additive manner. Figure 2 illustrates such an effect function. Then, the net result would depend on how treatments are distributed over the intervention points. Figure 3 illustrates how outcomes would be affected over different allocations of the treatment given that effects take the form

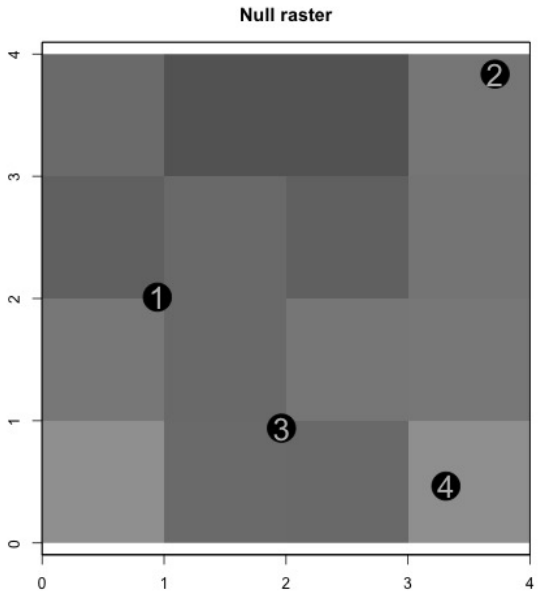


Figure 1: Illustration of a "null raster," with $N = 4$ intervention points, none of which are assigned to treatment. The coloring of the raster cells corresponds to outcome levels.

as in Figure 2. We emphasize that in the analysis below, we do not assume that effects are additive or homogenous in form—this is done here merely to provide a simple illustration.

# 3 Defining a marginal spatial effect

As the potential outcome notation for $Y_x(\mathbf{z})$ indicates, the outcome at any point may depend on the full vector of realized treatment assignments, $\mathbf{z}$. Similarly, the potential outcome notation for the circle average, $\mu_i(\mathbf{Y}(\mathbf{z}); d)$, suggests that the realized circle average for point $i$ may depend on treatment assignments for points other than $i$. As such, the circle averages are potentially subject to causal interference.

We now define a spatial effect that we call the "marginalized individualistic response" (MIR). The MIR accounts for interference. It is a marginal effect (Rubin, 2005) that bears resemblance to the "average direct causal effect" defined by Hudgens and Halloran (2008) and the "expected average treatment effect" defined by Savje et al. (2018). The usual definition

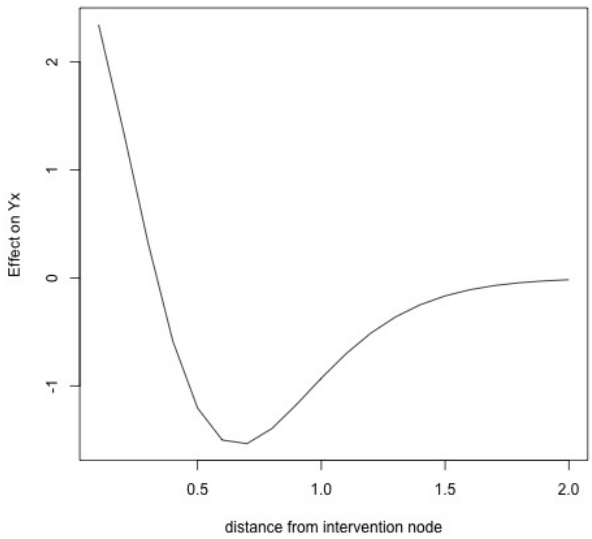


Figure 2: Illustration of a possible effect function that such that treatments transmit effects non-monotonically in distance. When multiple intervention points are treated, these effects accumulate.

of a unit-level treatment effect takes the difference between a unit's potential outcomes under one treatment condition versus under another treatment condition. A unit-level marginal effect is different, because it takes the difference between the average of a unit's potential outcomes over a set of potential outcomes versus the average over another set. We apply this idea to the spatial setting. In doing so, we consider effects that may bleed out in ways that are not necessarily contained within pre-defined group boundaries, as in Hudgens and Halloran (2008).

To define the spatial MIR, first rewrite the potential outcome at point $x$ as $Y_x(z_i, \mathbf{z}_{-i})$, where $\mathbf{z}_{-i}$ is vector equalling $\mathbf{z}$ except that the value for point $i$ is omitted. This allows us to pay special attention to how variation in treatment at point $i$ relates to potential outcomes at point $x$, given variation in treatment values in $\mathbf{z}_{-i}$. We can marginalize over variation in $\mathbf{z}_{-i}$ to define an "individualistic" average of potential outcomes for point $x$, holding the

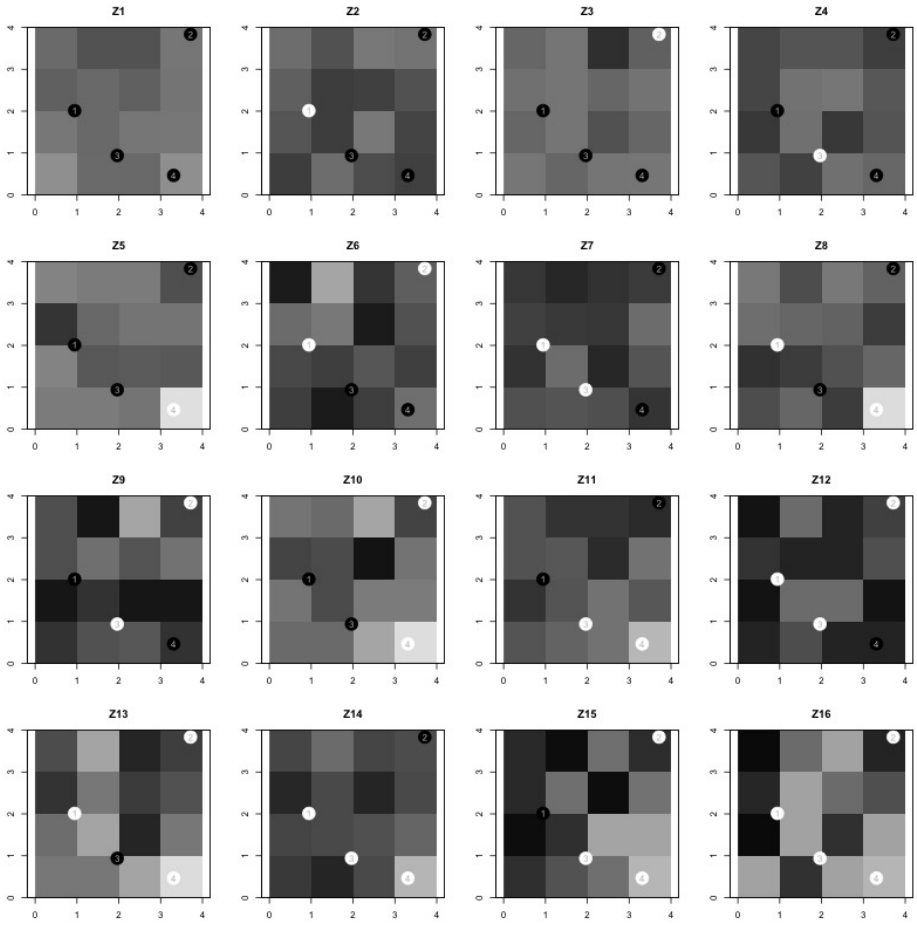


Figure 3: Illustration of how outcomes are affected given different allocations of treatment and effects that take the form as presented in Figure 2. Treated intervention points are white, while non-treated intervention points are black.

