Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Samantha M. Noreen Date

Quantifying the Impact of Local SUTVA Violations in Spatiotemporal Causal Models

By

Samantha M. Noreen

Doctor of Philosophy

Biostatistics

Lance A. Waller, Ph.D.
Advisor

W. Dana Flanders, Ph.D.
Committee Member

Vicki Hertzberg, Ph.D.
Committee Member

Michael Kramer, Ph.D.
Committee Member

Accepted:

Lisa A. Tedesco, Ph.D.
Dean of the James T. Laney School of Graduate Studies

Date

Quantifying the Impact of Local SUTVA Violations in Spatiotemporal Causal Models

By

Samantha M. Noreen

B.S., B.S.Ed, Miami University of Ohio, 2012

M.S., Emory University, 2015

Advisor: Lance A. Waller, Ph.D.

Committee Members:

W. Dana Flanders, Ph.D.

Vicki Hertzberg, Ph.D.

Michael Kramer, Ph.D.

An abstract of

A dissertation submitted to the Faculty of the

James T. Laney School of Graduate Studies of Emory University

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in Biostatistics

2018

Abstract

Quantifying the Impact of Local SUTVA Violations in Spatiotemporal Causal Models

By

Samantha M. Noreen

Various causal frameworks have been built upon and extended to deal with the complexities and intricacies that arise when quantifying causal relationships that may exist in natural or quasi-experimental settings. Particular care is needed when data are sensitive to or defined by spatial differences and dependencies, as fundamental causal assumptions may be violated. Interaction and heterogeneity frequently witnessed as part of spatial data settings, such as in health-related policy or program evaluation and event point data, can often violate the Stable Unit Treatment Value Assumption (SUTVA), a key assumption in the counterfactual causal framework. With such a violation, it is difficult to assess treatment effects that are at the heart of the problem of interest.

We propose a potential outcomes framework in the context of spatio-temporal point processes, developing a theoretical spatial framework that extends existing methods of causal inference with spatial point process theory. Specifically, we focus on bringing this viewpoint to event point data looking at conflict in Afghanistan, used as the illustrative setting for the matched wake analysis (MWA) approach. We reshape the causal hypothesis that indiscriminate insurgent violence using improvised explosive devices (IEDs) increases civilian handover events of unexploded ordnances to U.S. troops, compared to selective insurgent violence from a focus on counting events to quantify a causal effect, to interpreting the intensity, representing the expected events per unit volume of space-time, to quantify a local causal effect. Framing the effect of changing rates of an outcome event over time under treatment or control intervention in a stochastic point process perspective allows us to take advantage of convenient properties to inform the estimation and specification of spatio-temporal areas of influence for each unit. This impacts the definition of the units themselves in addition to estimation of an unbiased causal effect.

By bringing together two methological and computational approaches, we consider the misspecification of the radius of spatial influence that is needed to define the spatio-temporal wakes of each intervention event in the MWA approach. The assignment mechanism is based on geography, where changes can occur in one place and time but not another. As such, it is important to consider the local impact and spatial definitions of the causal relationship that is being quantified. This approach raises methodological challenges; however, we illustrate how a space-time point process stochastic framework allows novel insight as well as a theoretical basis for heuristic approaches for determining the local space-time scale of effects.

In order to address the violations of SUTVA that occur in this spatio-temporal

setting due to interference and the treatment definition based on geography, we propose a novel approach considering a space-time point process stochastic framework combined with the structure of interference with geographic features. This solution to the challenges posed by spatial interaction allows for a more in-depth examination of the underlying causal relationships of intervention efficacy. This problem is of interest due the spatial and temporal nature of the data and motivating questions of interest in the conflict data set. Geographic impact analysis accounting for selection bias, spatial dependence and spillovers, and spatial heterogeneity is becoming more and more necessary in this age of increasingly available observational, natural, and quasi-experimental data. This work contributes to an on-going conversation and area of focus that continues to grow across multiple disciplines.

Quantifying the Impact of Local SUTVA Violations in Spatiotemporal Causal Models

By

Samantha M. Noreen
B.S., B.S.Ed, Miami University of Ohio, 2012
M.S., Emory University, 2015

Advisor: Lance Waller, Ph.D.

A dissertation submitted to the Faculty of the
James T. Laney School of Graduate Studies of Emory University
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy
in Biostatistics
2018

Acknowledgment

This journey of earning a doctorate has seen many individuals both come and go, and throughout all stages these people have given me never-ending support, encouragement, and wisdom to come out on the other side. The many individuals that make up this entire experience are too many to call out, but to all I give my sincere thanks for the part that you have played in helping me reach this milestone. I truly would not have been able to achieve all that I have without each and every one of you in this experience.

A very large thank you to Lance Waller, for not only leading the department as a compassionate, engaging chair but also squeezing in the time and commitment to see me through to the finish line. Your genuine passion for learning and tackling new challenges with optimism was the inspiration that I needed to reignite my own excitement and dedication to research. Your support, dedication, and collaboration has helped me grow rapidly during our time working together.

Thank you to Vicki Hertzberg, for opening the doors and showing me a future that is now my reality. From allowing me to call you over six years ago when I was making that final decision of which graduate school to attend and listening while I was deciding between Emory and another school, to Taqueria del Sol lunches 'with Jill' - I cannot thank you enough. Your fiery spirit and continued accomplishments have been (and continue to be) the shining example of what I hope to be even a fraction of in my career. The unwaivering support you have provided since the Plum Crazy days has meant the world to me.

Thank you to my committee members Dana Flanders, Vicki Hertzberg, and Michael Kramer for their input, assistance, time, and guidance in bringing this work to fruition.

Thank you to Sebastian Schutte and Karsten Donnay, my collaborators, for their unbounded enthusiasm and willingness to begin this effort. This dissertation would not have been possible without the resources, discussions, and excitement they have provided in establishing a forward trajectory for this and continued work.

The faculty, staff and my peers in the department have been invaluable to my development as a knowledgeable statistician, and more importantly as a better person. There was always someone willing to share their technical knowledge, pass along words of wisdom, struggle through foreign problems, and commiserate over copious amounts of coffee. To even get to this point, I want to thank Mr. Bringardner. Who would have known that such a love and curiosity for mathematics and statistics would have come out of an honors geometry class that at first I thought so intimidating I cried myself to sleep every night those first few weeks after hours of homework, review, and studying. After requesting to be placed with him for subsequent classes, it is clear that my dedication and solid foundation truly started in that class and for that I am forever grateful.

For trusting that I was worth the effort and going above and beyond in offering support and guidance before it was due, thank you to Bob Carrico and Amber Wilk at United Network for Organ Sharing.

To all of the wonderful individuals that I have been lucky enough to call friends from each stage of my life and this journey, I cannot thank you enough for all of the ways you enriched my life and kept me sane. To Lauren, Emily, Gina, Kelsie, and Kendra, I could not ask for a more brilliant, strong, and magnetizing group of individuals that have become forever-friends. Thank you to Kylie, for always being a caring friend in all circumstances and sharing the experiences (the good, bad, and everything in-between) of graduate school. To Megan, my 'big sis' - I have never known a more selfless, good-hearted person and am so lucky to have met you in that first year. For the persistent and always much needed checking in since we left our college days, thank you to Carrie. Our talks and your encouragement and laughter always seem to come at the most needed times. Thank you to Emme, for giving me confidence and thinking more of me for setting out on this endeavor than I deserve.

Stephen, for being my rock, my shoulder to lean on, my cheerleader, and many, many other roles. You gracefully (and perhaps unknowingly) accepted this journey and all of the joys and heartaches that came along with it without a second thought. This would not have been remotely possible without you. Thank you. Words cannot express the love, admiration, and esteem that I have for you.

Lastly, I would like to thank my family. My sister, Sarah, for keeping me grounded and broadening my perspectives outside of the narrow focus of a life in the books. To my dad. I may have outgrown you being able to help with my math problems, but I will never outgrow your judicious, always generous, advice and guidance. You have always kept me honest, and I will always want you to read and help review my work - your opinion always gives me a sense of calm and acceptance. And to my mom - you pulled me through the knothole more times than I can count, and always gave me the push I needed. I will always look to you as an example of what I want to be capable of. To both of you, thank you cannot even begin to cover everything. It is a certainty that I would not have gotten to the completion of this without you. I am so lucky to have two such amazing, accomplished role models in my life to look up to and strive to learn from. Through the highs and lows, struggles and and lessons, tears and celebrations, and this largest celebration of them all, I owe my strength and continued success to your love and support. This accomplishment is dedicated to you.

This journey would not have been possible alone, and so many people have contributed to earning this degree. I will always be indebted to each and every one of the individuals that have helped me get here.

> *Life is not easy for any of us. But what of that? We must have perseverance and above all, confidence in ourselves. We must believe that we are gifted for something and that this thing must be obtained. – Marie Curie*

Contents

List of Figures

List of Tables

Chapter 1

Introduction and Background

1.1 Literature Review

1.1.1 Observational studies

The gold standard in treatment comparison is a randomized clinical trial design. This is not always feasible so many times we work with observational studies. In observational studies, there is no randomization of treatment so one cannot guarantee that the groups of treated and untreated are directly comparable. There may be systematic differences between the groups, not necessarily related to the treatment effect; thus, we cannot make conclusions about the cause of difference in outcomes due to the treatment effect. This phenomenon is due to (1) confounding, and (2) selection bias. Confounding manifests itself in characteristics that are correlated with both the response and exposure or other covariates, clouding the actual effect of treatment. Selection bias refers to the systematic differences that can occur in observational studies when particular kinds of patients tend to receive or not receive the treatment. The field of causal inference deals with these types of problems to allow us to make direct comparisons and determine specific treatment effects or comparisons in effects of particular covariates (Austin, 2011).

1.1.2 Causal Inference Framework

We follow the common causal inference framework, referred to as the Rubin Causal Model (Rubin, 1974; Holland, 1986). Rubin's Causal Model frames questions from the field of causal inference in a way such that methods to find unbiased estimation of treatment effects can be applied when we have non-randomized studies. This widely used and important framework identifies a causal effect of treatment through the use of potential outcomes.

Potential Outcomes

Each unit i is assumed to have a pair of potential outcomes; $Y_i(1)$ the outcome un-

der active treatment and $Y_i(0)$ outcome under control treatment. These potential outcomes are hypothetical responses, but are never observed simultaneously for an individual. One potential outcome is observed $Y_i(Z_i)$ for indicator of treatment received $Z_i = 0, 1$, where

$$Y_i = Z_i Y_i(1) + (1 - Z_i) Y_i(0)$$

while the other is called the counterfactual outcome. While this framework is convenient to define the precise statement of the questions of interest, it introduces us to the fundamental problem of casual inference (Holland, 1986). Here, t represents units in the treated group, c represents units in the control group, and u is the realized value of treatment to which the unit is exposed.

> " It is impossible to *observe* the value of $Y_t(u)$ and $Y_c(u)$ on the same unit and, therefore, it is impossible to *observe* the effect of t on u. " (Holland, 1986)

We would like to be able to say that a noticeable effect is attributable to the treatment; however, since it is impossible to observe these individual effects, we instead compare the mean response if the entire population received treatment to that if the entire population received control. This difference in means is known as the average causal treatment effect, $ACE = \mathbb{E}[Y(1) - Y(0)]$. Another often-used population-level comparison, the average treatment effect for the treated, $ATT = \mathbb{E}[Y(1) - Y(0)|Z = 1]$, considers the average effect of treatment if the entire population receives treatment. These causal effects can be estimated following some assumptions on the treatment assignment mechanism and how it relates to the outcome of interest.

Assumptions

For observational studies in particular, treatment exposure may be confounded by other factors. Subject characteristics may play a part in whether or not a subject

receives treatment and these characteristics may also be associated with the way the subject might respond under treatment or control. In order to account for the non-randomness of the treatment assignment mechanism and appropriately adjust for confounding, some untestable assumptions are still required to achieve the goal of the desired causal interpretation of the ACE. Before diving into these assumptions, a common rule of inference following from the counterfactual framework is employed called consistency (Robins, 1987). This states that the potential response of some patient, $Y_i(t)$ to a hypothetical treatment t must coincide with the patient's observed response Y_i whenever the actual treatment T happens to be t. A more formal definition following similar notation to Galles and Pearl (1998),

$$T = t \Longrightarrow Y_i(t) = Y_i \tag{1.1}$$

Following the implications of this rule, other important and useful assumptions may be made; the first of these is the stable-unit treatment value assumption (SUTVA) of Rubin (1986).

Assumption 1 (Stable-Unit Treatment Value (SUTVA))
The a priori assumption that the value of outcome variable Y_i for unit i when exposed to treatment Z_i will be the same no matter what mechanism is used to assign treatment Z_i to unit i and no matter what treatments the other units receive, this holds for all $i = 1, \ldots, N$ and all $Z_i \in T_i$.

This guarantees that there is no interference between the treatment received for one unit i and the potential outcomes of another unit j, $i \neq j$, and is essentially implying independence between all units in the study. Also by SUTVA, there is no variation in the treatments so the treatment of all units is comparable. Rubin (1986) states that "SUTVA is violated when, for example, there exist unrepresented versions of treatments ($Y_t(u)$ depends on which version of treatment t was received) or interference

between units ($Y_t(u)$ depends on whether unit u' received treatment t or t')".

Strongly ignorable treatment assignment, or "or unmeasured confounders", assumes for observational studies that treatment assignment may be strongly ignorable based on conditional independence; formally, following Rosenbaum and Rubin (1983),

Assumption 2 (Strongly Ignorable Treatment Assignment)
Treatment assignment is strongly ignorable given a vector of pre-treatment covariates, X_i, considered to be confounders if

$$(Y(1), Y(0)) \perp\!\!\!\perp Z|X, \;\; 0 < Pr(Z = 1|X) < 1 \; \forall X \tag{1.2}$$

This produces the result that, for X containing all confounders, the treatment assignment among individuals with a specific value of X is random, such that individuals that share the same value of X show no association between treatment received and potential outcomes. Strongly ignorable treatment assignment assumes all variables that affect the treatment assignment and outcome have been measured. This extends the random assignment mechanism of randomized clinical trials, $(Y(0), Y(1)) \perp\!\!\!\perp Z$ to the observational study scenario $(Y(0), Y(1)) \perp\!\!\!\perp Z|X$. The concept of strongly ignorable treatment is similar to that of exchangeability, which Greenland and Robins (1986) define as equivalence of response type, or "if the exposure states of the two individuals had been exchanged, the same data distribution would have resulted".

1.1.3 Controlling for Pre-Treatment Confounders

Covariate Adjustment

A number of different methods of adjustment for pre-treatment confounders are available. First, a largely used method involves a straightforward covariate adjustment. For parametric covariate adjustment, pre-treatment confounders are controlled for by including them as predictors in the regression model, or by performing matching.

Both the traditional maximum likelihood or Bayesian inference framework can be used, and the ACE can be directly estimated by fitting a regression model. An important caveat to using these covariate adjustment techniques is that the regression model must be correctly specified. If it is not, the average difference and the coefficient representing this difference in the regression model will not necessarily estimate the ACE correctly and some amount of bias will be introduced. Non-parametric models can be used for covariate adjustment to alleviate such problems with model misspecification; however, other issues arise in the form of the "curse of dimensionality" (which plagues many high-dimensional analyses, where there is a much larger number of covariates to include in a model than the number of observations reasonably allows). With only a particular number of units or subjects in the study, these units quickly become scattered and ill-equipped to cover the space of interest well. As more covariates are considered as necessary to adjust for, it becomes exponentially more problematic in obtaining statistically reliable results.

Propensity Scores

A solution to the problems encountered with covariate adjustment or a large number of predictors is to use propensity score methods. These scores, representing the propensity of an individual to receive treatment given all confounding factors are taken into account, allow formulation of a separate model of the necessary covariates. This scalar value can be used in a number of ways, including in a regression adjustment, in place of the covariates themselves or in addition to the most important covariates. Propensity scores are particularly useful for dimension reduction when a large number of confounders and other covariates need to be taken into account for the analysis.

Matching

Another solution, either using the covariates themselves or propensity scores, is to match treated and control units to create a "pseudo-population" of individuals that are similar in characteristics. By choosing well-matched samples of the original treated

and control groups, bias due to covariates is reduced. Stuart (2010) provide a thorough review of existing and new matching methods, providing details on matching methods in practice. In the sections below, we use coarsened exact matching (CEM) (Iacus et al., 2012). The benefits of this particular method for our proposed work include better computational efficiency and avoiding the need to manually readjust the model after post-estimation analysis.