treatment at intervention point $i$ to treatment value $z$:

$$Y_{ix}(z;\alpha) = \mathrm{E}_{\mathbf{Z}_{-i}}[Y_x(z,\mathbf{Z}_{-i})] = \sum_{\mathbf{z}_{-i}\in\mathcal{Z}_{-i}} Y_x(z,\mathbf{z}_{-i})\Pr(\mathbf{Z}_{-i}=\mathbf{z}_{-i}|Z_i=z,\alpha),$$

where $\alpha$ is the experimental design parameter that governs the distribution of $\mathbf{Z}$ and $\mathcal{Z}_{-i}$ is the set of possible values that $\mathbf{Z}_{-i}$ can take. This is the individualistic marginal potential outcome at point $x$ given that point $i$ is assigned to treatment condition $z$, marginalizing over possible assignments to other points. We can use Figure 3 to illustrate. To construct $Y_{1x}(0;\alpha)$, one would take a weighted average of the potential outcomes at point $x$ under assignments labeled in the figure as Z1, Z3, Z4, Z5, Z9, Z10, Z11, and Z15, where the weights would be proportional to the probability of each assignment.

We can define a similar marginal quantity at the level of the circle averages:

$$\bar{\mu}_i(z;d,\alpha) = \mathrm{E}_{\mathbf{Z}_{-i}}[\mu_i(\mathbf{Y}(z,\mathbf{Z}_{-i});d)] = \sum_{\mathbf{z}_{-i}\in\mathcal{Z}_{-i}} \mu_i(\mathbf{Y}(z,\mathbf{z}_{-i});d)\Pr(\mathbf{Z}_{-i}=\mathbf{z}_{-i}|Z_i=z,\alpha),$$

where we use $\mathbf{Y}(z,\mathbf{z}_{-i})$ to denote the set of potential outcomes over points in $\mathcal{X}$ that obtain under treatment assignment $(z,\mathbf{z}_{-i})$. This is the potential $d$-radius circle average around point $i$, given that point $i$ is assigned to treatment condition $z$ and marginalizing over possible assignments to other points. The quantity $\bar{\mu}_i(z;d,\alpha)$ is simply the circle average of the $Y_{ix}(z;\alpha)$ values, given that $\Pr(\mathbf{Z}_{-i}=\mathbf{z}_{-i}|Z_i=z,\alpha)$ is constant over $x$.

We can now define an *individualistic response* at point $x$ of intervening on point $i$, allowing other points to vary as they otherwise would under $\alpha$:

$$\tau_{ix}(\alpha) = Y_{ix}(1;\alpha) - Y_{ix}(0,\alpha).$$

This defines the response at point $x$ of switching point $i$ from no treatment to active treatment, averaging over possible treatment assignments to the points other than $i$. At the level

of circle averages, we can define

$$\tau_i(d;\alpha) = \bar{\mu}_i(1;d,\alpha) - \bar{\mu}_i(0;d,\alpha),$$

which is the average of individualistic responses for points along the circle of radius $d$ around point $i$. Using Figure 3 to illustrate, one would construct $\tau_1(d;\alpha)$ by working with the $d$-radius circle averages around intervention point 1, taking the difference between the mean of the circle averages under assignments Z1, Z6, Z7, Z8, Z12, Z13, Z14, and Z16 minus the mean of circle averages under assignments Z1, Z3, Z4, Z5, Z9, Z10, Z11, and Z15.

Finally, define the *marginalized individualistic response* (MIR) for distance $d$ by taking the mean over the intervention points:

$$\tau(d;\alpha) = \frac{1}{N}\sum_{i=1}^{N}\left[\bar{\mu}_i(1;d,\alpha) - \bar{\mu}_i(0;d,\alpha)\right].$$

The interpretation of the MIR for distance $d$ is the average effect of switching a point at distance $d$ from no treatment to active treatment, marginalizing over any effects emanating from treatments that could be assigned to other points under the experimental design.

Our analysis focuses on experimental designs that use Bernoulli assignment, in which case the $\mathcal{Z}$ consists of the $2^N$ possible vectors that could be obtained from $N$ (possibly differentially weighted) coin flips. This allows for a relatively clean definition of causal effects. This is because Bernoulli assignment ensuress that $(1, \mathbf{z}_{-i})$ and $(0, \mathbf{z}_{-i})$ each have positive probability of occurring. In this case, the marginal quantities $Y_{ix}(1;\alpha)$ and $Y_{ix}(0;\alpha)$ are defined by marginalizing over the same sets of $\mathbf{z}_{-i}$ values, and the individualistic response has a clear *ceteris paribus* interpretation with respect to variation in the treatment assignment at intervention point $i$. Things are different under complete random assignment, where a fixed number $N_1$ of points are assigned to treatment. Then, for $Y_{ix}(1;\alpha)$, one marginalizes over assignments with $N-1$ units assigned to treatment, while for $Y_{ix}(0;\alpha)$, one marginalizes over

assignments with $N$ units assigned to treatments. As $N$ grows, these differences between Bernoulli and complete random assignment lessen. As such, at least under the regularity conditions that we propose below, for large $N$ one could consider a design that uses complete random assignment as an approximation to the cleaner Bernouilli case.

# 4 Identification, estimation, and inference

In this section we derive our key theoretical results. All proofs are contained in the appendix. We begin with the following assumptions:

**C1.** (Bernoulli design.) $Z_i \sim \text{Bernoulli}(p)$ for $i = 1, ..., N$.

**C2.** (Bounded potential outcomes.) $|Y_x(\mathbf{z})| < b$ for some finite real constant $b$ and all $x \in \mathcal{X}$ and $\mathbf{z} \in \mathcal{Z}$.

**C3.** (Local interference.) For all pairs of intervention nodes $i$ and $j$ in $\mathcal{N}$ with locations $x(i)$ and $x(j)$ in $\mathcal{X}$, and for distances $d$ in the interval $[0, \bar{d}]$, there exists a constant $h(d)$ such that if $||x(i) - x(j)|| - d > h(d)$, then $\text{Cov}\left[\mu_i(Y(z, \mathbf{Z}_{-i}); d), \mu_j(Y(z', \mathbf{Z}_{-j}); d)\right] = 0$ for all $z, z'$.

Condition C1 defines the experimental design. As discussed above, condition C1 ensures that individualistic responses are *ceteris paribus* for variation in treatment assignment at a given node. We work with the assumption that the assignment probability, $p$, is constant over intervention nodes, although extending this to cases where assignment probabilities could be done by working through suitable application of inverse probability weights. Condition C1 has the following implications for a function $f(\cdot)$.

**C1.1** $E_{\mathbf{Z}}\left[Z_i^k f(\mathbf{Z})\right] = pE_{\mathbf{Z}_{-i}}\left[f(1, \mathbf{Z}_{-i})\right]$ for any integer $k$.

**C1.2** $E_{\mathbf{Z}}\left[Z_i^k Z_j^l f(\mathbf{Z})\right] = p^2 E_{\mathbf{Z}_{-(i,j)}}\left[f(1, 1, \mathbf{Z}_{-(i,j)})\right]$ for any integer $k$ and $l$.

Moreover, condition C1 implies that treatment is independent of potential outcomes:

$$Y_x(\mathbf{Z}) \perp\!\!\!\perp \mathbf{Z}. \text{ for any } x \in \mathcal{X}.$$

Conditions C2 and C3 are the regularity conditions on outcomes that allow us to define reliable inferential procedures below. Condition C2 is a regularity condition on the potential outcomes that typically holds for real-world data. It also ensures boundedness of higher order moments for the distribution of functions of the potential outcomes. The bound $b$ affects rates of asymptotic convergence. Condition C3 means that there are limits to the spatial extent of the interference. For intervention point $j$ to satisfy C3 with respect to $i$, it would require that the circle average at $i$ is unaffected (at least to the extent that is consequential for the covariance) by not only $j$'s treatment value but also the treatment values at any intervention points that affect the circle average at $j$. In other words, $i$ and $j$ cannot share any sources of variation in their circle averages. Condition C3 allows us to define an asymptotic growth process in which the extent of independent information also increases, as stated below in Proposition 2. The upper bound of the interval $\bar{d}$ defines the largest circle mean radius for which this non-dependence condition might hold.

Our first two results show that a $d$-radius MIR admits unbiased estimator under C1 and that under C1-C3, a simple non-parametric Horvitz-Thompson estimator is consistent for $\tau(d;\alpha)$.

**Proposition 1** (Identification)**.** *Under C1,*

$$\tau(d;\alpha) = \mathrm{E}_{\mathbf{Z}}\left[\frac{1}{Np}\sum_{i=1}^{N} Z_i\mu_i(\mathbf{Y};d) - \frac{1}{N(1-p)}\sum_{i=1}^{N}(1-Z_i)\mu_i(\mathbf{Y};d)\right].$$

The terms inside the expectations on the right-hand side consist of either known design parameters ($N$ and $p$) or observable quantities. We can thus define an unbiased non-parametric

Horvitz-Thompson estimator for the $d$-radius MIR as follows:

$$\widehat{\tau}(d) = \frac{1}{Np}\sum_{i=1}^{N} Z_i\mu_i(\mathbf{Y};d) - \frac{1}{N(1-p)}\sum_{i=1}^{N}(1-Z_i)\mu_i(\mathbf{Y};d).$$

We also define two terms that count the number of dependent circle means for each intervention point, based on condition C3:

$$c_i(d) = |\{j : ||x_i - x_j|| - d \leq h(d)\}|, \text{ and } c(d) = \max_i c_i(d).$$

In the analysis that follows, we consider the convergence properties of

$$\sqrt{N}(\widehat{\tau}(d) - \tau(d;\alpha)) = \sqrt{N}\left(\frac{1}{Np}\sum_{i=1}^{N} Z_i\mu_i(\mathbf{Y};d) - \frac{1}{N(1-p)}\sum_{i=1}^{N}(1-Z_i)\mu_i(\mathbf{Y};d) - \tau(d;\alpha)\right).$$

By proposition 1 this has mean zero. The variance is equal to,

$$\begin{aligned} N\text{Var}&\left[\frac{1}{Np}\sum_{i=1}^{N} Z_i\mu_i(\mathbf{Y};d) - \frac{1}{N(1-p)}\sum_{i=1}^{N}(1-Z_i)\mu_i(\mathbf{Y};d)\right] \\ &= N\text{Var}\left[\frac{1}{Np}\sum_{i=1}^{N} Z_i\mu_i(\mathbf{Y};d)\right] \\ &\quad + N\text{Var}\left[\frac{1}{N(1-p)}\sum_{i=1}^{N}(1-Z_i)\mu_i(\mathbf{Y};d)\right] \\ &\quad - 2N\text{Cov}\left[\frac{1}{Np}\sum_{i=1}^{N} Z_i\mu_i(\mathbf{Y};d), \frac{1}{N(1-p)}\sum_{i=1}^{N}(1-Z_i)\mu_i(\mathbf{Y};d)\right]. \end{aligned} \tag{1}$$

The next result establishes sufficient conditions for the consistency of the non-parametric estimator.

**Proposition 2** (Consistency)**.** *Suppose C1-C3 and that as* $N \to \infty$, $\frac{c(d)}{N} \to 0$. *Then, as* $N \to \infty$, $\widehat{\tau}(d) - \tau(d;\alpha) \xrightarrow{p} 0$.

The condition $\frac{c(d)}{N} \to 0$ means that for each intervention point, the share of other intervention points with dependent $d$-radius circle averages goes to zero. By C3, this requires that the asymptotic growth process for the experimental design adds new intervention points that are at increased distance from those already in the sample. Moreover, the condition C3 allows that the value of $h(d)$, which defines the boundary separating dependent and independent circle averages, may vary in $d$.

In the appendix, we offer an explicit characterization of the variance in terms of the fixed potential circle averages, which would allow one to define an unbiased Horvitz-Thompson estimator for the variance as well. That said, we can also use a duality that exists in this estimation problem to re-express the estimator $\widehat{\tau}(d)$ as a regression estimator. This allows us to recast the problem of statistical inference in terms of a spatial regression. The relevant duality here is that the mean of the circle averages can be re-expressed in terms of a regression that works with the observations at each observation point, $x$. This is based on the idea that we use a finite approximation to evaluate the circle average integrals. That is, we work with

$$\mu_i(\mathbf{Y}(\mathbf{z}); d) \approx \frac{\sum_{x \in \tilde{\mathcal{X}}} Y_x 1(d_i(x) = d)}{\sum_{x \in \tilde{\mathcal{X}}} 1(d_i(x) = d)},$$

where $\tilde{\mathcal{X}}$ is a finite set of outcome evaluation points in $\mathcal{X}$ and the approximation error can be made arbitrarily small by choosing a more granular set of evaluation points.

**Proposition 3.** *Define the* $|\tilde{\mathcal{X}}| \times 1$ *vector,*

$$\mathbf{W}_d = \left( \frac{1}{N} \sum_{i=1}^{N} \frac{Z_i - p}{p(1-p)} \frac{1(d_i(x) = d)}{\sum_{x \in \tilde{\mathcal{X}}} 1(d_i(x) = d)} \right)_{x \in \tilde{\mathcal{X}}}$$