1.1.4 Controlling for Post-Treatment Confounders

While these methods deal with confounding and selection bias via pre-treatment covariates, separate methods consider post-treatment confounders, or *intermediate outcome variables*. Conditioning on a post-treatment covariate or estimating a causal effect based on a post-treatment covariate may result in additional bias. Frangakis and Rubin (2002) developed the principal stratification framework as a cross-classification of individuals by the joint potential values of the post-treatment variable of interest. Since membership within a principal stratum is only reflective of a unit's characteristics, strongly ignorable treatment assignment (along with the other necessary assumptions) still holds in a modified version. Thus, each principal strata is unaffected by treatment assignment and principal effects, causal effects within principal strata, can be found. Principal stratification has been shown to be the best approach to causal inference with post-treatment covariates and provides useful insights; however, it is important to keep in mind that it does not always answer the question of primary interest. VanderWeele (2011) can be referenced for a more in-depth review of this approach.

1.1.5 Spatial Point Processes

Spatial point processes are used to detect patterns and to draw inferences about the distribution of the locations of mapped point data. Each event is an occurrence of

something of interest, with its own event location. Each of these events may occur at a point, or location, within the area of interest; an event location is where an event actually did occur within the area of interest. One of the three fundamental patterns of events is *complete spatial randomness*, in which every event has an equal likelihood of occurring at any location within the pre-specified area of interest, regardless of where the other points are located. This situation allows two key properties to occur: 1) events are uniformly distributed within the area of interest, and 2) events are independent of one another. The two other fundamental pattens are clustering patterns and regular patterns.

More specifically, probabilistic models can be defined to represent spatial patterns and detect what the pattern of events may be. A spatial point process (SPP) is a probabilistic model defined by a set of random variables, where each random variable represents the location of an event in space (Waller and Gotway, 2004).

A particular set of spatial point processes include the stationary *homogeneous spatial Poisson point processes*. Stationary in this sense means that the process is invariant to translation within d-dimensional space (Waller and Gotway, 2004). Homogeneous implies that the nature of the point process is constant across the entire area of interest. More specifically, Diggle (1983) (p.50) and Stoyan et al. (1995) (p.33) define a stationary homogeneous spatial Poisson point process by:

1. The number of events in a finite region A is a random variable following a Poisson distribution with mean $\lambda|A|$ for some constant $\lambda > 0$ and $|A|$ denoting the area of A. This defines the notion of intensity, denoted λ, representing the expected number of events per unit area.

2. Given N=total number of events occurring within an area A, the locations of the N events represent an independent uniform sample of N locations, where each point is equally likely to be chosen as an event. This is equivalent to the concept of complete spatial randomness (CSR) where events are uniformly

distributed across the area.

An advantage of assuming the Poisson distribution is that it allows the total number of events to vary between realizations and at the same time, maintains a fixed number of expected events per unit area. Thus, the estimate of intensity from any given realization of the process is then $\widehat{\lambda} = N/|A|$, the total number of events observed divided by the total area. Cressie (1993) (p.634) provides an equivalent but more foundational definition of both the intensity and the associated homogeneous Poisson process as:

1. The numbers of events in non-overlapping regions are statistically independent.

2. For any region $A \subseteq D$,

$$\lim_{|A|\to 0} \frac{P\left(\text{exactly one event in } A\right)}{|A|} = \lambda > 0$$

where $|A|$ is the area of region A, D is the domain of interest, and

3.

$$\lim_{|A|\to 0} \frac{P\left(\text{two or more events in } A\right)}{|A|} = 0$$

The primary goal for comparisons of intensity functions is to detect local differences between the spatial pattern in event incidence observed in the cases from the spatial pattern observed in the controls. Further details and theory of spatial point processes from numerous fields can be found in Diggle (1983), Cressie (1993) (Ch.1), and Ripley (1991) (Ch.8).

1.1.6 Motivating Problem: Conflict Analysis of Civilian Collaboration in the Afghanistan War

The motivation for a causal inference framework in a spatio-temporal setting in this work is to provide a more robust theoretical foundation for the *matched wake analysis* approach, in addition to adding a spatial theory perspective. The goal of the conflict analysis in Schutte and Donnay (2014) is to better understand deterrence-based and alienation-based approaches to population-centric warfare, and how a mix of these approaches can provide a more complete picture of the conflict. Conflict literature often seeks to separately estimate the coercive or alienating effects of indiscriminate violence in (counter)insurgencies, typically falling under two main camps. The first is deterrence-based, claiming that larger quantities of violence against insurgents disrupt their ability to mobilize. This in turn creates fear in civilian witnesses, deterring their alignment with the uprising, and the increased risk for individuals leads to a more severe collective action problem in the existing rebel movement. The second and contrasting claim is that indiscriminate violence leads to more rebel mobilization, due to civilians joining the rebel forces out of revenge for innocent bystanders that are harmed.

Kalyvas (2006) provides an excellent background on the how and why of violence being applied against civilians in civil wars. Deterring defection and enforcing collaboration through violence is the main focus, and how civilian alienation feeds back into the dynamics of conflict is not touched upon. The underlying assumptions are on mobilization and civilian collaboration as endogenous to military control. On the other side of this argument is reactive mobilization, which argues instead that violence, rather than weakening the military opponent, has an opposite effect. More civilians are *alienated* from the attacker, and collaborate with the opponent. Sheehan (1998) writes about this observation in the Vietnam War,

> *"This is a political war and it calls for discrimination in killing. The best weapon for killing would be a knife, but I'm afraid we can't do it that way. The worst is an airplane. The next worst is artillery. Barring a knife, the best is a rifle - you know who you're killing."* (p. 317)

This quote highlights an ordinal scale for the accuracy of applied tactics, in which violence is more or less prone to harming innocent bystanders and not necessarily classified as binary selective or indiscriminate.

To circle back and summarize, there are two dominant theoretical approaches to reactive mobilization: deterrence- and alienation-based explanations. *Deterrence-based* ideas suggest a negative effect of indiscriminate violence on mobilization for the opponent (Lyall, 2009). There is a risk-reward consideration where survival is the main goal that only works if non-participation actually entails a lower risk for individuals than participation. The collective action problem is imposed on the opponent. If we assume a null hypothesis of no change in civilian cooperation in the presence of indiscriminate violence, the deterrence-based concept can be summarized via the following alternative hypothesis:

> H_{A1}: Indiscriminate violent events lead to more civilian cooperation with the perpetrator of this violence.

In contrast, *alienation-based* ideas assume that, in response to indiscriminate violence in insurgency, there is a reactive collaboration with the military opponent (Ellsberg, 1970). The collective action problem can be solved through selective incentives, and revenge is a strong utility that potentially outweighs the perceived risks. The hypothesis summarizing this ideal is:

> H_{A2}: Indiscriminate violent events lead to more civilian cooperation with the opponent.

Numerous empirical studies have been conducted to solve the long-standing dispute among advocates of deterrence and alienation reactive patterns in insurgencies (Downes, 2007; Kalyvas and Kocher, 2009; Lyall, 2009; Linke et al., 2012; Braithwaite and Johnson, 2012); however, support has been found in favor of both camps of thinking and the discussion remains unsolved.

1.2 Outline

Following this introduction to relevant literature and concepts, this dissertation contains three chapters. First, we discuss causal inference in spatial settings and present the necessary preliminaries in Chapter 2. Chapter 3 introduces a spatial point process theoretical framework for causal inference with geographically-defined treatment. In addition, the focus of Chapter 3 also considers an analytical framework for determining the reach of spatial influence of the exposure event of interest. Next, the framework is extended to include solutions to the possible violation of SUTVA, and an interference effect for this context is proposed in Chapter 4. Then, the implications and current solutions in literature are compared and contrasted in Chapter 5 regarding geographically defined treatments and SUTVA violations. Future work and potential avenues of next steps for this research are also discussed in Chapter 5.

Chapter 2

Causal Inference and Spatial Settings

2.1 Introduction

2.1.1 Causal Effects of Spatial Events

The fundamental ideas of causal inference in the counterfactual framework rely on key assumptions that can be hard to justify when applied to spatial data and geographic questions. When treatments are defined by geographic features (e.g., via local policy decisions), understanding the effects of these treatments through observational studies can be challenging. The same can be said for disaggregated event data and the effects of specific types of events on future events defined geographically. In such instances, the research design and methods to account for the geographic proximity of events are important to consider.

Matching is an intuitive and flexible form of statistical adjustment that can often accommodate the combination of spatio-temporal methods and causal assumptions to obtain an estimated effect of interest. Geographic proximity of events can be taken into consideration in a causal framework by incorporating appropriate corresponding measures into the matching algorithm chosen.

We introduce a spatial point process framework for a specific causal problem of interest, namely, an investigation of local causal impacts within a given radius of particular events occurring at specific point locations and at specific times. A novel approach to this problem comes in the conceptual understanding, notation, and subsequent estimand following from point process theory. We investigate the effect of mis-specifying the sliding spatio-temporal windows used to define spatio-temporal "wakes" of intervention events. We then examine how this may affect bias in the estimation of of treatment and control intervention event intensities and bias in the estimation of the overall causal effect in terms of intensities.

2.1.1.1 Spatially Defined Treatment

How we define the treatment or intervention being investigated is an important determining factor in what methods may be appropriate in analysis. We limit the scope of this chapter to the discussion of one treatment setting: treatment assignment for each unit is determined by an event point and associated distance from the event point in space and time. When considering geography in a causal context, *local* treatments are of most interest - individuals *here* are considered treated, compared to individuals *there* are not. Typically, this spatial definition of treatment manifests itself in one of two ways - units themselves can be defined as treated or not-treated (control) by a spatial measure, or by a fixed spatial boundary and membership to one side or another of this boundary. In either scenario, leveraging the geographical nature of treatment assignment to more appropriately compare units is the common goal of interest. Here, we will focus on units being defined as treated or control in space and time.

2.1.1.2 Motivating Problem: Conflict Data Analysis

In research exploring causes and effects of human conflict, the drivers of local levels of violence can be thought of and described in three main parts. There are *a priori* exposures of any location to violence that can be driven by exogenous geographic conditions. Additionally, levels of violence tend to vary over time, which can be described as the momentum of a conflict at a given time. It is imperative to be able to tease apart the effects of the true exposure and the momentum in order to understand and analyze the causal effect of specific interventions also contributing to levels of violence.

Schutte and Donnay (2014) investigate causal relationships in a specific application, namely, exploring whether civilians are more likely to deny insurgents access to explosives in response to local occurrences of indiscriminate violence. To our knowl-

edge, from an epidemiological or statistical perspective, little has been done in the current literature in regards to this setting, considering both spatial and temporal elements in a potential outcomes framework for causal inference.

2.2 Motivation: Adapting Causal Inference to a Spatial Setting

2.2.1 Notation and Framework

Here we lay out the notation and framework necessary to discuss spatio-temporal causal methods and develop a common framework with clarity and statistical rigor. To begin, we do not use individual-level data but rather point process event data. Rather than a set of individuals, we consider (dependent) outcome events with locations and times for which the rate or probability are impacted by the location and time of intervention events. One of the unique aspects of this spatio-temporal problem is that the defining measure of the treatment impacts the resulting potential outcomes. Below we outline notation following Rubin's causal model (Holland, 1986) and tying it to the problem of interest.

We refer to *intervention events* as the treatment or exposure in a traditional sense, denoted by Z_i, for $i = 1, \ldots, n$ units. A treatment intervention event is denoted $Z_i = 1$ and control intervention event is denoted $Z_i = 0$. In this context, consider each unit i to represent a spatio-temporal cylinder, or wake, centered around its corresponding experienced intervention event and having radius r. The potential outcomes are defined as rates of change in the number of *dependent events* before and after receiving a particular intervention event-type. Specifically, $Y_i(0)$ is the change over time in the number of dependent events for unit i within the radius of influence, where unit i experiences a control intervention event, and $Y_i(1)$ is the change in rate, over time, of

dependent events for unit i experiencing treatment intervention event. The observed outcome for unit i can be defined as a linear combination of the potential outcomes and intervention event status, $Y_i = Z_i \cdot Y_i(1) + (1 - Z_i) \cdot Y_i(0)$. A vector of baseline covariates, which can include geographic variables, is denoted $\boldsymbol{X}_i$ for each unit i. The individual treatment effect for unit i is $Y_i(1) - Y_i(0)$, which is the difference in rates of dependent events over time given that unit i experiences treatment versus control intervention event. The assumptions outlined below in Section 4.2.2 result in the average causal effect (ACE),

$$ACE = \mathbb{E}\{Y_i(1) - Y_i(0)\}, \tag{2.1}$$

which can then be estimated from the data. The ACE can also be thought of as the difference in the change in rate of dependent events over time if the wakes experienced treatment versus control intervention events. We interpret the ACE in this way to lend an easy transition between causal and spatial point process theory.

2.2.2 Assumptions

A major concern when estimating causal effects is the fulfillment of the assumptions that are required to identify estimand(s) of interest. Namely, the *stable unit treatment value assumption* (SUTVA) (Rubin, 1980), a form of the "no unmeasured confounders" assumption, and consistency of the potential outcomes (Robins, 1987). These are stated more formally as follows:

Assumption 1 (Stable Unit Treatment Value Assumption (SUTVA))
For any unit i,
(a) neither $Y_i(1)$ *nor* $Y_i(0)$ *is affected by the treatment any other unit* $j \neq i$ *received - that is, there is no interference between units (Holland, 1986); and*
(b) no matter how unit i received treatment $T_i = 1$*, the outcome that would be observed*

would be $Y_i(1)$ and similarly for treatment $T_i = 0$; i.e., there are no hidden versions of treatments.

The strongly ignorable treatment assumption Rosenbaum and Rubin (1983), or "no unmeasured confounders", is given as,

Assumption 2 (Strongly Ignorable Treatment Assignment)
Treatment assignment is strongly ignorable given a vector of pre-treatment covariates, X_i, considered to be confounders if

$$(Y(1), Y(0)) \perp\!\!\!\perp Z|X, \;\; 0 < Pr(Z = 1|X) < 1 \; \forall X \tag{2.2}$$

Assumption 3 (Consistency of potential outcomes)
The consistency of potential outcomes holds if the potential response of some patient, $Y_i(t)$ to a hypothetical treatment t must coincide with the patient's observed response Y_i whenever the actual treatment T happens to be t.

A more formal definition following similar notation to Galles and Pearl (1998),

$$T = t \Longrightarrow Y_i(t) = Y_i \tag{2.3}$$

2.2.3 Our Goals

It is clear that while there has been some work done to understand causal implications in spatial settings, the setting discussed in Schutte and Donnay (2014) can benefit from a structured statistical framework. As application-specific extensions of causal inference methods to spatial and spatio-temporal settings emerge, it is important to define common elements to identify a core basis for spatial-causal inference. Here we connect the theory of spatial point processes with a causal problem in a spatio-temporal setting. We also quantify the impact of estimating the size of spatio-

temporal cylindrical radii, or the space and time influence of each intervention event, on the bias of the results.

2.3 Current Methods

2.3.1 Matched Wake Analysis

The matched wake analysis (MWA) of Schutte and Donnay (2014) quantifies causal relationships in point location spatio-temporal event data, specifically, when one is interested in how different interventions affect subsequent levels of reactive events in nearby areas. By combining spatial and causal methods, the authors address a number of challenges that arise with spatio-temporal event data with a particular emphasis on conflict research.

Many times, aggregate counts from artificial units of analysis (such as regions, counties, etc.) are used in place of missing natural point-level spatial units of analysis, such as villages. However, by using these aggregate units of analysis two problems arise common to the geographic literature.

First, arbitrary sizes of cells for the units of analysis directly impact the number of observations. The smaller the cell sizes, the more cell-level observations available for analysis if the intensity per cell stays the same. This increases the overall sample size for geographically smaller cells such that for any null hypothesis test, even the smallest of signals can become statistically significant. In contrast, if the overall intensity stays the same and you cut it into smaller and smaller cells, the power per cell will *decrease*.