*Then,*

$$\widehat{\tilde{\tau}}(d) = \mathbf{W}_d' \mathbf{Y}.$$

This representation also allows us to define a regression through which we can estimate

a series of MIRs for a series of $d$ values in one step. For distance values $d = d_1, ..., d_D$ and outcome evaluation points indexed by $x = 1, ..., |\mathcal{X}|$,

$$\hat{\tau} = \begin{pmatrix} \sum_{i=1}^{N} \frac{\frac{Z_i - p}{p(1-p)} 1(d_i(1)=d_1)}{\sum_{x \in \mathcal{X}} 1(d_i(1)=d_1)} & \sum_{i=1}^{N} \frac{\frac{Z_i - p}{p(1-p)} 1(d_i(2)=d_1)}{\sum_{x \in \mathcal{X}} 1(d_i(2)=d_1)}, & \cdots & \sum_{i=1}^{N} \frac{\frac{Z_i - p}{p(1-p)} 1(d_i(|\mathcal{X}|)=d_1)}{\sum_{x \in \mathcal{X}} 1(d_i(|\mathcal{X}|)=d_1)} \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{i=1}^{N} \frac{\frac{Z_i - p}{p(1-p)} 1(d_i(1)=d_D)}{\sum_{x \in \mathcal{X}} 1(d_i(1)=d_D)} & \sum_{i=1}^{N} \frac{\frac{Z_i - p}{p(1-p)} 1(d_i(2)=d_D)}{\sum_{x \in \mathcal{X}} 1(d_i(2)=d_D)} & \cdots & \sum_{i=1}^{N} \frac{\frac{Z_i - p}{p(1-p)} 1(d_i(|\mathcal{X}|)=d_D)}{\sum_{x \in \mathcal{X}} 1(d_i(|\mathcal{X}|)=d_D)} \end{pmatrix} \mathbf{Y}$$
$$= \begin{pmatrix} \mathbf{W}_1' \\ \vdots \\ \mathbf{W}_D' \end{pmatrix} \mathbf{Y} \equiv \mathbf{W}\mathbf{Y},$$

where $\hat{\tau} = (\hat{\tau}(d_1)..., \hat{\tau}(d_D))'$. Moreover, defining $\mathbf{X} = \mathbf{W}'(\mathbf{W}\mathbf{W}')^{-1}$, we have, $\hat{\tau} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y}$. Under condition C3, the $\mathbf{X}$ and $\mathbf{Y}$ terms exhibit bounded spatial dependence. For inference, we can thus apply the spatial heteroskedasticity and autocorrelation-robust (spatial HAC) variance estimator due to Conley (1999) to estimate the design-induced variance $\hat{\tau}$, with the kernel boundary for the estimator set to be at least length $\max_d h(d)$.

# 5 Extensions

## 5.1 Smoothing

The Horvitz-Thompson estimator proposed in the previous section is nonparametric and makes no assumption on model specification. But it is often not unreasonable to assume that the effect curve is a smooth function of the distance to intervention points. In other words, $\tau(d)$ could be seen as continuous on $[d_1, d_D]$ and estimated flexibly using methods like local polynomial regression.

Following the approach exploited in Hainmueller et al. (2019), we minimize the loss function below:

$$L = \sum_{d'=d_1}^{d_D} \left[Y_{ixd} - \alpha_{d'} - \tau_{d'} Z_i - \beta_{d'}(d - d') - \delta_{d'} Z_i * (d - d')\right]^2 K\left(\frac{d - d'}{h}\right)$$

where $d' \in [d_1, d_2, \ldots, d_D]$ indicates the evaluation points that are set to be all the unique distance values; $Y_{ixd}$ is the outcome variable's value at point $x$ when we calculate the circle average at distance $d$ for intervention point $i$; $Z_i$ is intervention node $i$'s treatment status, and $K(.)$ is the chosen kernel function. The bandwidth is selected via cross-validation. The estimated $\hat{\tau}_d$ unveils how the treatment effect varies smoothly with the distance $d$. We can calculate the standard error and the corresponding confidence interval of $\hat{\tau}_d$ at each point as in the regression representation.

## 5.2 Permutation test

In section 4, we have demonstrated our regression approach to make statistical inference of the Horvitz-Thompson estimator. This approach offers a convenient way to construct pointwise confidence interval for each $\hat{\tau}(d)$. Nevertheless, it is less helpful when researchers are interested in whether the effect is statistically significant on a particular interval rather than at some point. For this purpose, we propose to use Fisher's permutation test to get the confidence interval for whatever statistic of interest. Under the sharp null hypothesis, we know the full distribution of potential outcomes, i.e. $Y_x(\mathbf{z})$ for any $\mathbf{z}$.

Denote the statistic of interest as $T(\mathbf{Y}, \mathbf{Z})$. Examples include each $\tau(d)$ and the average of $\tau(d)$ on interval $[d_1, d_D]$. As all the potential outcomes are known, we can redraw the assignment $\mathbf{z}$ for $P$ times and calculate the corresponding $T(\mathbf{Y}, \mathbf{Z}_p)$. The empirical distribution of $T(\mathbf{Y}, \mathbf{Z}_p)$ will approximate the distribution of $T(\mathbf{Y}, \mathbf{Z})$ and allows us to determine its statistical significance. $T(\mathbf{Y}, \mathbf{Z})$ is statistically significant if $\frac{1}{P}\sum_p \mathbf{1}\{T(\mathbf{Y}, \mathbf{Z}_p) \geq T(\mathbf{Y}, \mathbf{Z})\} \leq \alpha$, where $\alpha$ is the selected significance level.

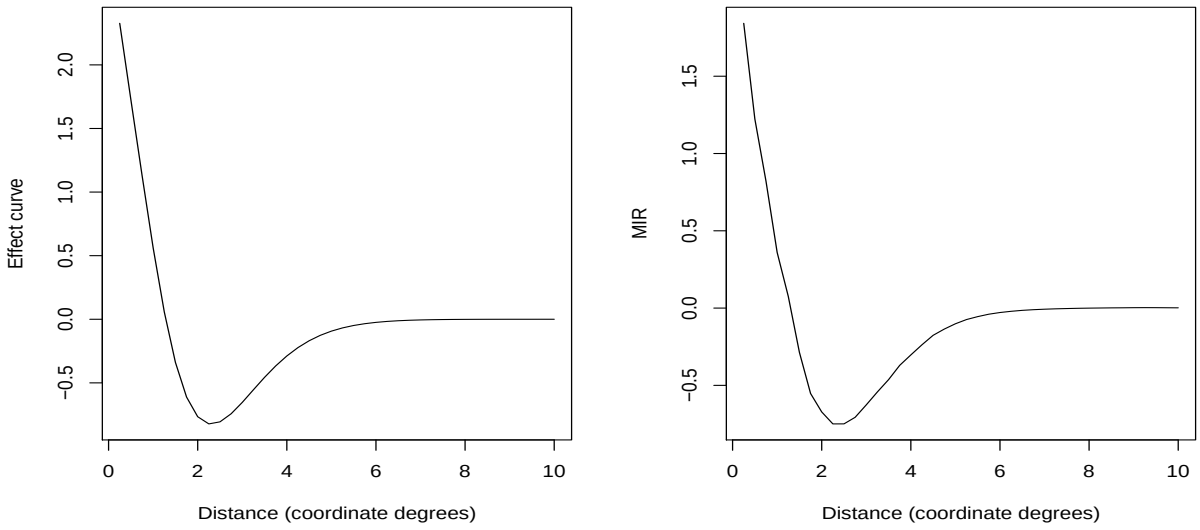


Figure 4: The comparison of the effect curve (left) and the true MIR curve (right).

# 6 Simulation

In this section, we use simulated datasets to illustrate propositions introduced in the previous section and examine the estimator's performance. For the purposes of illustration, we exploit the hypothetical non-monotonic effect function introduced in section 3, which is constructed by mixing two gamma-distribution kernels. The datasets' structure approximates our toy example in Figure 1 and Figure 3, but has more outcome points and intervention points.

To get the true MIR, we marginalize over all of the ways that treatment could be applied. We first calculate $\tau_{ix}(\alpha)$ for each pair of $(i, x)$, the expectation of the effect at outcome node $x$ induced by the treatment status' change at intervention node $i$, where the expectation is taken over the treatment status at other $N-1$ intervention nodes. Taking average of all the $\tau_{ix}(\alpha)$ at distance $d$ renders us $\tau(d)$ on the true MIR curve. The comparison between the effect curve and the true MIR curve is shown in Figure 4. The two seem very similar to each other. Actually, we prove in the appendix that the two curves are the same under Bernoulli randomization when the effect is additive.

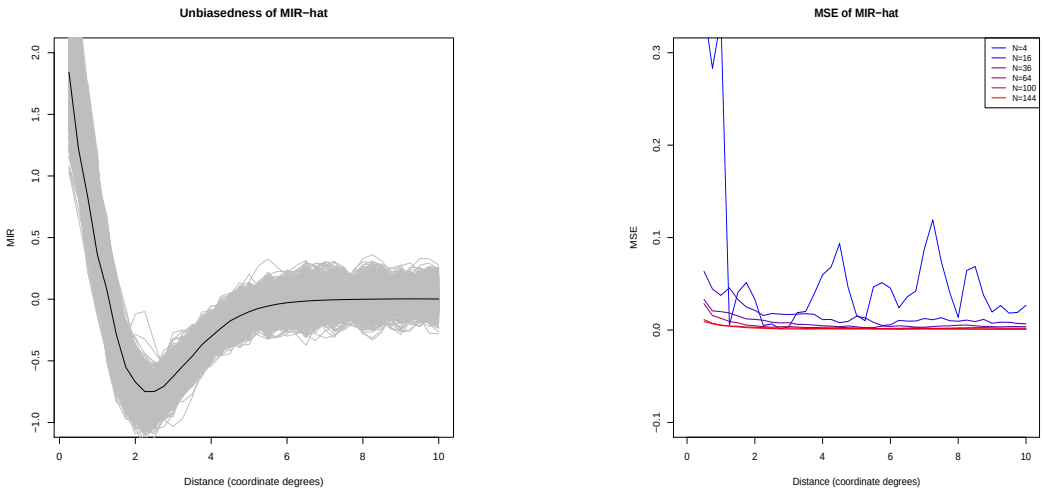


Figure 5: The unbiasedness (left) and consistency (right) of the Horvitz-Thompson estimator. The black curve on the left figure below indicates the true MIR curve and each grey curve represents the estimate under one assignment using the estimator. Each curve on the right figure shows the estimate's MSE at each distance.

We first verify the unbiasedness of the Horvitz-Thompson estimator via repeated assignments. The dataset examined has 1600 outcome points and 16 intervention points. It is clear from the left plot in Figure 5 that the true MIR curve resides in the middle of all the MIR estimates from repeated assignments. The right plot in Figure 5 indicates that as samle size (the number of intervention points) rises[1], the MSE of MIR estimates converge to zero at all distance values.

We then illustrate the two proposed approaches for statistical inference, Spatial HAC (Conley) standard errors based on regression representation and permutation test. In the first row of Figure 6, estimates from one particular assignment are shown along with their 95% confidence interval constructed from Conley standard errors and the 95% confidence interval under the sharp null distribution. Both approaches indicate that the effect is significantly positive near 0 and significantly negative around 2. The left plot does not use smoothing while the right one does. In the second row of Figure 6, we compare the aver-

[1]The number of outcome points increase proportionally

age converage of the 95% confidence interval constructed from Conley standard errors over repeated assignments with the real 95% confidence interval. We can see that regression representation returns a slightly conservative estimate of the uncertainty's magnitude.

# 7 Examples

We demonstrate how to apply our method in empirical studies using two examples. The first example comes from Paler et al. (2015) in which the authors conduct a field experiment in Aceh, Indonesia by randomly setting up monitoring stations to 28 villages. There are 14 treated villages and 14 control ones. The outcome of interest is forest coverage in the district (ranging from 0 to 1) and the dataset takes the form of a raster object. Each tile represents a grid on the map, as presented in the left plot of Figure 7. As rangers from the stations patrol around each village, we expect their existence affects forest coverage in nearby areas. The result using our method is presented in the right plot of Figure 7. We can see that although the estimated MIR curve takes positive values near 0.2, the effect is not statistically significant at any distance. Result from the permutation test suggests that the average effect is also insignificant on the interval $[0.2, 0.4]$.

The second example is based on Miguel and Kremer (2004). In this study, pupils from 25 randomly picked schools out of 50 in a Kenyan area received anti-worm treatment, as illustrated in the left plot of Figure 8. We know the average effect of the treatment at school level and we want to understand its diffusion effect in the whole area. For this purpose, kriging is first used to extrapolate the outcome variable's value to each spatial point. The MIR is then estimated using these predicted values. From the right plot of Figure 8 we can see that the effect is significantly negative around zero based on the permuation test, but not significant under the 95% confidence intervals constructed from Conley standard

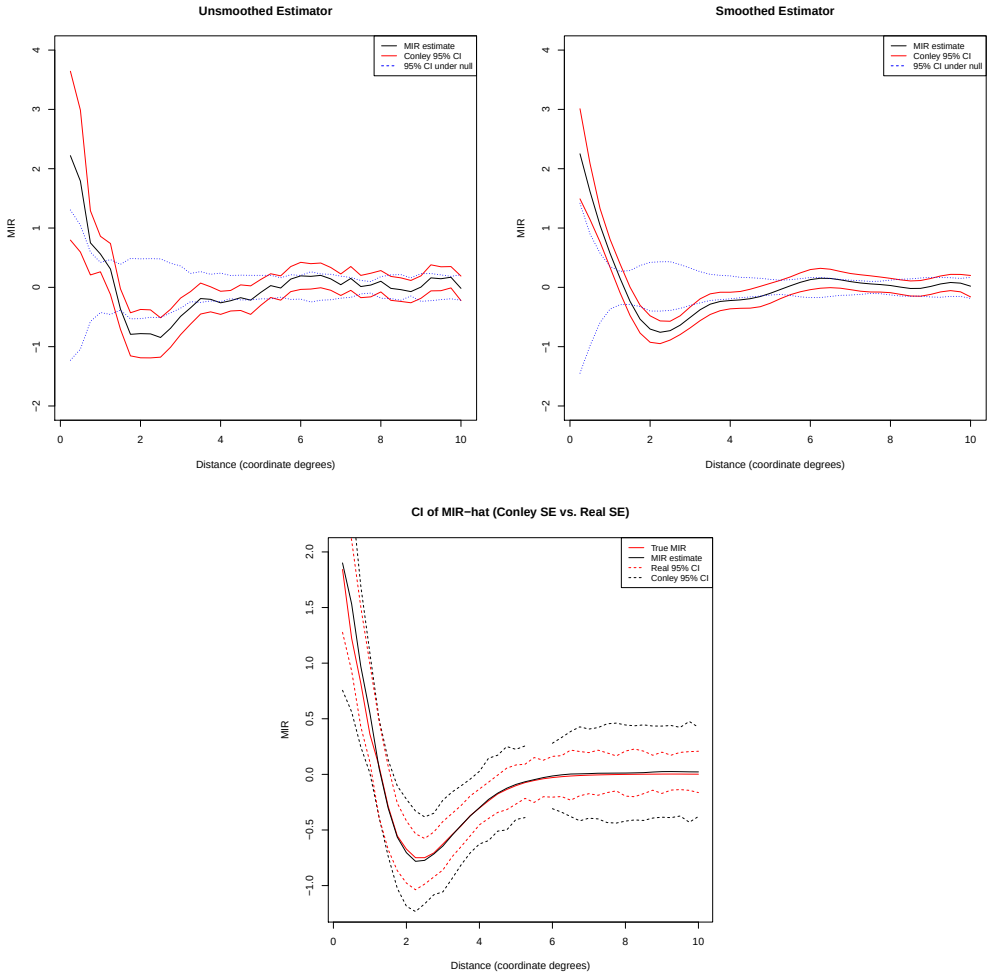


Figure 6: The two plots on the top show the estimate and its statistical significance from one assignment. The left one uses the nonparametric Horvitz-Thompson estimator and the right one exploits local regression. The black curve represents the MIR estimate. The red curves are 95% confidence intervals constructed from Conley standard errors. The blue lines are the 95% confidence intervals under sharp null. On the bottom the red curves are the real 95% confidence intervals and the dotted curves are based on Conley SE.