Second, there is the "modifiable areal unit problem" (MAUP) (Openshaw and Taylor, 1979), which refers to the fact that the selection of artificial cell sizes drives spatial inference. This can be tied to Simpson's Paradox (Simpson, 1951) and the ecological fallacy (Selvin, 1958). Simpson's Paradox can occur when a result that

appears in groups of the data disappears or is reversed when combining the groups of data together. Similarly, an ecological fallacy can occur when an analysis on grouped data leads to different results than those from an analysis on individual data. The MAUP expresses the idea that an analysis on a neighborhood or region (or any arbitrarily defined space) does not yield the same results on a different measure of space, say ZIP code, or on the individuals in that area. In short, as stated by Waller and Gotway (2004), the MAUP is a "geographic manifestation of the ecological fallacy in which conclusions based on data aggregated to a particular set of districts may change if one aggregates the same underlying data to a different set of districts".

While there are a number of approaches that have been considered to remedy these issues, a common method is to use a sliding spatial and temporal window design similar to SaTScan (Kulldorff, 1997), a space-time scan statistic. SaTScan was originally applied to epidemiological problems for testing whether there may be an elevated risk of disease in a certain region but where the extent of the risk elevation is unknown. In general, the method reveals whether a cluster of events is consistent with chance allocation under the null hypothesis.

The MWA method proceeds in three steps.

1. Treatment and control (point) intervention events occur and are identified in space and time.

2. A radius is defined for the spatio-temporal cylinders, or "wakes", around each of the intervention events (Figure 2.1). This defines a *unit* for which the change in rate of dependent events before and after is of interest.

3. The coarsened exact matching (CEM) (Iacus et al., 2012) algorithm is used to match treated intervention event wakes and control intervention event wakes, accounting for appropriate covariates and confounding factors.

After matching is performed, Schutte and Donnay (2014) count the dependent

events prior to the intervention event within each wake and determine the trend of dependent events prior to and following the intervention event. This is then used in the analysis of the outcome of interest, following a difference-in-differences (DD) design (Angrist and Pischke, 2009) to estimate the causal effect of the treatment and assess within-subject before and after change in local rates of dependent events.

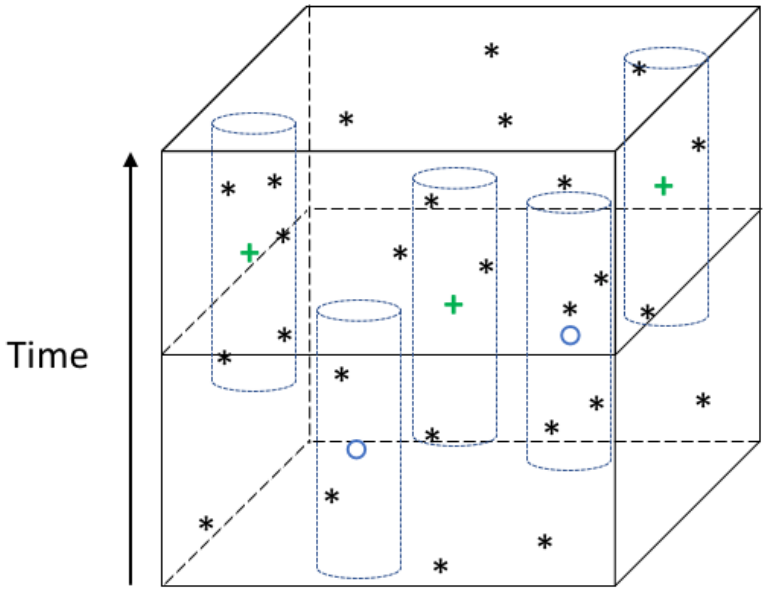


Figure 2.1: An illustration of spatio-temporal wakes in a predefined neighborhood, where '+' represents treatment intervention events, '∘' represents control intervention events, and '∗' represents dependent events. All wakes are of equal radii and time periods before and after intervention events.

2.3.2 Spatial Complications in Definitions of Elements of Standard Causal Framework

While MWA provides a novel approach to evaluating and estimating causal relationships of conflict events that could be used for other types of georeferenced event data, there are some limitations to the approach. First, the spatio-temporal cylinders defining the units that experience treated or control intervention events can potentially overlap with one another, violating SUTVA. Violating this assumption could lead to

biased effect estimates. Two scenarios could result from MWA: (A) multiple units of the same intervention event value overlap in space and time, biasing estimates upward (assuming a positive treatment effect from each); or (B) a unit of each of a treatment intervention and control intervention could overlap and bias the estimated treatment effect downward. Second, spatial concerns relate to defining the wakes themselves, and how to optimally choose the radii of the wakes to allow for unbiased estimation of the causal effect. We address this second limitation and provide a heuristic approach for determining the underlying radius in the sections below.

2.3.3 Spatio-temporal Point Processes

A spatial point process is a stochastic process in which the locations of some events of interest are modeled via random variables and observed within a bounded region. The *intensity* is a first order property of the random process, describing the expected density of events in any location of the region (Waller and Gotway, 2004). In MWA, we assume a causal effect of treatment intervention events results in the intensity inside the cylinder being different than the intensity outside, providing a convenient probabilistic infrastructure for the problem. This allows for *local* insights - i.e., where do patterns appear to differ? - versus *global* insights - i.e., what are the general patterns of clustering and/or regularity with respect to complete spatial randomness?

For a piecewise homogeneous Poisson process in space and time, the expected number of dependent events within each cylinder is intensity times volume. This allows us to take advantage of two properties of the Poisson point process: 1) locations of events are uniform in space and time, and 2) the total number of events within a polygon can be determined (Waller and Gotway, 2004). If the radius is known, we can connect spatial point process theory using the intensity, λ, to the causal difference-in-differences estimator for ACE. This can be framed by the following hypothesis:

H_0: The intensity of dependent events is constant before and after treatment intervention event, i.e., the difference in intensity before and after the intervention event is the same for treatment and control intervention events.

H_A : The intensity of dependent events is higher after treatment intervention event, i.e., the difference in intensity before and after an intervention event is larger for treatment intervention event than control intervention event.

Using these definitions, we can frame the causal difference-in-differences estimator in terms of the true, observed, and expected intensity functions, allowing us to derive the potential bias that occurs in the estimation of the change in intensities of the cylinders, as well as the bias in the estimation of the causal effect. We provide explicit definitions and illustrations for independent treatment events in Chapter 3 and definitions of SUTVA violations and potential adjustments in Chapter 4.

Chapter 3

Linking Spatial Point Process Theory and Causal Inference

3.1 Proposed Spatial Point Process Causal Framework

In this chapter, we outline the necessary assumptions and framework for two scenarios using our spatial point process framework and consider the choice of radius for the spatio-temporal wakes. The scenarios posed below begin by assuming the naïve case in which the radius of both treatment and control intervention event wakes are equivalent and known. In practice, the radius of intervention event wakes is not known and must be chosen by the analyst. We propose a stochastic model approach, in which the dependent events within each cylinder follow a Poisson point process and which allows heuristic estimation of the causal effect radius.

3.1.1 Scenario 1: Constant, Known Spatio-temporal Cylinder Radius.

Here we will draw parallels in notation under spatial point process theory similar to that in Section 2.2.1 above. Let Z_i represent the exposure for unit i, $i = 1, \ldots, n$, with $Z_i = 1$ a treatment intervention event and $Z_i = 0$ a control intervention event. Let $\Delta_i(0) = Y_i(0)$ represent the change in intensity (rate of dependent events) before and after an intervention event for unit i over time, given a control intervention event occurs, $\lambda_{1i}(0) - \lambda_{0i}(0)$. Similarly, $\Delta_i(1) = Y_i(1)$ is the change in intensity over time for unit i corresponding to a treatment intervention event occurring, with $\Delta_i(1) = \lambda_{1i}(1) - \lambda_{0i}(1)$. The observed outcome is then $\Delta_i = Z_i \cdot \Delta_i(1) + (1 - Z_i) \cdot \Delta_i(0)$. Each unit i has an associated vector of pre-intervention event covariates and risk factors, $\boldsymbol{X}_i$. We use spatial intensities λ and changes in spatial intensities Δ in lieu of the traditional outcome notation Y. Assume that consistency holds, in which a unit assigned to treatment intervention event has an observed outcome that is the

potential outcome under the treatment intervention event (and vice versa). Formally,

$$\text{For } i \neq j,\ Z_i = Z_j = 1 \Rightarrow \Delta_i = \Delta_i(1) \text{ and } Z_i = Z_j = 0 \Rightarrow \Delta_i = \Delta_i(0).$$

We also assume positivity, that there is a positive, nonzero probability of receiving every level of exposure for every combination of values of exposure and confounders that occur among individuals in the population (Cole and Frangakis, 2009), and SUTVA (Assumption 1), that there is no interference between units as well as no hidden versions of intervention event types. In these scenarios, we initially assume that there is no potential violation of SUTVA to consider. To identify the causal estimand of interest and obtain unbiased estimates, an ignorability assumption is necessary. Typically, the assumption states that treatment is ignorable conditional on a set of observed covariates, meaning that there are no unmeasured confounders. Below is a version of this statement that can be used when considering spatial settings (Keele and Titiunik, 2017), modified to reflect the new notation:

Assumption 4 (Conditional Geographic Treatment Ignorability)

$$\{\Delta_i(0), \Delta_i(1)\} \perp\!\!\!\perp Z_i | \boldsymbol{X}_i.$$

However, as Keele et al. (2015) point out, there is no *a priori* reason to assume that geographic variation of treatment in such a way is justified, or that it is justified at every possible location and time. As such, instead assume that treatment assignment can be considered as-if randomized as in Assumption 5. This can be seen as a weaker version of Assumption 4, similar to Imbens (2000); Imai and Van Dyk (2004), originally stemming from Rosenbaum and Rubin (1983).

Assumption 5 (Conditional Mean Independence in Local Neighborhood (Keele and Titiunik, 2017))

For all units that reside in a narrow area for which the boundary of the spatio-temporal

cylinder(s) does not encounter any heterogeneity,

$$\mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i, Z_i = 1\} = \mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i, Z_i = 0\}$$

$$\mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 1\} = \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 0\}$$

That is, assume that there exists a small neighborhood around the set of wakes, where the average potential outcomes are mean independent of intervention event type, given the covariates. In this setting, the causal estimand of interest is identifiable:

$$\begin{aligned}
\tau &= \mathbb{E}\{\Delta_i(1) - \Delta_i(0)\} \qquad (3.1)\\
&= \mathbb{E}\left[\mathbb{E}\{\Delta_i(1) - \Delta_i(0)|\boldsymbol{X}_i\}\right]\\
&= \mathbb{E}\left[\mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i\} - \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i\}\right]\\
&= \mathbb{E}\left[\mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i, Z_i = 1\} - \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 0\}\right]\\
&= \mathbb{E}\left[\mathbb{E}\{\Delta_i|\boldsymbol{X}_i, Z_i = 1\} - \mathbb{E}\{\Delta_i|\boldsymbol{X}_i, Z_i = 0\}\right]
\end{aligned}$$

For the Difference-in-Differences (DD) design, the observed outcome can then be modeled by

$$\begin{aligned}
\Delta_i &= \beta_0 + \beta_1 n_{pre} + \tau Z_i + \epsilon_i\\
&= \boldsymbol{X}_i^T\boldsymbol{\beta} + \tau Z_i + \epsilon_i
\end{aligned} \qquad (3.2)$$

where Δ_i is the observed change in intensity over time, τ represents the difference between experiencing treatment or control intervention event in change in intensity over time for unit i, n_{pre} is the number of dependent events prior to the intervention event occurring, and

$$\begin{aligned}\epsilon_i =& Z_i \left[\mathbb{E}\{\Delta_i(1) - \Delta_i(0)|\boldsymbol{X}_i\} - \mathbb{E}\{\Delta_i(1) - \Delta_i(0)\}\right] \\ &+ (1 - Z_i) \left[\Delta_i(0) - \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 0\}\right] \\ &+ Z_i \left[\Delta_i(1) - \mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i, Z_i = 1\}\right].\end{aligned}$$

Under consistency and Assumptions 1 and 5, $\mathbb{E}\left(\epsilon_i|\boldsymbol{X}_i, Z_i\right) = 0$ and τ can be consistently estimated.

3.1.2 Scenario 2: Constant, Unknown Spatio-temporal Cylinder Radius.

The 'unknown' element of this scenario refers to the radial size of the intervention event wakes. Unlike Scenario 1, this must be defined by the analyst. We follow the idea of Schutte and Donnay (2014) and use methods similar to SaTScan (Kulldorff, 1997) to define the optimal radius size, which we assume is the same for both treated and control intervention events. A sliding window design is used to find the radius corresponding to the greatest average difference in dependent event intensities.

The overall causal framework, including assumptions that are made, is the same as in Scenario 1. In addition, we discuss the impact of estimating the size of the spatio-temporal cylinders and potential biases in the context of spatial intensities of these cylinders here.

Consider the background intensity of the area to be constant $\tilde{\lambda}$. The overall temporal length of cylindrical wakes is denoted t, with intervention event i occurring at $\frac{t}{2}$ so the length of time before and after the event is the same. The intensity prior to the event for unit i is denoted $\lambda_{0i}(Z_i)$, where $Z_i = 0$ or 1 for a control or treatment intervention event, respectively. Similarly, the intensity after a control or treatment intervention event for unit i is denoted $\lambda_{1i}(Z_i)$. The causal effect of a treatment

intervention event versus a control intervention event for unit i, $i = 1, \ldots, n$, is the DD estimator as in Scenario 1, equation (3.2).

Next, we assume that the intensities before and after a control intervention event are equivalent (i.e., control events have no effect on the outcome). That is, $\mathbb{E}\left[\Delta_i(0)\right] = \mathbb{E}\left[\lambda_{1i}(0) - \lambda_{0i}(0)\right] = \mathbb{E}\left[\lambda_i(0) - \lambda_i(0)\right] = 0$, for all i. Also assume a constant change over time for treatment intervention event, $\mathbb{E}\left[\Delta_i(1)\right] = \mathbb{E}\left[\lambda_{1i}(1) - \lambda_{0i}(1)\right] = c$. Thus the ACE τ can be written under these simplifications and for these purposes as

$$
\begin{aligned}
\tau &= \mathbb{E}\{\Delta_i(1) - \Delta_i(0)\} \qquad (3.3)\\
&= \mathbb{E}\left[\mathbb{E}\{\Delta_i(1) - \Delta_i(0)|\boldsymbol{X}_i\}\right]\\
&= \mathbb{E}\left[\mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i\} - \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i\}\right]\\
&= \mathbb{E}\left[\mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i, Z_i = 1\} - \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 0\}\right]\\
&= \mathbb{E}\left[\mathbb{E}\{\lambda_{1i}(1) - \lambda_{0i}(1)|\boldsymbol{X}_i, Z_i = 1\} - \mathbb{E}\{\lambda_{1i}(0) - \lambda_{0i}(0)|\boldsymbol{X}_i, Z_i = 0\}\right]\\
&= c
\end{aligned}
$$

There are then three cases when considering the radius of the cylindrical wake: 1) the radius is correctly specified, $\hat{r} = r_{true}$, 2) the proposed radius is smaller than the truth, $\hat{r} < r_{true}$, and 3) the proposed radius is larger than the truth, $\hat{r} > r_{true}$. Figures 3.1 and 3.2 illustrate the implications of 2) and 3), respectively. Below are the analytical derivations of the bias within the outcome measure under each intervention event, $\Delta(0)$ and $\Delta(1)$, as well as the bias of the effect itself. In each case, this effect can still be calculated as a difference-in-difference estimate.

When the proposed radius of the wake is *equal* to the truth, r, our estimate is unbiased.

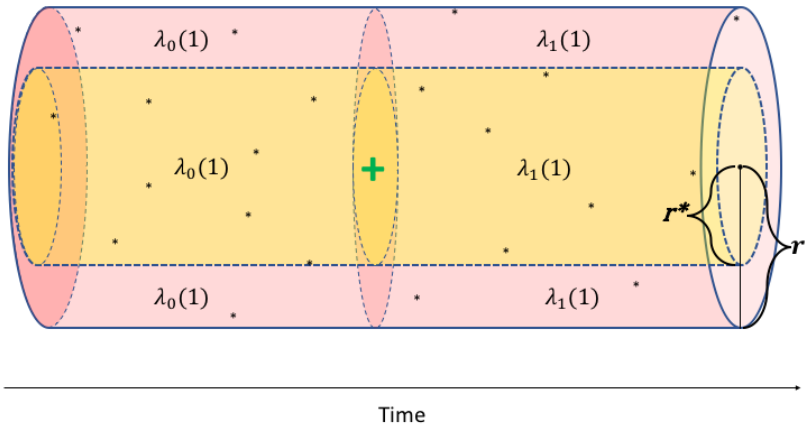


Figure 3.1: Illustrating the concept of smaller-than-truth radius for cylindrical wakes. The cylinder in red represents the true wake of a treatment intervention event, with pre-intervention event intensity $\lambda_0(1)$ and post-intervention event intensity $\lambda_1(1)$. The yellow cylindrical wake includes these pre- and post-intervention event intensities, though for a smaller overall volume.

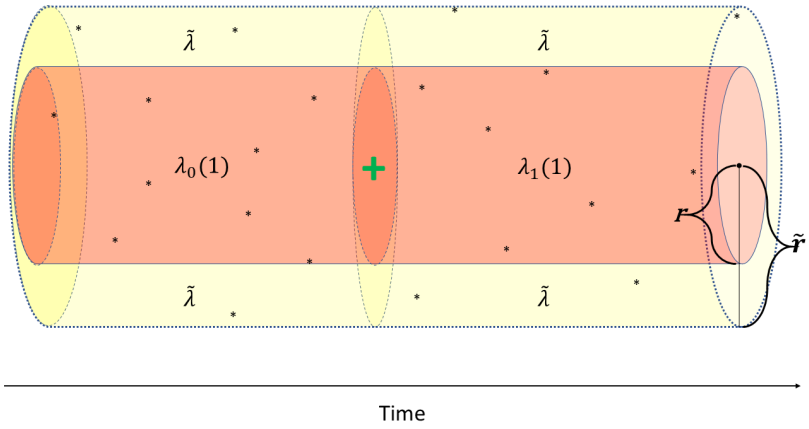


Figure 3.2: Illustrating the concept of larger-than-truth radius for cylindrical wakes. The cylinder in red represents the true wake of a treatment intervention event, with pre-intervention event intensity $\lambda_0(1)$ and post-intervention event intensity $\lambda_1(1)$. The yellow cylindrical wake includes these pre- and post-intervention event intensities, in addition to background intensity $\tilde{\lambda}$ in both the pre- and post-intervention event sections. The inclusion of the background intensity for a portion of the wake is realized as the decreasing slope after the true spatial radius in Figures 3.3 and 3.4.

$$\mathbb{E}\left[\widehat{\lambda}_{0i}(Z_i)\right] = \mathbb{E}\left[\frac{\#\text{ dependent events before intervention event for unit } i}{Volume}\right] = \frac{\pi r^2(t/2)\lambda_0(Z_i)}{\pi r^2(t/2)} = \lambda_{0i}(Z_i)$$

$$\mathbb{E}\left[\widehat{\lambda}_{1i}(Z_i)\right] = \mathbb{E}\left[\frac{\#\text{ dependent events after intervention event for unit } i}{Volume}\right] = \frac{\pi r^2(t/2)\lambda_{1i}(Z_i)}{\pi r^2(t/2)} = \lambda_{1i}(Z_i)$$

$$\mathbb{E}\left[\widehat{\Delta}_i(0)\right] = \mathbb{E}\left[\widehat{\lambda}_{1i}(0) - \widehat{\lambda}_{0i}(0)\right] = \lambda_{1i}(0) - \lambda_{0i}(0) = 0$$

$$\mathbb{E}\left[\widehat{\Delta}_i(1)\right] = \mathbb{E}\left[\widehat{\lambda}_{1i}(1) - \widehat{\lambda}_{0i}(1)\right] = \lambda_{1i}(1) - \lambda_{0i}(1) = c$$

$$Bias = \mathbb{E}\left[\widehat{\Delta_i(1) - \Delta_i(0)}\right] - \mathbb{E}\left[\Delta_i(1) - \Delta_i(0)\right] = 0$$

When the proposed radius of the wake is *smaller* than the truth, $r^* < r$ ($r = r^* + x$), our estimate remains unbiased.

$$\mathbb{E}\left[\widehat{\lambda}_{0i}(Z_i)\right] = \mathbb{E}\left[\frac{\#\text{ dependent events before intervention event}}{Volume}\right] = \frac{\pi r^{*2}(t/2)\lambda_{0i}(Z_i)}{\pi r^{*2}(t/2)} = \lambda_{0i}(Z_i)$$

$$\mathbb{E}\left[\widehat{\lambda}_{1i}(Z_i)\right] = \mathbb{E}\left[\frac{\#\text{ dependent events after intervention event}}{Volume}\right] = \frac{\pi r^{*2}(t/2)\lambda_{1i}(Z_i)}{\pi r^{*2}(t/2)} = \lambda_{1i}(Z_i)$$

$$\mathbb{E}\left[\widehat{\Delta}_i(0)\right] = \mathbb{E}\left[\widehat{\lambda}_{1i}(0) - \widehat{\lambda}_{0i}(0)\right] = \lambda_{1i}(0) - \lambda_{0i}(0) = 0$$

$$\mathbb{E}\left[\widehat{\Delta}_i(1)\right] = \mathbb{E}\left[\widehat{\lambda}_{1i}(1) - \widehat{\lambda}_{0i}(1)\right] = \lambda_{1i}(1) - \lambda_{0i}(1) = c$$

$$Bias = \mathbb{E}\left[\widehat{\Delta}_i(1) - \widehat{\Delta}_i(0)\right] - \mathbb{E}\left[\Delta_i(1) - \Delta_i(0)\right] = 0$$

When the proposed radius of the wake is *larger* than the truth, $\tilde{r} > r$ ($\tilde{r} = r + (\tilde{r} - r)$), our estimate will include data from the true higher-intensity cylinder **and** data from the background intensity, resulting in bias; specifically,

$$\begin{aligned}
\mathbb{E}\left[\widehat{\lambda}_{0i}(Z_i)\right] &= \mathbb{E}\left[\tfrac{\text{\# dependent events before intervention event}}{Volume}\right] &&= \tfrac{\pi r^2(t/2)\lambda_{0i}(Z_i)+\pi(\tilde{r}-r)^2(t/2)\tilde{\lambda}}{\pi\tilde{r}^2(t/2)} \\
&&&= \tfrac{r^2\lambda_{0i}(Z_i)+\left(\tilde{r}^2-2\tilde{r}r+r^2\right)\tilde{\lambda}}{\tilde{r}^2} \\
&&&= \tfrac{r^2\left(\lambda_{0i}(Z_i)+\tilde{\lambda}\right)+\left(\tilde{r}^2-2\tilde{r}r\right)\tilde{\lambda}}{\tilde{r}^2} \\
\mathbb{E}\left[\widehat{\lambda}_{1i}(Z_i)\right] &= \mathbb{E}\left[\tfrac{\text{\# dependent events after intervention event}}{Volume}\right] &&= \tfrac{\pi r^2(t/2)\lambda_{1i}(Z_i)+\pi(\tilde{r}-r)^2(t/2)\tilde{\lambda}}{\pi\tilde{r}^2(t/2)} \\
&&&= \tfrac{r^2\lambda_{1i}(Z_i)+\left(\tilde{r}^2-2\tilde{r}r+r^2\right)\tilde{\lambda}}{\tilde{r}^2} \\
&&&= \tfrac{r^2\left(\lambda_{1i}(Z_i)+\tilde{\lambda}\right)+\left(\tilde{r}^2-2\tilde{r}r\right)\tilde{\lambda}}{\tilde{r}^2}
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}\left[\widehat{\Delta}_i(0)\right] &= \mathbb{E}\left[\widehat{\lambda}_{1i}(0)-\widehat{\lambda}_{0i}(0)\right] &&= \left\{\tfrac{r^2\left(\lambda_{1i}(0)+\tilde{\lambda}\right)+\left(\tilde{r}^2-2\tilde{r}r\right)\tilde{\lambda}}{\tilde{r}^2}\right\}-\left\{\tfrac{r^2\left(\lambda_{0i}(0)+\tilde{\lambda}\right)+\left(\tilde{r}^2-2\tilde{r}r\right)\tilde{\lambda}}{\tilde{r}^2}\right\} \\
&&&= \tfrac{r^2\left(\lambda_i(0)+\tilde{\lambda}\right)-r^2\left(\lambda_i(0)+\tilde{\lambda}\right)}{\tilde{r}^2} \\
&&&= 0 \\
\mathbb{E}\left[\widehat{\Delta}_i(1)\right] &= \mathbb{E}\left[\widehat{\lambda}_{1i}(1)-\widehat{\lambda}_{0i}(1)\right] &&= \left\{\tfrac{r^2\left(\lambda_{1i}(1)+\tilde{\lambda}\right)+\left(\tilde{r}^2-2\tilde{r}r\right)\tilde{\lambda}}{\tilde{r}^2}\right\}-\left\{\tfrac{r^2\left(\lambda_{0i}(1)+\tilde{\lambda}\right)+\left(\tilde{r}^2-2\tilde{r}r\right)\tilde{\lambda}}{\tilde{r}^2}\right\} \\
&&&= \tfrac{r^2\left(\lambda_{1i}(1)-\tilde{\lambda}\right)-r^2\left(\lambda_0(1)-\tilde{\lambda}\right)}{\tilde{r}^2} \\
&&&= \tfrac{r^2(\lambda_{1i}(1)-\lambda_{0i}(1))}{\tilde{r}^2} \\
&&&= c\cdot\tfrac{r^2}{\tilde{r}^2}
\end{aligned}$$

$$Bias = \mathbb{E}\left[\widehat{\Delta}_i(1)-\widehat{\Delta}_i(0)\right]-\mathbb{E}\left[\Delta_i(1)-\Delta_i(0)\right] = c\cdot\left(\frac{r^2}{\tilde{r}^2}-1\right)<0, \text{ a parabola in } r.$$

We note that the background effect need not be the same as the control effect - that is, the intensity observed within a control wake can be a different level than the overall background intensity of the entire space. It is clear that overestimating the size

of the cylindrical wakes that are used in the matching process and thus estimation of the treatment effect introduces bias. Under the circumstances of no change in intensity before and after a control intervention event, or $\Delta(0) = 0$, the bias in the treatment effect estimate comes from the estimation of the change in intensity before and after a treatment intervention event. By assuming that $\tilde{\lambda} < \lambda_0(Z_i) \leq \lambda_1(Z_i)$, the estimated change in intensity for treatment wakes is smaller than the true causal effect due to the inclusion of the moderating portion of the wake with intensity $\tilde{\lambda}$. This carries over to the bias observed in the treatment effect.

3.2 Results

3.2.1 Simulations

Simulations were conducted mimicking the conflict data from Schutte and Donnay (2014). A total of 100 treatment and 200 control intervention events were distributed within an approximately 222 km by 222 km area occurring within one year. The background intensity was specified to be 0.002, and dependent events were generated following a Poisson distribution with mean intensity 0.002 times the volume of the whole cube minus the summation of the volume of the cylindrical wakes to represent the background events. Background dependent events were generated in this way so as to avoid overlapping placement with dependent events occurring within wakes of interest and avoid skewing estimation of intensities. A random Poisson distribution with mean equal to intensity times volume of the cylinder was used to generate the number of dependent events pre- and post- intervention event. We assume a pre-intervention event intensity of 0.2, with an equivalent post-intervention event intensity for controls and an increased post-intervention event intensity for treatment of 0.2×1.75, or 0.35. The true causal effect is then 0.15, which represents an increase of 0.15 expected dependent events per unit volume following a treatment intervention event

compared to control intervention event, on average.

The effect size of 1.75 represents the change in intensity of dependent events from pre- to post- intervention event, given a unit experiences treatment versus control intervention events. For both treatment and control intervention wakes, the true size is set with a temporal window of 8 days informed by the results of Schutte and Donnay (2014) and spatial windows ranging from approximately 1.11 km to 11.11 km. This represents the temporal lag before and after the intervention event, and the spatial radius of associated influence for an intervention event. Ten iterations were performed for each spatial window size, corresponding to the radius of the wakes, ranging from 0.01 to 0.1 by 0.01 increments, which represents approximately 1.11 km to 11.11 km increasing in size by 1.11 km steps, as the spatial influence of the wakes.

Our goal is to distinguish a signal at the true spatial radius for the wakes. The average estimated change over time in intensity for each of control and treatment intervention events, representing the change in rate of dependent events within the volume of the wake before to after each intervention event wake is illustrated, in addition to the difference in the changes in intensity over time, on average, given a wake experiences a treatment intervention event in comparison to if it experienced a control intervention event ($\widehat{ACE} = \mathbb{E}\left[\widehat{\Delta}(1) - \widehat{\Delta}(0)\right]$). Thus, we illustrate the estimated average causal effect over all iterations.

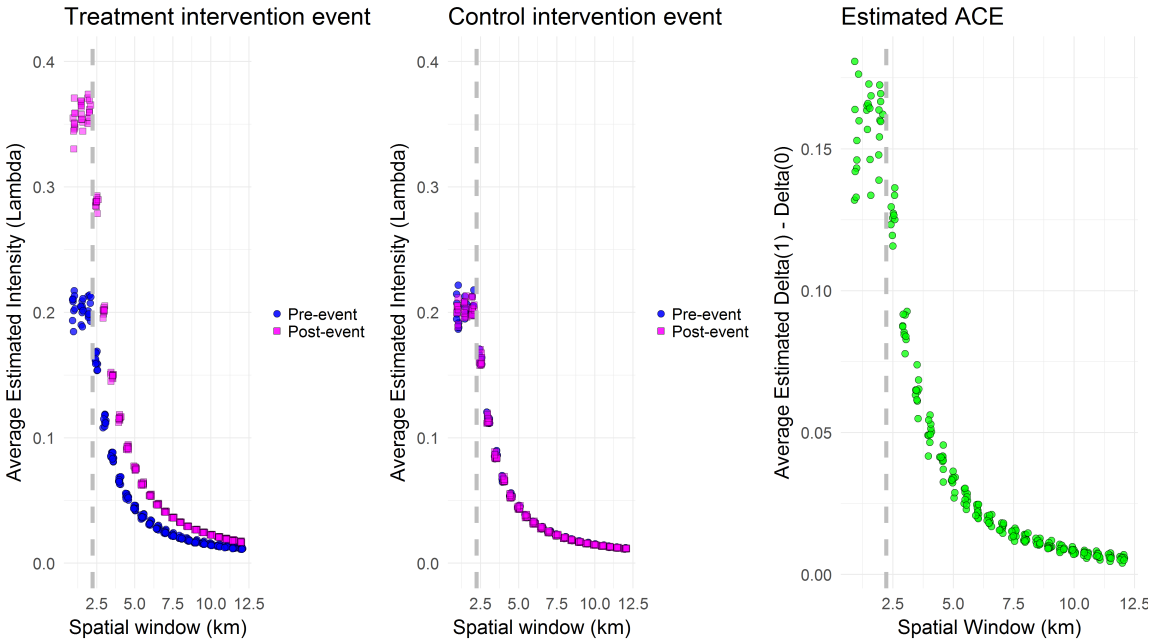
Four of the simulation scenarios are given in Figures 3.3 and 3.4 for illustration purposes. The rest of the scenarios can be found in Appendix A, showing similar patterns as detailed here. There is a clear delineation in the pattern of the average estimated $\widehat{\lambda}_{\cdot}(1)$ and $\widehat{\lambda}_{\cdot}(0)$ before and after the true radius size in each plot. The dotted vertical line represents the true spatial window, and is marking a change between the radius when smaller than and larger than the truth.

In Figures 3.3 (a) and (b) and 3.4 (c) and (d), the variation in the estimate increases as the radius decreases as one would expect with decreasing sample size.