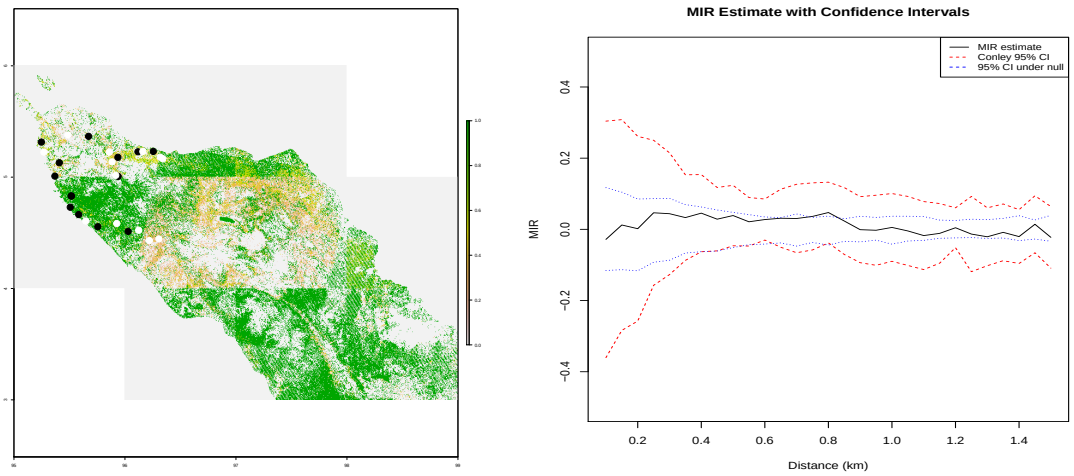


Figure 7: The plot on the left demonstrates both the treatment status and outcome values in the experiment. White points are villages under control and black points are treated villages. Greener colors indicate higher forest coverage rate. The plot on the right presents results using our method. The black curve represents the MIR estimate. The red curves are 95% confidence intervals constructed from Conley standard errors. The blue lines are the 95% confidence intervals under sharp null.

errors. Hence there is some evidence that the anti-worm effect diffuses to households nears the schools.

# 8 Conclusion

Interference is a common phenomenon in spatial experiments. In practice, researchers have attempted various ways to identify the diffusion effect of their treatment. Yet, rigrous analysis under the Neyman-Rubin framework has not be established in extant studies. This paper constructs a well-defined causal quantity— the marginalized individualistic response (MIR)— and proposes a Horvitz-Thompson estimator for its estimation. We show that the estimator is unbiased and consistent under regularity conditions, and provide both the Neyman variance and Fisher permutation test for inference. Evidence from simulations and empirical studies suggest that the method behaves well and can be helpful in different

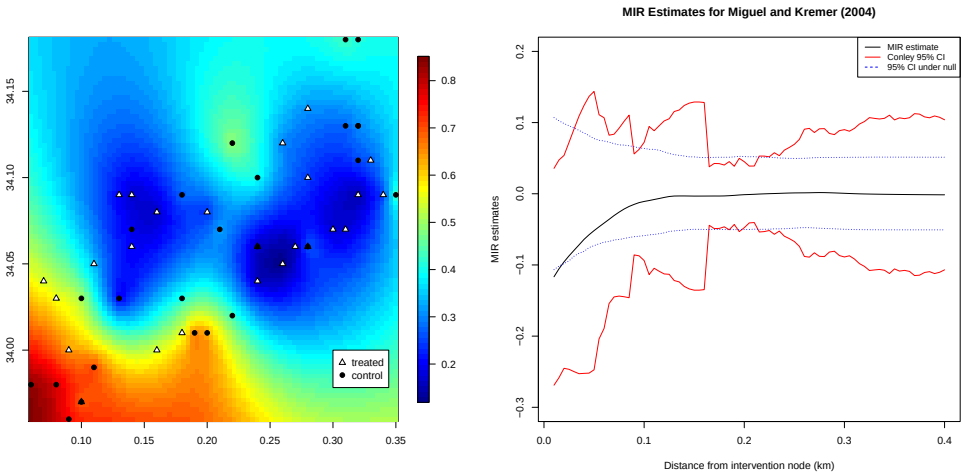


Figure 8: The plot on the left demonstrates both the treatment status and predicted outcome values from kriging in the experiment. White triangles are schools under control and black circle are treated schools. The color on the map indicates the infection rate. The plot on the right presents results using our method. The black curve represents the MIR estimate. The red curves are 95% confidence intervals constructed from Conley standard errors. The blue lines are the 95% confidence intervals under sharp null.

research settings.

# A Proofs

## A.1 Proposition 1

*Proof.*

$$
\begin{aligned}
\tau(d;\alpha) &= \frac{1}{N}\sum_{i=1}^{N}\bar{\mu}_i(1;d,\alpha) - \frac{1}{N}\sum_{i=1}^{N}\bar{\mu}_i(0,d;\alpha) \\
&= \frac{1}{Np}\sum_{i=1}^{N}p\bar{\mu}_i(1;d,\alpha) - \frac{1}{N(1-p)}\sum_{i=1}^{N}(1-p)\bar{\mu}_i(0;d,\alpha) \\
&= \frac{1}{Np}\sum_{i=1}^{N}p\mathrm{E}_{\mathbf{Z}_{-i}}[\mu_i(\mathbf{Y}(1,\mathbf{Z}_{-i});d)] - \frac{1}{N(1-p)}\sum_{i=1}^{N}(1-p)\mathrm{E}_{\mathbf{Z}_{-i}}[\mu_i(\mathbf{Y}(0,\mathbf{Z}_{-i});d)] \\
&= \frac{1}{Np}\sum_{i=1}^{N}\mathrm{E}_{\mathbf{Z}}[Z_i\mu_i(\mathbf{Y}(1,\mathbf{Z}_{-i});d)] - \frac{1}{N(1-p)}\sum_{i=1}^{N}\mathrm{E}_{\mathbf{Z}}[(1-Z_i)\mu_i(\mathbf{Y}(0,\mathbf{Z}_{-i});d)] \\
&= \frac{1}{Np}\sum_{i=1}^{N}\mathrm{E}_{\mathbf{Z}}[Z_i\mu_i(\mathbf{Y};d)] - \frac{1}{N(1-p)}\sum_{i=1}^{N}\mathrm{E}_{\mathbf{Z}}[(1-Z_i)\mu_i(\mathbf{Y};d)] \\
&= \mathrm{E}_{\mathbf{Z}}\left[\frac{1}{Np}\sum_{i=1}^{N}Z_i\mu_i(\mathbf{Y};d) - \frac{1}{N(1-p)}\sum_{i=1}^{N}(1-Z_i)\mu_i(\mathbf{Y};d)\right],
\end{aligned}
$$

where the fourth equality follows from implication C1.1 and the fifth equality follows from the definition of the potential outcomes and corresponding circle averages. □