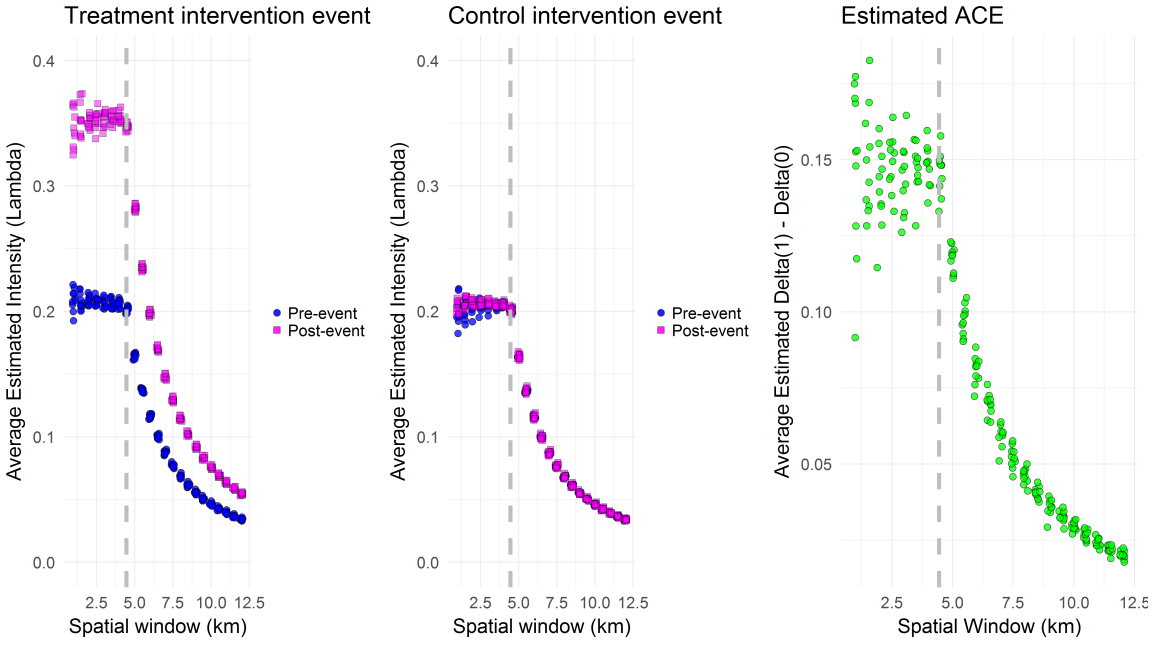
This empirical unbiasedness when the proposed radius of the wake is smaller than the truth follows the analytical derivation in Section 3.1.2. For larger spatial windows than the truth, the steeply decreasing slope reflects the bias of a larger proposed wake than the truth. Again, this empirical evidence supports the analytical derivation of bias, equal to $0.35 \times (\frac{r^2}{\tilde{r}^2} - 1) < 0$, where 0.35 is the true treatment effect, r is the true radius and $\tilde{r}$ is the larger, proposed radius. When the spatial window is larger than the truth, the background intensity is included in the wake intensity estimation, essentially diluting the true, higher intensities seen within the wake. This is illustrated in Figure 3.2 by the yellow cylinder, representing a treatment wake with a larger-than-true radius, surrounding the red treatment wake with the correct radius.

The leftmost panel of each plot illustrates the estimated average change in intensity of dependent events over time given units experience treatment intervention event ($\widehat{\Delta}(1) = \widehat{\lambda}_1(1) - \widehat{\lambda}_0(1)$). This shows a large difference in average estimated intensities before and after the intervention event, as was expected. The middle panel shows this estimated change in intensity of dependent events over time for units, given they experience a control intervention event ($\widehat{\Delta}(0) = \widehat{\lambda}_1(0) - \widehat{\lambda}_0(0)$). There is complete overlap of the distribution of pre- and post-intervention event average estimated intensities, indicating there is no change in intensity over time under control. Again, this illustrates what was expected. The estimated average effect of treatment on the change in intensity of dependent events over time is illustrated in the rightmost panel of each plot. The vertical line at the true radius of the wakes marks the end of unbiasedness in this estimated effect.

Figure 3.5 also confirms that the empirical bias follows the analytically derived bias in each of these scenarios.

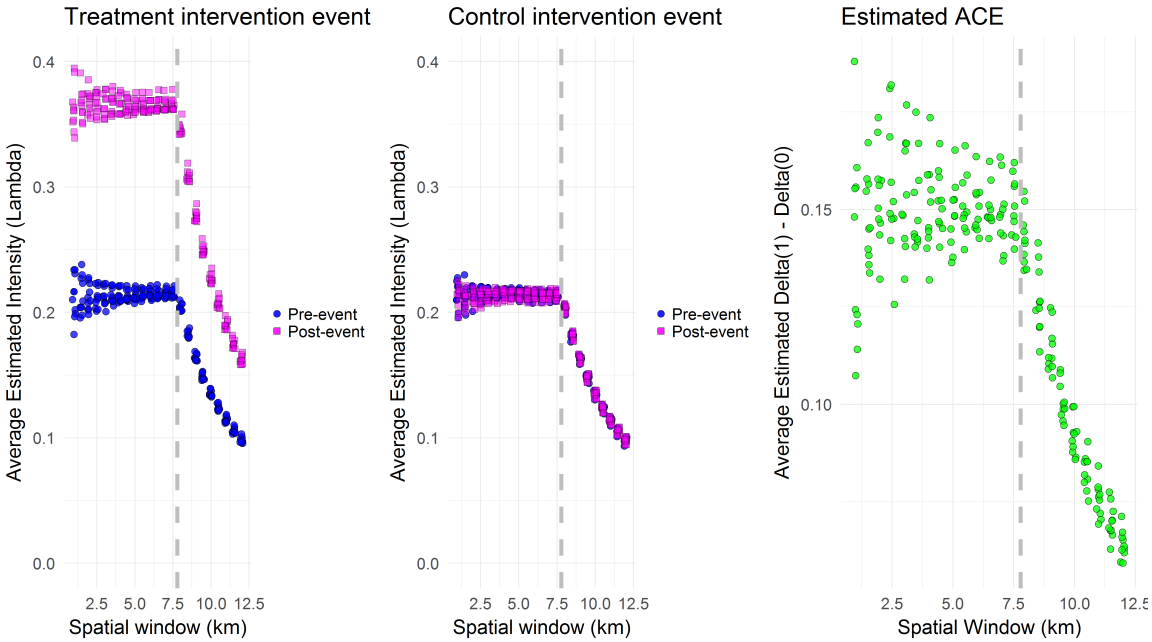


(a)

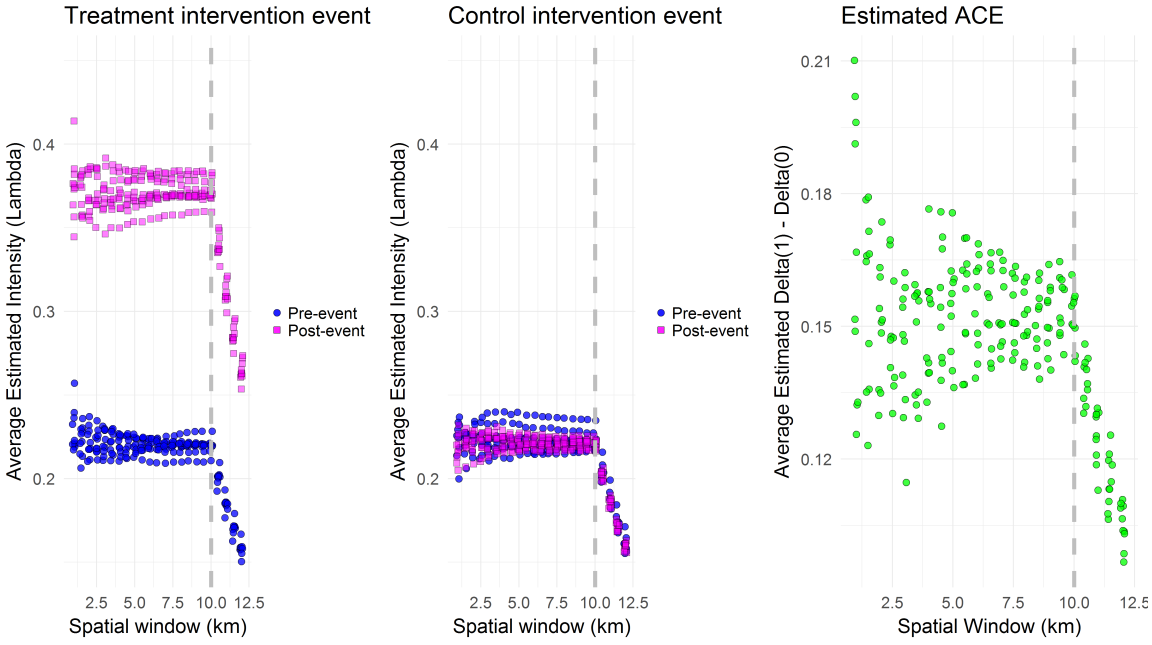


(b)

Figure 3.3: Illustrations of the average $\widehat{\lambda}_1(1)$ and $\widehat{\lambda}_0(1)$ with 100 treatment intervention events (left panels), average $\widehat{\lambda}_1(0)$ and $\widehat{\lambda}_0(0)$ with 200 control intervention events (middle panels), and average difference in $\widehat{\Delta}(1)$ and $\widehat{\Delta}(0)$ (right panels). The dotted gray lines represent the true spatial window size.



(c)



(d)

Figure 3.4: Continued illustrations of the average $\widehat{\lambda}_1(1)$ and $\widehat{\lambda}_0(1)$ with 100 treatment intervention events (left panels), average $\widehat{\lambda}_1(0)$ and $\widehat{\lambda}_0(0)$ with 200 control intervention events (middle panels), and average difference in $\widehat{\Delta}(1)$ and $\widehat{\Delta}(0)$ (right panels). The dotted gray lines represent the true spatial window size.

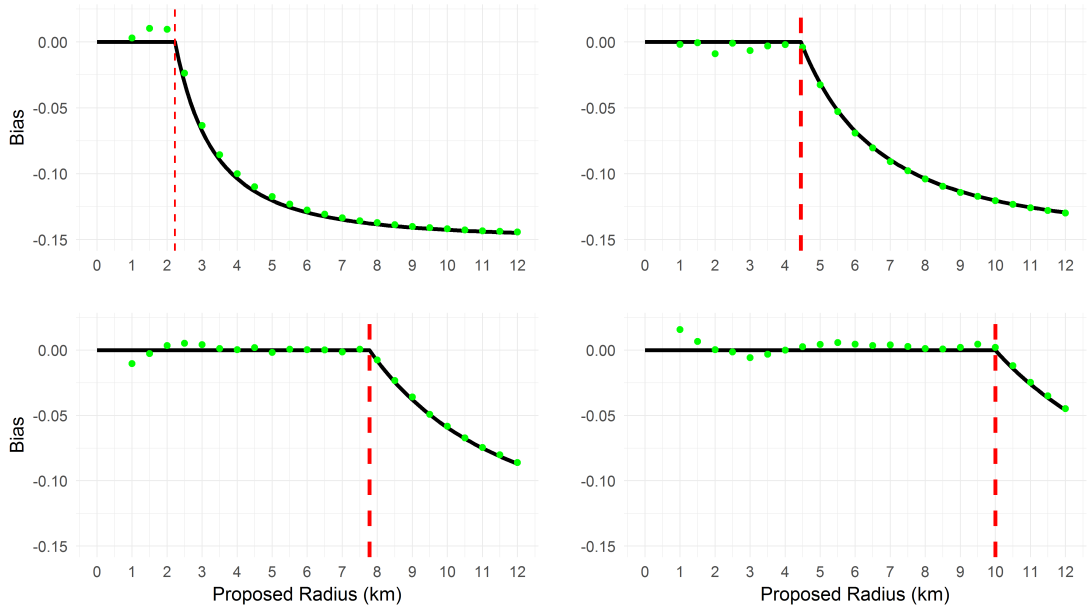


Figure 3.5: Analytically derived bias for each true radius, denoted by the red vertical dotted line, is shown by the black solid line. Empirical averages of the bias at each spatial window included in the simulation are shown as green dots.

3.2.2 Conflict Data: Civilian Collaboration in Afghanistan

The spatial and temporal distribution of conflict events (O'Loughlin et al., 2010) as well as the clustering of such events in space and time (Braithwaite and Johnson, 2012) have been topics of interest recently when considering the conflicts in Afghanistan and Iraq. In 2010, wikileaks.org released data coded by the U.S. military of temporally and spatially referenced conflict events in Afghanistan and Iraq, known as SIGACT ("Significant Activity") (SIGACT, 2010). It is the largest, and arguably most complete and unbiased data set about the Afghan War to date, providing information on both insurgent and incumbent activity, encompassing 154 types of events, and spanning from 2004 to 2009. However, it is not without its limitations. The four key limitations of the dataset are (Schutte, 2017):

1. Soldiers in the field report civilian casualties without a second, independent confirmation which may lead to lower counts

2. Activities of other coalition troops, private contractors, and other U.S. service branches are not systematically recorded in the data

3. Indirect fire or air strikes may harm civilians without ground troops' knowledge

4. Insurgent-civilian relationships are not recorded in the data

Figure 3.6 illustrates the complex relationships that occur, and the capabilities of the SIGACT data set in capturing these. We consider an analysis similar to those completed in Schutte and Donnay (2014) and Schutte (2017). Events were analyzed in the statistical aggregate and no marginal effects were estimated for geographic covariates to protect individuals, institutions, and political actors involved in these conflict events. The measure of civilian collaboration is the turning-in of unexploded ordnances or other explosive remnants of war that could be used by insurgents against U.S. forces, since these are largely what is used in attacks on both civilians and military by insurgents in Afghanistan. Here, variation in the levels of civilian support to the incumbent (U.S. forces) is used to analyze the effects of violence. In general, civilians have three options when a violent event occurs that potentially harms fellow civilians/bystanders:

1. Remain passive - do nothing, and unexploded ordnance would thus be obtained by insurgents

2. Alert insurgents to the presence of the unexploded ordnance

3. Alert incumbents of the threat, sparing both lives and materials on the incumbent side (as well as potentially fellow civilians)

Choosing to act in options 2 or 3 would allow civilians to take sides in a civil war without having to publicly declare loyalties, which can be risky. Table 3.1 details the event types that are selected as indiscriminate and selective insurgent and incumbent violence and civilian assistance, since events are not as densely clustered as during

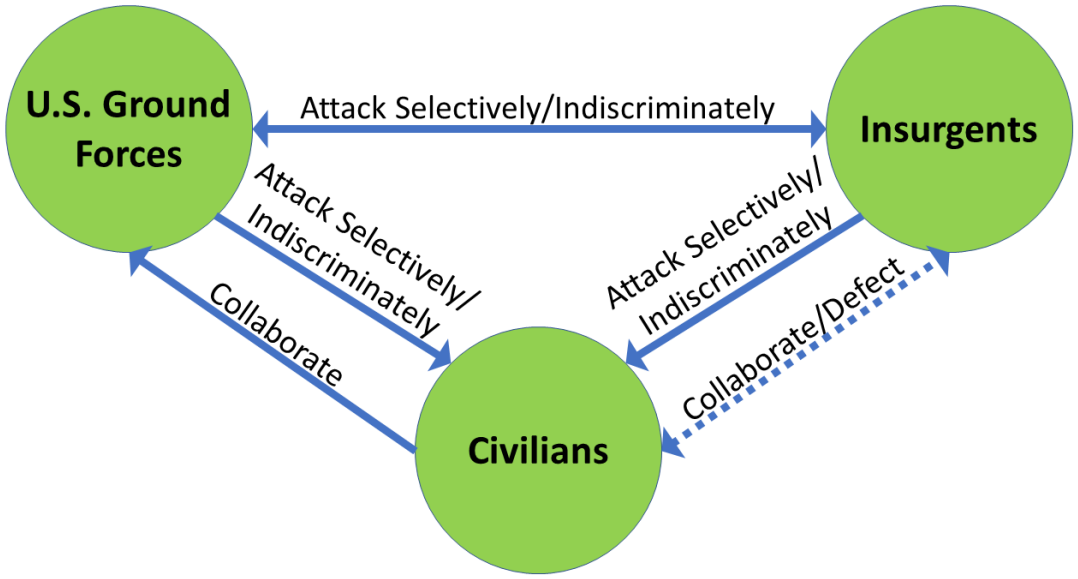


Figure 3.6: Adapted from Schutte (2017). A chart outlining the relationships and intricacies of relationships in the Afghan War, and what the SIGACT data set includes. The dotted arrow between civilians and insurgents signifies that these relationships are unable to be captured in the data.

the most intense violence in earlier years and lend themselves well to the hypothesis of interest. Treatment intervention events are instances of indiscriminate insurgent violence using IEDs - those that led to civilian casualties, and control events are instances of selective insurgent violence using IEDs - those that did not claim civilian lives. Dependent events are instances of civilian assistance, which can occur before and after treatment and control events.