## A.2 Proposition 2

*Proof.* Considering expression 1, for the first term we have,

$$\begin{aligned} N\text{Var}\left[\frac{1}{Np}\sum_{i=1}^{N} Z_i\mu_i(\mathbf{Y};d)\right] &= \frac{1}{Np^2}\text{Var}\left[\sum_{i=1}^{N} Z_i\mu_i(\mathbf{Y};d)\right] \\ &= \frac{1}{Np^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\text{Cov}\left[Z_i\mu_i(\mathbf{Y};d), Z_j\mu_j(\mathbf{Y};d)\right] \\ &= \frac{1}{Np^2}\sum_{i=1}^{N}\sum_{j=1}^{N}\text{Cov}\left[Z_i\mu_i(\mathbf{Y}(1,\mathbf{Z}_{-i});d), Z_j\mu_j(\mathbf{Y}(1,\mathbf{Z}_{-j});d)\right] \\ &\leq \frac{1}{Np^2}Nc(d)k = \frac{c(d)k}{p^2}, \end{aligned}$$

for $k > 0$. By C2, $k$ is finite and bounded, and by C3, c(d) is finite and bounded. By the same logic,

$$N\text{Var}\left[\frac{1}{Np}\sum_{i=1}^{N} Z_i\mu_i(\mathbf{Y};d)\right] \leq \frac{c(d)k}{(1-p)^2}.$$

Finally,

$$\begin{aligned} -2N\text{Cov}&\left[\frac{1}{Np}\sum_{i=1}^{N} Z_i\mu_i(\mathbf{Y};d), \frac{1}{N(1-p)}\sum_{i=1}^{N}(1-Z_i)\mu_i(Y;d)\right] \\ &= -\frac{2}{Np(1-p)}\sum_{i=1}^{N}\text{Cov}\left[Z_i\mu_i(\mathbf{Y}(1,\mathbf{Z}_{-i});d), (1-Z_i)\mu_i(\mathbf{Y}(0,\mathbf{Z}_{-i});d)\right] \\ &\quad -\frac{2}{Np(1-p)}\sum_{i=1}^{N}\sum_{j\neq i}\text{Cov}\left[Z_i\mu_i(\mathbf{Y}(1,\mathbf{Z}_{-i});d), (1-Z_j)\mu_j(\mathbf{Y}(0,\mathbf{Z}_{-j});d)\right] \\ &\leq \frac{2k(1+c(d))}{p(1-p)} \end{aligned}$$

by C2 and C3. Then, the condition $\frac{c(d)}{N} \to 0$ implies that that as $N \to \infty$, $\text{Var}\left[\widehat{\tau}(d) - \tau(d;\alpha)\right] \to 0$. Combining these results with the unbiasedness result from Proposition 1 and Chebychev's inequality we have that $\widehat{\tau}(d) - \tau(d;\alpha) \xrightarrow{p} 0$. □

## A.3 Proposition 3

$$
\begin{aligned}
\hat{\tau}(d) &= \frac{1}{Np}\sum_{i=1}^{N} Z_i \mu_i(\mathbf{Y}; d) - \frac{1}{N(1-p)}\sum_{i=1}^{N}(1-Z_i)\mu_i(\mathbf{Y}; d) \\
&= \frac{1}{N}\sum_{i=1}^{N}\frac{Z_i - p}{p(1-p)}\mu_i(\mathbf{Y}; d) \\
&= \frac{1}{N}\sum_{i=1}^{N}\frac{Z_i - p}{p(1-p)}\frac{\sum_{x\in\tilde{\mathcal{X}}} Y_x 1(d_i(x) = d)}{\sum_{x\in\tilde{\mathcal{X}}} 1(d_i(x) = d)} \\
&= \sum_{x\in\tilde{\mathcal{X}}} Y_x \frac{1}{N}\sum_{i=1}^{N}\frac{Z_i - p}{p(1-p)}\frac{1(d_i(x) = d)}{\sum_{x\in\tilde{\mathcal{X}}} 1(d_i(x) = d)}
\end{aligned}
$$

## A.4 The equivalence between effect function and the MIR curve

Suppose that for each outcome node $x$, its outcome value is generated additively: $h(x) = \sum_i^N Z_i g_i(x) + f(x)$, where $f(x)$ captures spatial trends in the absence of any intervention, and then $g_i(x)$ captures effects that emanate, perhaps idiosyncratically, from each of the intervention nodes. Note that this model captures more restrictive models of homogenous spatial effects. Then we have,

$$\begin{aligned}
\tau_{ix} =& E_{\mathbf{Z}_{-\mathbf{i}}|Z_i=1}\left[h(x)|Z_i=1\right] - E_{\mathbf{Z}_{-\mathbf{i}}|Z_i=0}\left[h(x)|Z_i=0\right] \\
=& E_{\mathbf{Z}_{-\mathbf{i}}}\left[g_i(x) + \sum_{j\neq i}^{N} Z_j g_j(x) + f(x)\right] - E_{\mathbf{Z}_{-\mathbf{i}}}\left[\sum_{j\neq i}^{N} Z_j g_j(x) + f(x)\right] \\
=& g_i(x)
\end{aligned}$$

in which case the MIR estimates the average of the accumulated effects emanating from the $N$ intervention points.

# B Expressing the variance in terms of potential circle averages

We develop the derivation for the component in expression 1 that corresponds to the variance of the treated intervention points' circle averages. Results for the control intervention points and then the covariance of the two would follow. First, define $\mathcal{B}(i;d)$ as the set of intervention points whose treatment statuses affect the potential circle averages for intervention point $i$. (Note that intervention points outside of $\mathcal{B}(i;d)$ do not necessarily satisfy condition C3 with respect to $i$.) Let $\mathbf{z}_{\mathcal{B}(i;d)}$ and $\mathcal{Z}_{\mathcal{B}(i;d)}$ correspond to realized and feasible ordered treatment vectors for units in $\mathcal{B}(i;d)$. Then, let $I_j(\mathbf{z}_{\mathcal{B}(i;d)})$ be an indicator variable for whether the treatment indicator at intervention point $j$ equal's point $j$'s entry in the vector $\mathbf{z}_{\mathcal{B}(i;d)}$. With these terms, we can rewrite the expression for the total of treated circle averages in terms of random selection indicators and then fixed potential circle averages:

$$\begin{aligned}
\sum_{i=1}^{N} Z_i \mu_i(\mathbf{Y};d) &= \sum_{i=1}^{N} Z_i \sum_{\mathbf{z}_{-i}\in\mathcal{Z}_{-i}} I(\mathbf{Z}_{-i}=\mathbf{z}_{-i})\mu_i(\mathbf{Y}(1,\mathbf{z}_{-i});d) \\
&= \sum_{i=1}^{N} Z_i \sum_{\mathbf{z}_{\mathcal{B}(i;d)}\in\mathcal{Z}_{\mathcal{B}(i;d)}} \prod_{j\in\mathcal{B}(i;d)} I_j(\mathbf{z}_{\mathcal{B}(i;d)})\mu_i(\mathbf{Y}(1,\mathbf{z}_{\mathcal{B}(i;d)});d) \\
&\equiv \sum_{i=1}^{N} T_i.
\end{aligned}$$

In the expression above, only $Z_i$ and $I_j(\mathbf{z}_{\mathcal{B}(i;d)})$ are random, while $\mu_i(\mathbf{Y}(1,\mathbf{z}_{\mathcal{B}(i;d)});d)$ is a fixed potential circle average. Then we have,

$$\text{Var}\left[\sum_{i=1}^{N} T_i\right] = \sum_{i=1}^{N} \text{Var}\left[T_i\right] + \sum_{i=1}^{N}\sum_{j\neq i} \text{Cov}\left[T_i, T_j\right].$$

By Bernoulli assignment, we have that $\text{Cov}\left[T_i, T_j\right] = 0$. (The logic is a straightforward extension of Frank, 1977.) Then, applying the expression for the variance of a product of

independent variables due to Goodman (1962), we have,

$$
\begin{aligned}
\text{Var}\left[\sum_{i=1}^{N} T_i\right] &= \sum_{i=1}^{N} \text{Var}\left[Z_i \sum_{\mathbf{z}_{\mathcal{B}(i;d)} \in \mathcal{Z}_{\mathcal{B}(i;d)}} \prod_{j \in \mathcal{B}(i;d)} I_j(\mathbf{z}_{\mathcal{B}(i;d)}) \mu_i(\mathbf{Y}(1, \mathbf{z}_{\mathcal{B}(i;d)}); d)\right] \\
&= \sum_{i=1}^{N} \Bigg\{ \sum_{\mathbf{z}_{\mathcal{B}(i;d)} \in \mathcal{Z}_{\mathcal{B}(i;d)}} \text{Var}\left[Z_i \prod_{j \in \mathcal{B}(i;d)} I_j(\mathbf{z}_{\mathcal{B}(i;d)})\right] \mu_i(\mathbf{Y}(1, \mathbf{z}_{\mathcal{B}(i;d)}); d)^2 \\
&\quad + \sum_{\mathbf{z}_{\mathcal{B}(i;d)} \in \mathcal{Z}_{\mathcal{B}(i;d)}} \sum_{\mathbf{z}'_{\mathcal{B}(i;d)} \neq \mathbf{z}_{\mathcal{B}(i;d)}} \text{Cov}\left[Z_i \prod_{j \in \mathcal{B}(i;d)} I_j(\mathbf{z}_{\mathcal{B}(i;d)}), Z_i \prod_{j \in \mathcal{B}(i;d)} I_j(\mathbf{z}'_{\mathcal{B}(i;d)})\right] \\
&\qquad \times \mu_i(\mathbf{Y}(1, \mathbf{z}_{\mathcal{B}(i;d)}); d) \mu_i(\mathbf{Y}(1, \mathbf{z}'_{\mathcal{B}(i;d)}); d) \Bigg\} \\
&= \sum_{i=1}^{N} \Bigg\{ \sum_{\mathbf{z}_{\mathcal{B}(i;d)}} \Bigg[ (p(1-p) + p^2) \prod_{j \in \mathcal{B}(i;d)} (p(1-p) + \text{E}\left[I_j(\mathbf{z}_{\mathcal{B}(i;d)})\right]^2) \\
&\qquad\qquad - p^2 \prod_{j \in \mathcal{B}(i;d)} \text{E}\left[I_j(\mathbf{z}_{\mathcal{B}(i;d)})\right]^2 \Bigg] \mu_i(\mathbf{Y}(1, \mathbf{z}_{\mathcal{B}(i;d)}); d)^2 \\
&\qquad\qquad - p^2 \Bigg[ \sum_{\mathbf{z}_{\mathcal{B}(i;d)} \in \mathcal{Z}_{\mathcal{B}(i;d)}} \sum_{\mathbf{z}'_{\mathcal{B}(i;d)} \neq \mathbf{z}_{\mathcal{B}(i;d)}} \prod_{j \in \mathcal{B}(i;d)} \text{E}\left[I_j(\mathbf{z}_{\mathcal{B}(i;d)})\right] \text{E}\left[I_j(\mathbf{z}'_{\mathcal{B}(i;d)})\right] \\
&\qquad\qquad\qquad \times \mu_i(\mathbf{Y}(1, \mathbf{z}_{\mathcal{B}(i;d)}); d) \mu_i(\mathbf{Y}(1, \mathbf{z}'_{\mathcal{B}(i;d)}); d) \Bigg] \Bigg\}.
\end{aligned}
$$

Given the Bernoulli design, the $\text{E}\left[I_j(\mathbf{z}_{\mathcal{B}(i;d)})\right]$ terms will equal $p$ or $1-p$ depending whether the assignment pattern in $\mathbf{z}_{\mathcal{B}(i;d)}$ requires $z_j = 1$ or $z_j = 0$.

# References


Aronow, P. M. and C. Samii (2017). Estimating average causal effects under general interference, with application to a social network experiment. *Annals of Applied Statistics 11*(4), 1912–1947.

Bowers, J., M. M. Fredrickson, and C. Panagopolous (2013). Reasoning about interference between units: A general framework. *Political Analysis 21*(1), 97–124.

Conley, T. G. (1999). GMM estimation with cross sectional dependence. *Journal of Econometrics 92*, 1–45.

Cox, D. R. (1958). *Planning of Experiments*. Wiley.

Frank, O. (1977). A note on bernoulli sampling in graphs and horvitz-thompson estimation. *Scandanavian Journal of Statistics 4*(4), 178–180.

Goodman, L. A. (1962). The variance of the product of k random variables. *Journal of the American Statistical Association 57*(297), 54–60.

Hainmueller, J., J. Mummolo, and Y. Xu (2019). How much should we trust estimates from multiplicative interaction models? simple tools to improve empirical practice. *Political Analysis 27*(2), 163–192.

Hudgens, M. G. and M. E. Halloran (2008). Toward causal inference with interference. *Journal of the American Statistical Association 103*(482), 832–842.

Manski, C. F. (2012). Identification of treatment response with social interactions. *The Econometrics Journal* (In press).

Miguel, E. and M. Kremer (2004). Worms: identifying impacts on education and health in the presence of treatment externalities. *Econometrica 72*(1), 159–217.

Paler, L., C. Samii, M. Lisiecki, and A. Morel (2015). Social and environmental impact of the community rangers program in aceh.

Rosenbaum, P. R. (2007). Interference between units in randomized experiments. *Journal of the American Statistical Association 102*(477), 191–200.

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association 100*, 322–331.

Savje, F., P. M. Aronow, and M. G. Hudgens (2018). Average treatment effects in the presence of unknown interference. *arXiv:1711.06399 [math.ST]*.

Tchetgen-Tchetgen, E. J. and T. J. VanderWeele (2010). On causal inference in the presence of interference. *Statistical Methods in Medical Research 21*(1), 55–75.

Zigler, C. M. and G. Papadogeorgou (2018). Bipartite causal inference with interference. *arXiv:1807.08660 [stat.ME]*.