Type of Event In our Analysis	**Event Category in SIGACT**	**N**
Civilian Assistance	turn-in, evidence turn-in/received, ERW/turn-in	899
Insurgent Selective Violence	direct fire	15,438
Insurgent Indiscriminate Violence	mine strike, indirect fire	7,151
Incumbent Selective Violence	direct fire	823
Incumbent Indiscriminate Violence	indirect fire, close air support	471

Table 3.1: Summary of event category definitions from SIGACT for selective and indiscriminate violence (intervention events), as well as civilian assistance (dependent events) in Afghanistan. (Total number of events is 24,782.)

Figure 3.7 shows a map of Afghanistan in relation to the data points in a 3-dimensional cloud. In this analysis, we consider the effect of experiencing a nearby indiscriminate intervention event on the change in *intensity* of civilian assistance, in comparison to the change in intensity of civilian assistance before and after an intervention event, given it had been a selective violent event. Spatial windows from 1 to 20 km and temporal windows from 5 to 75 days (every 5 days) were considered. The hypothesis of interest is:

> H_A: Indiscriminate insurgent violence (treatment) using IEDs increases civilian handover of unexploded ordnances to U.S. troops compared to selective insurgent violence (control) using IEDs.

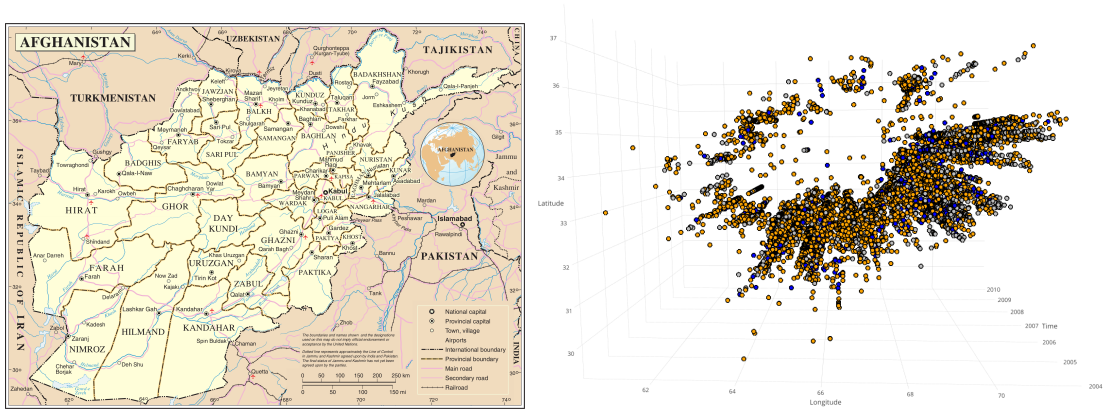


Figure 3.7: A comparison of a map of Afghanistan based on UN map no. 3958 Rev. 7 June 2011 to the 3-dimensional illustration of the data points in our data set. Patterns of heavier events throughout time occur in the south and southeaster portions of Afghanistan, following the border.

Treatment and control events are associated with geographic covariates through nearest-neighbor mapping (Schutte and Donnay, 2014). We include geographic variables similarly considered in Schutte (2017), informed by theory and correlation with insurgent and incumbent control. This is important so that comparisons of the treatment and control interventions do not lead to false inferences and to take into consideration potential vastly different conditions. Distance to Pakistan and Kabul are

included to account for remoteness from the state's power center and the ability to seek refuge across international borders, along with spatially disaggregated wealth (GECON) (Nordhaus et al., 2006) since wealthier regions may be better protected by state power, population numbers for the year 2000 (CIESIN and CIAT, 2005) to correlate with the number of conflict events in a region, spatially referenced data to account for elevation and inaccessible terrain important for providing shelter (Gesch et al., 1999), an indicator for the predominant ethnic groups in the region and Pashtun tribal areas where the heaviest fighting occurs, and a line-of-sight dataset (details of its generation found in Schutte (2017)) to explain the use of indirect strikes or air attacks. Pretreatment trends in civilian assistance and previous treatment and control intervention are also matched on, the latter pattern included to mitigate any issues overlapping wakes may cause (Schutte and Donnay, 2014). A summary of all variables that are adjusted for is given in Table 3.2.

	Mean (S.D.) or N (%)	**Minimum**	**Maximum**
Distance to Pakistan (km)	125.44 (103.85)	0.00	610.32
Distance to Kabul (km)	331.46 (184.34)	0.12	806.12
Elevation above sea level (km)	1382.05 (580.28)	290.36	4542.00
Population count in 2000	1300.90 (6558.65)	2.76	123730.24
GECON	0.35 (0.59)	0.01	3.12
Line of sight	165.85 (137.76)	9.00	925.00
Pashtun tribal areas	20398 (82.31)	-	-
Hazara tribal areas	343 (1.38)	-	-

Table 3.2: Summary statistics of covariates included in the matching algorithm for estimating the causal effect (N=24,782).

3.2.2.1 Results

Evaluation of the estimated intensities in Figure 3.8 indicate that the most appropriate spatial window would be within the first 3 km. Given that there is no constant 'plateau' as we saw in the simulated scenarios, the smallest radius seems to be the most fitting choice for this data. As such, it further confirms that this analysis as-

suming no interference is worthwhile as an initial given the rarity of turn-in events in comparison to intervention events and the small spatial window of influence that is identified.

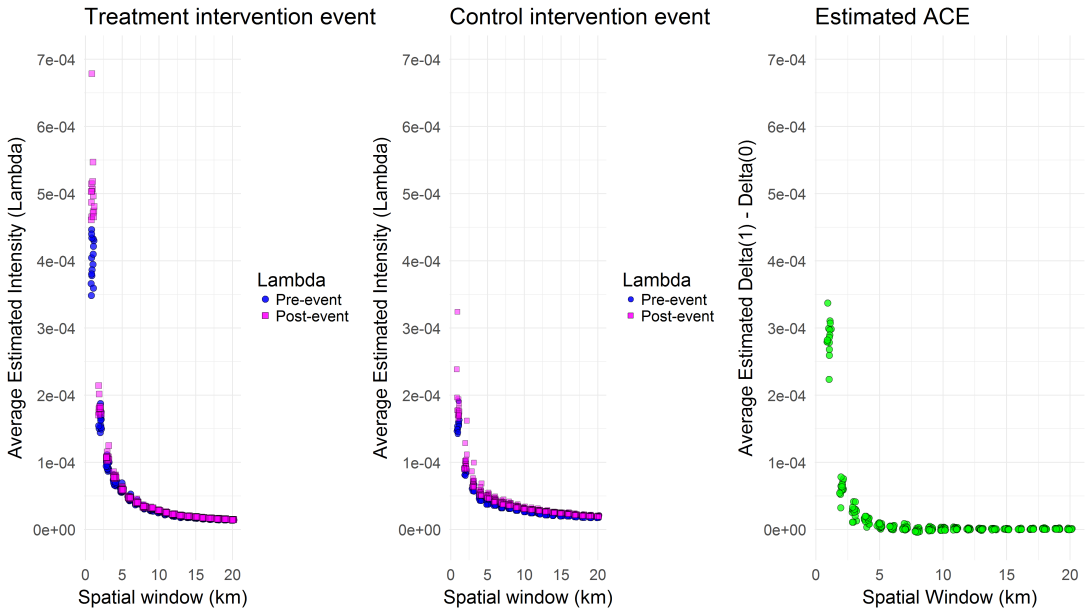


Figure 3.8: Estimated intensities and average causal effect to determine spatial radius that should be used.

Further examination of the contour plot (Figure 3.9) showing the before-and-after average causal effect of the DD regression coincides with spatial windows in the immediate vicinity (1-3 km) of violent events, indicating statistically significant events for wakes of this size. The largest effect estimates occur in this interpretable area, where $p < 0.05$ suggests substantial evidence in favor of the alternative. The strongest effects occur across all temporal windows investigated within the immediate vicinity as intensities of 2.6×10^{-5} to 3.4×10^{-4}.

It is interesting that, when considering larger spatial window sizes, a robust (though smaller) effect appears at 20-40 days and 17-19 km for wake sizes. This second bubble of significant effects does not appear in Figure 3.8.

We also examine the matching metrics for the CEM algorithm (Tables 3.3, 3.4, 3.5), namely, common support (empirical overlap between treatment and control

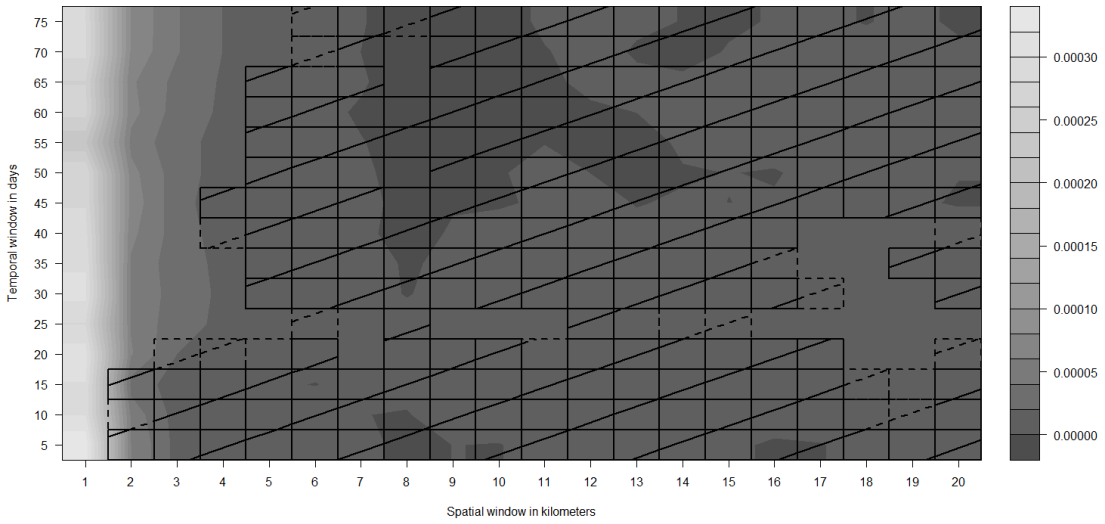


Figure 3.9: This contour plot shows the average causal effect estimates of the difference-in-differences design for the outcome model. The outcome is the change in intensity (expected number of dependent events per unit volume). Treatment intervention events are instances of indiscriminate violence and control intervention events are instances of selective violence. Areas with solid lines correspond to estimates with p-values larger than 0.1, areas with dotted lines correspond to estimates with p-values between 0.05 and 0.1, and clear areas correspond to estimates with p-values less than 0.05.

units) and multivariate $L1$ distance (measured between 0 and 1). Larger values for common support are better, while smaller values for the $L1$ distance are better. Results and summary statistics are also given for each spatio-temporal window with a significant estimated effect ($p < 0.05$). Quantifying the overlaps (same intervention type as 'SO' and mixed intervention types as 'MO') is included to assess the frequency of SUTVA violations. Again, smaller values are better. Overlap is defined as the situation when two (or more) intervention events occur within a wake, meaning that the distance between the intervention events at the center of each wake is less than r. For all spatial windows 5 km or smaller, the percentage of common support ranges from 7.1% to 36.9% after matching (in comparison to 7.8% to 20.4% before matching). There is a similar range of $L1$ values pre- to post-matching, from (0.689, 0.877) to (0.608, 0.885). Lower percentages of common support and higher $L1$ values occur for larger spatial windows within each temporal window. These two measures highlight the effectiveness of CEM in balancing the treated and control intervention event wakes through decreases in $L1$ distance and increases in the common support, across all spatio-temporal windows.

While matching on previous interventions was performed to prevent the impact of SUTVA violations from driving the results (Schutte and Donnay, 2014), there is still a considerable amount of overlap that occurs, more so even for the same intervention event wakes than for mixed, suggesting spatial clustering of treatment intervention events and control intervention events. The amount of overlap (same and mixed) increases as the temporal windows increase; however, the percentage of overlap tends to decrease after matching. The percentage of same-intervention event overlap after matching for spatial windows of 5 km and smaller ranges from 25.8% to 77.8% and the percentage of mixed-intervention event overlap after matching for spatial windows of 5 km or smaller ranges from 6.7% to 61.9%. The higher percentages of overlap correspond to larger temporal windows with the spatial windows of 5 km or less.

When considering spatio-temporal windows less than 25 days and 5 km and smaller, the mixed-intervention event overlap ranges from 6.7% to 42.5% after matching.

		Results		Before Matching						After Matching					
Time	Space	Estimate	P-value	N(control)	N(treat)	L1	%CS	%SO	%MO	N(control)	N(treat)	L1	%CS	%SO	%MO
5	1	0.0003370	0.005	15318	7130	0.69	20.4	29.2	7.4	11273	5950	0.61	36.9	25.8	6.7
10	1	0.0002983	< 0.001	15207	7098	0.71	18.7	36.4	10.6	11071	6014	0.64	31.1	33.6	9.9
15	1	0.0002992	< 0.001	15103	7070	0.73	17.5	40.6	13.1	10591	5785	0.64	30.2	38.2	12.5
20	1	0.0003104	< 0.001	15056	7042	0.74	16.1	43.3	15.2	10283	5630	0.65	28.1	41.0	14.6
	2	0.0000617	0.004	15056	7042	0.79	13.1	53.2	23.0	9584	5870	0.71	22.0	51.0	22.3
	7	0.0000077	0.012	15056	7042	0.86	9.0	70.3	46.5	8386	5743	0.83	11.2	68.4	45.7
	18	0.0000017	0.031	15056	7042	0.88	6.5	82.4	69.2	7464	5193	0.88	6.7	81.4	68.5
	19	0.0000018	0.014	15056	7042	0.88	6.3	83.0	70.3	7311	5162	0.88	6.6	82.2	69.5
25	1	0.0002792	< 0.001	14950	7017	0.75	15.5	45.3	17.0	10182	5706	0.67	26.6	43.2	16.4
	2	0.0000588	0.002	14950	7017	0.79	13.1	55.4	25.6	9368	5699	0.72	20.6	53.3	25.0
	3	0.0000255	0.015	14950	7017	0.82	10.6	61.3	32.5	8911	5683	0.77	16.5	59.3	32.0
	4	0.0000168	0.013	14950	7017	0.85	9.0	65.4	38.9	8955	5808	0.80	13.5	63.5	38.2
	5	0.0000101	0.034	14950	7017	0.85	9.5	68.4	43.3	8718	5692	0.83	11.1	66.7	42.5
	7	0.0000087	0.001	14950	7017	0.87	7.7	72.8	50.6	8346	5659	0.85	9.5	71.3	50.2
	9	0.0000047	0.021	14950	7017	0.89	6.3	76.1	56.9	8021	5549	0.87	8.4	74.8	56.3
	10	0.0000036	0.04	14950	7017	0.88	6.2	77.6	59.4	7823	5487	0.88	7.7	76.3	58.9
	11	0.0000031	0.04	14950	7017	0.86	7.9	78.8	61.6	7677	5465	0.88	7.3	77.6	61.3
	16	0.0000021	0.015	14950	7017	0.89	6.2	83.2	70.1	7527	5234	0.89	6.4	82.3	69.7
	17	0.0000022	0.006	14950	7017	0.89	5.8	84.0	71.6	7363	5208	0.89	6.0	83.1	71.2
	18	0.0000026	< 0.001	14950	7017	0.89	5.7	84.7	72.7	7228	5110	0.90	5.8	83.8	72.3
	19	0.0000023	< 0.001	14950	7017	0.90	5.5	85.2	73.8	7103	5052	0.90	5.7	84.5	73.3
	20	0.0000019	0.002	14950	7017	0.90	5.2	85.8	74.9	6991	5047	0.90	5.5	85.1	74.3

Table 3.3: Results and summary statistics of the MWA with SPP method for the conflict analysis. All combinations of spatial and temporal windows with significant effect estimates are included. Spatio-temporal wake sizes (time in days, space in km) and corresponding estimates with p-values are listed first. The counts of treated and control intervention event type wakes, along with $L1$ distance, percent of common support (CS), and percentages of same intervention event-type overlap (SO) and mixed intervention event-type overlap (MO) are then listed before matching was used and after matching was used.

		Results		Before Matching						After Matching					
Time	Space	Estimate	P-value	N(control)	N(treat)	L1	%CS	%SO	%MO	N(control)	N(treat)	L1	%CS	%SO	%MO
30	1	0.0003071	< 0.001	14789	6964	0.76	14.6	46.8	18.4	9989	5579	0.68	25.0	45.0	18.0
	2	0.0000592	0.002	14789	6964	0.80	12.2	57.0	27.7	9178	5731	0.74	19.1	55.3	27.4
	3	0.0000296	0.004	14789	6964	0.84	9.8	63.0	35.2	8640	5647	0.79	15.0	61.3	34.7
	4	0.0000167	0.013	14789	6964	0.85	9.9	67.0	42.1	8701	5769	0.82	11.7	65.5	41.3
	18	0.0000018	0.009	14789	6964	0.89	5.9	86.2	75.4	7039	5034	0.89	5.8	85.6	75.2
	19	0.0000017	0.01	14789	6964	0.89	5.8	86.7	76.5	6897	5025	0.89	5.6	86.2	76.1
35	1	0.0002968	< 0.001	14702	6921	0.77	14.2	48.2	19.5	9810	5555	0.68	24.3	46.6	19.5
	2	0.0000577	0.002	14702	6921	0.81	11.6	58.4	29.6	8969	5580	0.75	18.4	56.9	29.6
	3	0.0000279	0.006	14702	6921	0.85	9.1	64.5	37.6	8633	5631	0.80	13.8	63.1	37.1
	4	0.0000157	0.013	14702	6921	0.86	9.2	68.5	44.4	8376	5627	0.83	11.2	67.2	44.0
	17	0.0000016	0.025	14702	6921	0.90	5.6	86.7	76.3	6995	5032	0.90	5.3	86.3	76.3
	18	0.0000019	0.006	14702	6921	0.90	5.4	87.3	77.4	6843	4908	0.90	5.2	87.0	77.3
40	1	0.0002896	< 0.001	14570	6861	0.76	14.4	49.3	20.8	9563	5442	0.69	23.4	47.8	20.8
	2	0.0000648	< 0.001	14570	6861	0.82	10.9	59.6	31.4	8785	5483	0.76	17.2	58.2	31.6
	3	0.0000298	< 0.001	14570	6861	0.84	10.8	65.7	39.6	8344	5494	0.81	12.7	64.4	39.5
	17	0.0000016	0.019	14570	6861	0.90	5.2	87.7	78.0	6914	4958	0.91	4.6	87.4	78.2
	18	0.0000016	0.012	14570	6861	0.91	4.9	88.3	79.0	6735	4859	0.91	4.5	88.0	79.1
	19	0.0000012	0.031	14570	6861	0.91	4.7	88.9	79.9	6713	4822	0.92	4.2	88.5	79.9
45	1	0.0002781	< 0.001	14423	6823	0.76	14.0	50.3	22.0	9518	5462	0.69	22.8	48.9	22.1
	2	0.0000725	< 0.001	14423	6823	0.82	10.5	60.6	33.2	8731	5435	0.77	16.4	59.4	33.4
	3	0.0000276	0.002	14423	6823	0.85	10.4	66.8	41.5	8450	5482	0.82	11.9	65.6	41.5
50	1	0.0002793	< 0.001	14273	6781	0.77	13.2	51.1	22.9	9450	5397	0.70	21.2	49.9	23.2
	2	0.0000667	< 0.001	14273	6781	0.83	9.8	61.6	34.4	8589	5358	0.78	15.4	60.4	34.9
	3	0.0000286	< 0.001	14273	6781	0.85	9.6	67.7	43.0	8295	5423	0.83	11.1	66.8	43.2
	4	0.0000120	0.027	14273	6781	0.88	7.8	71.8	50.1	8382	5497	0.87	8.7	71.0	50.4
	8	-0.0000039	0.038	14273	6781	0.89	6.8	80.7	65.5	7607	5431	0.89	6.1	80.1	66.4

Table 3.4: Results and summary statistics of the MWA with SPP method continued.

		Results		Before Matching						After Matching					
Time	Space	Estimate	P-value	N(control)	N(treat)	L1	%CS	%SO	%MO	N(control)	N(treat)	L1	%CS	%SO	%MO
55	1	0.0002234	< 0.001	14118	6741	0.77	12.9	51.9	23.7	9332	5356	0.71	20.6	50.8	24.2
	2	0.0000540	< 0.001	14118	6741	0.82	11.8	62.4	35.6	8529	5328	0.79	14.6	61.4	36.3
	3	0.0000260	< 0.001	14118	6741	0.86	9.2	68.5	44.4	8248	5405	0.84	10.6	67.7	44.7
	4	0.0000141	0.003	14118	6741	0.86	8.5	72.5	51.4	8028	5459	0.87	8.3	71.9	52.0
60	1	0.0002681	< 0.001	13939	6698	0.78	12.7	52.5	24.5	9031	5288	0.71	20.5	51.5	25.2
	2	0.0000669	< 0.001	13939	6698	0.82	11.6	63.0	36.8	8474	5333	0.80	14.0	62.2	37.5
	3	0.0000302	< 0.001	13939	6698	0.86	9.0	69.2	45.7	8124	5316	0.85	10.1	68.6	46.0
	4	0.0000171	< 0.001	13939	6698	0.86	8.2	73.3	52.7	7948	5437	0.88	7.8	72.8	53.2
65	1	0.0002591	< 0.001	13785	6658	0.78	12.5	53.1	25.2	9030	5294	0.72	19.7	52.3	26.0
	2	0.0000634	< 0.001	13785	6658	0.83	11.1	63.6	37.7	8367	5308	0.81	13.3	63.0	38.6
	3	0.0000334	< 0.001	13785	6658	0.84	9.6	69.7	46.8	7956	5301	0.85	9.8	69.3	47.2
	4	0.0000176	< 0.001	13785	6658	0.87	8.0	73.9	53.8	7912	5384	0.86	8.8	73.5	54.6
	8	-0.0000036	0.041	13785	6658	0.90	6.1	82.7	68.9	7526	5288	0.88	6.4	82.4	70.2
70	1	0.0002862	< 0.001	13655	6627	0.78	12.0	53.5	25.9	8921	5219	0.72	19.2	52.8	26.8
	2	0.0000756	< 0.001	13655	6627	0.83	10.6	64.1	38.7	8231	5283	0.81	12.7	63.6	39.6
	3	0.0000414	< 0.001	13655	6627	0.84	9.2	70.3	47.9	7862	5236	0.86	9.4	69.9	48.3
	4	0.0000198	< 0.001	13655	6627	0.85	10.2	74.4	54.7	7679	5284	0.86	8.9	74.1	55.7
	5	0.0000079	0.02	13655	6627	0.87	8.7	77.5	59.5	7413	5199	0.88	7.6	77.1	60.6
	8	-0.0000040	0.022	13655	6627	0.88	7.2	83.1	69.7	7110	5177	0.89	6.0	83.1	71.1
75	1	0.0002821	< 0.001	13488	6585	0.79	11.6	53.9	26.5	8547	5056	0.72	19.3	53.5	27.5
	2	0.0000782	< 0.001	13488	6585	0.83	10.5	64.6	39.5	8131	5198	0.81	12.4	64.3	40.7
	3	0.0000413	< 0.001	13488	6585	0.85	9.2	70.8	48.9	7717	5183	0.86	9.2	70.7	49.6
	4	0.0000196	< 0.001	13488	6585	0.85	10.2	75.0	55.6	7627	5235	0.86	8.4	74.9	57.0
	5	0.0000089	0.007	13488	6585	0.87	8.4	78.1	60.5	7373	5145	0.88	7.1	77.8	61.9

Table 3.5: Results and summary statistics of the MWA with SPP method continued.

3.2.2.2 Conclusions

These results suggest evidence in favor of H_{A2}, that indiscriminate violence leads to more collaboration with U.S. forces. The strongest effects occur across all temporal windows investigated within the immediate vicinity as an intensity of 3.4×10^{-4} (maximum), meaning that there is an expected increase of 0.0003 dependent events (civilian turn-ins) per unit volume of space-time (1 km by 1 km, per day), on average, given the instance of a nearby intervention event is indiscriminate violence compared to selective violence. This relatively weak signal is also noted in Schutte (2017), and the small results may be due to civilians simply not having any information to pass on, or not having the opportunity or means to turn-in unexploded ordnances and information. With such low intensities and considering the smaller temporal windows, there are less mixed overlapping intervention event wakes ($< 30\%$) and thus assuming no interference is appropriate as a beginning step for this particular problem and data analysis.

Potential explanations of the second cluster illustrated in Figure 3.9 may include administrative reasons, given that the data is recorded by individuals. When considering further distances from the point intervention event locations, patrols may not be as frequent, and thus reports from civilians must coincide with opportunity to report unexploded ordnances to U.S. troops. It may also be the case that these are recorded on a monthly basis or there is an administrative deadline at that time. It would be of interest to investigate further the dispersion of intervention events to consider if the location and distance to more populated areas needs to be given even greater attention in how it is incorporated the analysis design. Additionally, further analysis of this problem would consider a more narrow focus in time rather than for the full length of data availability (2004-2009), given that conflicts tend to experience phases of intensity periods.

3.3 Discussion

In this work, we are able to provide further analytical evidence in identifying the radial influence of spatio-temporal wakes of intervention events. These results complement those of Schutte and Donnay (2014). By considering the causal problem in terms of spatial point process theory and utilizing the ideas of *spatial intensity* of dependent events, we bring a novel perspective to causal methodology. In addition, we appropriately incorporate spatial characteristics in the assumptions needed for valid causal inference. These two pieces work together to provide a solid framework for spatio-temporal causal problems in the future. The SPP viewpoint allows us to clearly delineate the spatial radius equal to or closest to the true value. This translates into unbiased estimates of the treatment effect.

Through numerical studies, we clearly illustrate the value of this perspective; however, we do note the limitations of the scenarios presented in this work. It is important to keep in mind that wakes are considered independent with no SUTVA violations; this is not practical to assume in the context of the conflict problem of interest. Additionally, we assume that the radii of both treatment and control wakes are equivalent, which may also not be the case in real life. Further study is needed to incorporate these problems in the SPP context we have begun here.

Chapter 4

Defining the Interference Effect With Spatial Point Process Theory

4.1 Introduction

4.1.1 Spatial-Causal Setting

The fundamental ideas of causal inference in the counterfactual framework rely on key assumptions that can be hard to justify when applied to spatial data and geographic questions. When treatments are defined by geographic features (e.g. via local policy decisions), understanding the effects of these treatments through observational studies can be challenging. The same can be said for disaggregated event data and the effects of specific types of events on future events defined geographically. In such instances, the research design and methods to account for the geographic proximity of events are important to consider.

We carefully review the ideas behind interference in geographic designs (Keele and Titiunik, 2017) under a geographic regression discontinuity (GRD) approach and the matched wake analysis (MWA) approach of Schutte and Donnay (2014) to define a clear, articulated framework for causal inference with spatio-temporal event data. We propose to transform the geographic interference structure for the particular type of data MWA is meant to deal with, incorporating ideas of spatial point process theory. Expanding the methods of the previous chapter, we specifically investigate the impact of overlapping "wakes" as a significant interference effect and separately defining this effect. Additionally, we consider the requirements and limitations for obtaining unbiased treatement effects under this analytic approach for spatio-temporally referenced data.

4.1.2 Types of Questions of Interest

How one poses a causal question including a geographical element can determine whether current methods are feasible to use. The spatio-temporal-causal methodology differs from traditional causal inference in quantifying the effect of treatment

when units are likely to choose any location within a fixed band defining exposure status, or a unit's location choice defines exposure status and is not fixed to an area defining this. Another, very different, question of interest is how to quantify and test the likelihood of receiving spillover effects and considering a causal effect with and without interference. However, all of these questions have important nuances stemming from the geographical elements of the problem and the methods used to obtain the estimated causal effects need to take this into consideration.

4.1.3 Spatially Defined Treatment

When considering geography in a causal context, *local* treatments are of most interest - individuals *here* are considered treated, compared to individuals *there* that are not. This spatial definition of treatment typically manifests itself in one of two ways - units can be defined as treated or not-treated (control) by a spatial measure, or by falling inside or outside some spatial boundary. In either scenario, leveraging the geographical nature of treatment assignment to more appropriately compare units is the common goal of interest.

4.2 Causal Inference in a Spatio-temporal or Spatial Setting

4.2.1 Notation and Framework

Here we lay out the notation and framework necessary to discuss current spatio-causal and spatio-temporal-causal methods and develop a common framework with clarity and statistical rigor. To begin, we do not use individual-level data but rather point process event data. Rather than a set of individuals, we consider (dependent) events with locations and times for which the rate or probability are impacted by

the location and time of other (intervention) events. One of the unique aspects of this spatio-temporal problem is that the defining measure of the treatment receipt impacts the potential outcomes. Below we outline notation following Rubin's causal model (Holland, 1986) and dive into the issues revolving around violations of key causal assumptions.

We refer to an *intervention event* that receives either treatment or non-treatment (control), denoted by Z_i, for $i = 1, \ldots, n$ units. A treatment intervention event is denoted $Z_i = 1$ and control intervention event is denoted $Z_i = 0$. In this context, consider each unit i to represent a spatio-temporal cylinder, or wake, centered around its corresponding intervention event Z_i. Rather than units representing individuals that exist as their own entity and then may be assigned treatment or control, units in this context are geographical areas defined by radius r from an intervention event with its own time series. Thus, units are, in part, defined by the exposure itself. The potential outcome $\Delta_i(Z_i)$, similar to the traditional $Y_i(Z_i)$ notation, represents the change in intensity of *dependent events* over time if the unit experienced the intervention event Z_i. Specifically, $\Delta_i(0) = \lambda_{1i}(0) - \lambda_{0i}(0)$ is the change in intensity, over time, of dependent events within radius r of the center intervention event for unit i, given a control intervention event occured ($Z_i = 0$). The change in intensity, over time, of dependent events within radius r of the centered intervention event for unit i, given a treatment intervention event occurred is $\Delta_i(1) = \lambda_{1i}(1) - \lambda_{0i}(1)$. Both $lambda_{1i}(Z_i)$ and $\lambda_{0i}(Z_i)$ are intensities of dependent events, for the space after intervention event Z_i occurs and the space preceding intervention event Z_i, respectively. The observed outcome can be defined as a linear combination of the potential outcomes and intervention event status, $\Delta_i = Z_i \cdot \Delta_i(1) + (1 - Z_i) \cdot \Delta_i(0)$. A vector of pre-intervention covariates, which can included geographic variables, is denoted $\boldsymbol{X}_i$ for each unit i.

The causal estimand of interest is the average causal effect,

$$ACE = \mathbb{E}\{\Delta_i(1) - \Delta_i(0)\}, \tag{4.1}$$

which is the difference in change in intensity of dependent events over time, given a unit experiences a treatment intervention event versus a control intervention event, on average.

4.2.2 Definition of SUTVA, SUTVA Violations

A major concern when estimating causal effects is the fulfillment of the assumptions that are required to identify estimand(s) of interest. Namely, we require a form of the "no unmeasured confounders" assumption (introduced later in the chapter), and the *stable unit treatment value assumption* (SUTVA) (Rubin, 1980). Particularly in spatial settings, the violation of SUTVA needs to be addressed. A formal statement of SUTVA is,

Assumption 1 (Stable Unit Treatment Value Assumption (SUTVA))
For any unit i,
(a) neither $Y_i(1)$ *nor* $Y_i(0)$ *is affected by the treatment any other unit* $j \neq i$ *received - that is, there is no interference between units; and*
(b) no matter how unit i received treatment $T_i = 1$*, the outcome that would be observed would be* $Y_i(1)$ *and similarly for treatment* $T_i = 0$*; i.e., there are no hidden versions of treatments.*

The manifestation of SUTVA violations differ depending on the context and problem at hand. Within spatial and spatio-temporal causal applications, two types of violations are of particular interest, which we detail below.

Interference Interference occurs when the potential outcomes of unit i are influenced by the treatment received by another unit, violating SUTVA. This is a problem when considering spatial proximity and the influence of social networks (Toulis and Kao, 2013). An example illustration adapted from Kao (2017) is illustrated in Figure 4.1. In the applications of Keele and others (Keele and Titiunik, 2014, 2016, 2017; Keele et al., 2015), the proximity of units in the control area could potentially influence the voting behavior of units located within a close distance in the treated area. This SUTVA violation can result in biased estimates of the treatment effect. The conflict example of Schutte and Donnay (2014) quantifies subsequent instances of turning-in unexploded ordnances by civilians to U.S. troops following indiscriminate insurgent violent events in comparison to selective insurgent violent events. Units, defined as spatio-temporal wakes centered around an insurgent violent event, may interfere with one another. Within this overlapping section of wakes, say wake i experiences a indiscriminate violent event and wake j experiences a selective violent event, there is a potential for individuals to withhold evidence or information of an IED (per the context of the specific problem) in parts of wake j that otherwise wouldn't (in the influence of an indiscriminate violent event). This close proximity leads to potential influence of selective violent event wakes on indiscriminate violent event wakes. While Schutte and Donnay (2014) show that overlap in small amounts still leads to a sound estimation of the treatment effect, the consequences of even a moderate amount of overlap are evident in their work.

Spillover Effect A spillover effect, or indirect exposure effect, occurs when the effect of treatment carries over, or "spills" over, to the control area or control units (or vice versa). An illustration of this is the crime-fighting program application of Verbitsky-Savitz and Raudenbush (2012). An effective crime-lowering program implemented in a certain neighborhood decreases crime due to a shift of this criminal

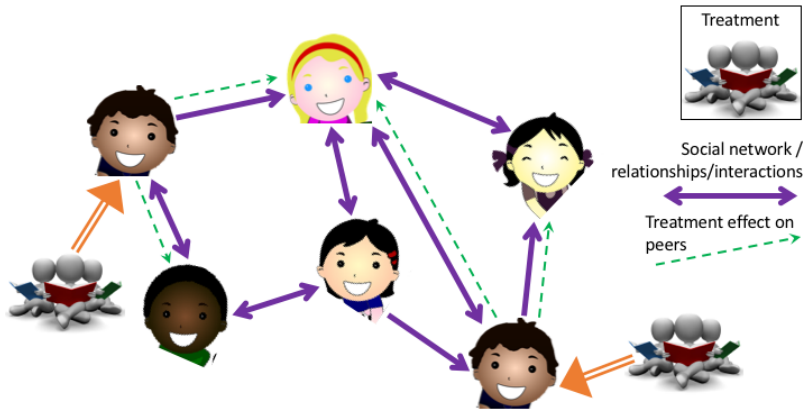


Figure 4.1: Adapted from Kao (2017), an experiment in which interference occurs in a social network. Not only does the treatment affect those that receive it, but it also affects peers through word of mouth and interactions.

activity to contiguous neighborhoods without the program; thus, the implementation of the intervention in one area increases the crime rate in contiguous areas. In general, the spillover effect can be seen when individuals defined as receiving the control or located in the control area actually receive the treatment, or may be located in the treatment area, but the magnitude of the effect differs from the effect of isolated treatment (or vice versa). This group of individuals does not receive the true effect of treatment, but are located in such a way that they receive part of both the defined treatment and control versions of exposure. This also violates SUTVA. The overlap in the conflict data setting (Schutte and Donnay, 2014) could also be seen to fit this type of SUTVA violation. If a wake centered around a treatment intervention event occurs within close proximity to wakes centered around a control intervention event, the control intervention event wakes could experience a spillover effect of the nearby treatment intervention event wake. However, we do not consider this scenario going forward.

4.2.3 Geographic Regression Discontinuity

The geographic regression discontinuity (GRD) design (Keele and Titiunik, 2014, 2016, 2017; Keele et al., 2015) gives rise to a matching framework that incorporates geographic proximity and observed covariates in a flexible manner. The goal is to minimize spatial distance between treated and control pairs while still preserving balance for observed covariates. This approach works best when the intended causal question of interest pertains to inferences regarding the effects of geographically varying treatments, where units in the treated area must be compared to units in the control area. While one can condition on a set of observed pretreatment covariates, there is still a potential for unmeasured confounders (Keele et al., 2015; Keele and Titiunik, 2017). GRD proposes to combine this conditioning on observed covariates with the exploitation of geographic proximity, detailed in Keele and Titiunik (2014); Keele et al. (2015). The idea behind this work is to use the regression discontinuity (RD) framework, where counterfactuals are the treated and control groups near the boundary, making local treatment effects identifiable. Units either sort around a boundary between the treated and control areas with error, or the boundary is arbitrary between the areas.

This GRD design deals with the issue of strong self-selection around some border of interest. The proposed solution is to assume that after conditioning on pretreatment covariates $\boldsymbol{X}$, treatment assignment is as-if randomized (independent) for those who live near the boundary. Geographic distance between treated and control observations to the boundary is minimized while enforcing balance in pretreatment covariates. The conditional geographic treatment ignorabilitity in local neighborhood assumption (Keele and Titiunik, 2017) states that a small neighborhood exists where potential outcomes and treatment assignment are conditionally independent given pretreatment covariates,

Assumption 2 (Conditional Mean Independence in Local Neighborhood (Keele and

Titiunik, 2017))

For all units that reside in a narrow area for which the boundary of the spatio-temporal cylinder(s) does not encounter any heterogeneity,

$$\mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i, Z_i = 1\} = \mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i, Z_i = 0\}$$

$$\mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 1\} = \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 0\}$$

The work of Keele and Titiunik (2017) generalizes and expands that of Lee (2008) and Cattaneo et al. (2015) to consider a small geographic neighborhood around the boundary separating treatment and control areas, such that the condition of as-if randomization, or independence, holds after conditioning on pretreatment covariates. Lee (2008) and Cattaneo et al. (2015) note that the emphasis on "after conditioning on pretreatment covariates" is important, since this as-if randomization may not hold unconditionally for the small geographic neighborhood around the boundary separating treatment and control areas. This assumption allows for the identification of a treatment effect at the boundary points and points included in the geographic neighborhood where it holds. The form of this assumption makes sense to use when treatment assignment is based on geography. Additionally, it is weaker than the assumption when only conditioning on observables (Assumption 3), as it requires conditional independence for a subset of the population.

Assumption 3 (Conditional Geographic Treatment Ignorability)

$$\{\Delta_i(0), \Delta_i(1)\} \perp\!\!\!\perp Z_i|\boldsymbol{X}_i.$$

4.2.4 Our Goals

It is clear that while there has been some work done to understand causal implications in spatial settings, the setting discussed in Schutte and Donnay (2014) can benefit

from a structured statistical framework. Additionally, there is some overlap between the current ideas and problem structure of MWA and interference in a geographic setting as has been explored by Keele and Titiunik (2017) and others. We work to bring these together in Section 4.3. As application-specific extensions of causal inference methods to spatial and spatio-temporal settings emerge, it is important to define common elements to identify a core basis for spatial-causal inference. The questions we are interested in answering deal with the impact of overlapping spatio-temporal cylinders on the treatment effects of interest. We will define the "*interference effect*" and the "*interference-free effect*" in this particular setting. Interestingly, we note that the implications of spatio-temporal SUTVA violations are themselves spatial; that is, the violations are *local* in space and time, may be limited to particular observations in particular locations, and as such, merit careful examination in this context.

4.3 Spatio-temporal Manifestations of SUTVA Violations

In this section, we outline the necessary assumptions and framework for scenarios in which SUTVA may be violated due to the spatio-temporal nature of the data and question of interest. Below is a chart of the problem from least complexity to most complexity, to the extent that will be addressed in this paper (Table 4.1). We note that this does not include all potential scenarios, but is merely a start to the vast reaches of the depth of this problem. This table begins with the inclusion of both treated and control intervention events rather than considering scenarios of only one type of intervention event, given the nature of the problem and inference of interest. Additionally, given our motivating application, we assume no compounded treatment going forward in the rest of this work. That is, we do not attempt to address the implications of compounded treatment in addition to the those of overlapping differing

treatments. The focus will be on scenarios in which overlapping control and treatment intervention wakes may potentially occur.

What is Estimated	Radius	Potential Overlap	Assumptions	Scenario
Differences in Intensities	Constant, Known	No	$r_T = r_C = r$ is known - no SaTScan elements necessary; Consistency is fulfilled; SUTVA holds; Conditional mean independence in local neighborhood	**1A**
		Yes	$r_T = r_C = r$ is known - no SaTScan elements necessary; Consistency is fulfilled; SUTVA does not hold-define interference and interference-free effect; Conditional mean independence in local neighborhood	**1B**
Differences in Intensities	Constant, Unknown	No	$r_T = r_C = r$ is found via SaTScan-esque technique; Consistency is fullfilled; SUTVA holds; Conditional mean independence in local neighborhood	**2A**
		Yes	$r_T = r_C = r$ is found via SaTScan-esque technique; Consistency is fulfilled; SUTVA does not hold-define interference and interference-free effect; Conditional mean independence in local neighborhood	**2B**

Table 4.1: An outline of the complexity of scenarios addressed in this chapter.

4.3.1 Complications of Spatial Setting for Definitions of Elements of Standard Causal Framework: SUTVA, Interference, and Spillover

While MWA provides a novel approach to evaluating and estimating causal relationships of conflict events that could be used for other types of georeferenced event data, there are some limitations to the approach. Namely, the spatio-temporal cylinders defining the units that receive treated or control intervention events can potentially overlap with one another, violating SUTVA. Violating this assumption could lead to biased effect estimates. Two scenarios could result from MWA: 1) radii from multiple units of the same intervention event type overlap in space and time, biasing estimates upward (assuming a positive treatment effect); or 2) a unit of each of a treatment intervention and control intervention could overlap and bias the estimated treatment effect downward. The chances for overlap decrease as the intensity of intervention events decreases and size of the wakes decreases.

In the context of their polling problem of interest, Keele and others consider

the possibility for interference, resulting in a violation of SUTVA. However, for this application there has been shown to be little evidence of treatment spillovers even withinin households (Sinclair et al., 2012). Thus, SUTVA violation in the form of voters in control area encouraging neighbors in treated area to vote, or the likelihood of interference across voters even within adjacent households, is slim. Regardless, if such a violation occured it would bias the effect towards 0, and positive effects would be conservative estimates.

4.3.2 Scenario 1A: When the Spatio-temporal Wake Radius is Known

We begin with similar notation as in Section 4.2.1 above. Let Z_i represent a treatment or control intervention event for unit i, $i = 1, \ldots, n$ with $Z_i = 1$ a treatment intervention event and $Z_i = 0$ a control intervention event. The potential outcomes $\Delta_i(0), \Delta_i(1)$ again represent the change in intensities of dependent events over time given unit i experiences a control intervention event and unit i has a treatment intervention event occur, respectively. The observed outcome is $\Delta_i = Z_i \cdot \Delta_i(1) + (1 - Z_i) \cdot \Delta_i(0)$. Each unit i has an associated vector of pre-intervention event covariates and risk factors, $\boldsymbol{X}_i$. Assume that consistency (Cole and Frangakis, 2009) holds, in which a unit assigned to treatment intervention event has an observed outcome that is the potential outcome under the treatment intervention event (and vice versa). Formally,

$$\Delta_i = \Delta_i(1) \text{ if } Z_i = 1 \text{ and } \Delta_i = \Delta_i(0) \text{ if } Z_i = 0.$$

We also assume that SUTVA (Assumption 1) initially holds - that there is no interference between units as well as no hidden versions of intervention event types. In this scenario, there is no potential violation of SUTVA to consider. To identify the

causal estimand of interest and obtain unbiased estimates, an ignorability assumption is necessary. Typically, the assumption states that treatment is ignorable conditional on a set of observed covariates, meaning that there are no unmeasured confounders. We first consider the version of this assumption that is appropriate for spatial settings (Keele and Titiunik, 2017) introduced earlier,

Assumption 3 (Conditional Geographic Treatment Ignorability)

$$(\Delta_i(0), \Delta_i(1)) \perp\!\!\!\perp Z_i | \boldsymbol{X}_i.$$

However, as Keele et al. (2015) point out, there is no *a priori* reason to assume that geographic variation of treatment in such a way is justified, or that it is justified at every possible location and time. As such, instead assume that treatment assignment can be considered as-if randomized as in the following assumption, also previously introduced.

Assumption 2 (Conditional Mean Independence in Local Neighborhood (Keele and Titiunik, 2017))
For all units that reside in a narrow area for which the boundary of the spatio-temporal cylinder(s) does not encounter any heterogeneity,

$$\mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i, Z_i = 1\} = \mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i, Z_i = 0\}$$

$$\mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 1\} = \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 0\}$$

That is, assume that there exists a small neighborhood around the set of wakes, where the average potential outcomes are mean independent of intervention event type, given the covariates. In this setting, the causal estimand of interest is identifiable:

$$\begin{aligned}\tau &= \mathbb{E}\{\Delta_i(1) - \Delta_i(0)\} \\ &= \mathbb{E}\left[\mathbb{E}\{\Delta_i(1) - \Delta_i(0)|\boldsymbol{X}_i\}\right] \\ &= \mathbb{E}\left[\mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i\} - \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i\}\right] \\ &= \mathbb{E}\left[\mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i, Z_i = 1\} - \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 0\}\right] \\ &= \mathbb{E}\left[\mathbb{E}\{\Delta_i|\boldsymbol{X}_i, Z_i = 1\} - \mathbb{E}\{\Delta_i|\boldsymbol{X}_i, Z_i = 0\}\right]\end{aligned} \tag{4.2}$$

For the Difference-in-Differences (DD) design of Schutte and Donnay (2014), the observed outcome is then modeled by

$$\begin{aligned}\Delta_i &= \beta_0 + \beta_1 n_{pre} + \tau Z_i + \epsilon_i \\ &= \boldsymbol{X}_i^T \boldsymbol{\beta} + \tau Z_i + \epsilon_i\end{aligned}$$

where Δ_i is the observed intensity of dependent events after intervention event, n_{pre} is the number of dependent events prior to the intervention event occuring, and

$$\begin{aligned}\epsilon_i =& Z_i \left[\mathbb{E}\{\Delta_i(1) - \Delta_i(0)|\boldsymbol{X}_i\} - \mathbb{E}\{\Delta_i(1) - \Delta_i(0)\}\right] \\ &+ (1 - Z_i)\left[\Delta_i(0) - \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 0\}\right] \\ &+ Z_i \left[\Delta_i(1) - \mathbb{E}\{\Delta_i(1)|\boldsymbol{X}_i, Z_i = 1\}\right].\end{aligned}$$

Under consistency and Assumptions 1 and 2, $\mathbb{E}\left(\epsilon_i|\boldsymbol{X}_i, Z_i\right) = 0$ and τ can be consistently estimated.

4.3.3 Scenario 1B: Accounting for Potential Overlap as Interference

When intervention events occur close to one another in space, there is a potential of spillovers or interference between units. Treated intervention event wakes may be influenced by control intervention wakes and vice versa, due to geographic proximity. Given the context of this problem, we assume that control intervention wakes influence treatment intervention wakes. As such, begin by assuming that interference is one-sided from control to treatment intervention wakes, similar to Verbitsky-Savitz and Raudenbush (2012) and Keele and Titiunik (2017). It is advantageous to narrow down the scope of potential outcomes that are truly possible and constructive for defining *a priori* causal estimands of interest. Allowing for any interference pattern means that an event's potential outcomes must be expressed as the full vector $\Delta_i(\boldsymbol{Z})$, so that each event's treatment status may be allowed to depend on the treatment status of every other event. Such a large number of causal effects per event makes this problem implausible.

Following Keele and Titiunik (2017), assume each intervention event's potential outcomes depend on its own treatment status, as well as any intervention event of opposite treatment status located within a specified distance of the event i. We use a modified functional indicator of interference (Keele and Titiunik, 2017), $g_i(\boldsymbol{Z};\delta)$, for fixed values of δ =distance, which defines the radius within which event i receives interference.

$$g_i(\boldsymbol{Z};\delta) = \mathbb{I}(Z_i = 1)\mathbb{I}(N_{i\delta} \geq 1),$$

where

$$N_{i\delta} = \sum_{j=1}^{n} \mathbb{I}(d(i,j) \leq \delta)\mathbb{I}(Z_i \neq Z_j),$$

$\mathbb{I}(\cdot)$ is the indicator function, $d(i,j)$ is a measure of distance between intervention event i and j's locations, and $\delta \in \mathbb{R}$ is a pre-specified distance (2 times the radius)

around i's location. If i is a treatment intervention event, then $g_i(\boldsymbol{Z};\delta) = 1$ if there is at least one control intervention event within a δ radius of the intervention event of wake i's location, and $g_i(\boldsymbol{Z};\delta) = 0$ if there are no control intervention events within a δ radius around i's location. Because we assume that interference is one-sided, $g_i(\boldsymbol{Z};\delta) = 0$ for any control intervention event. This indicator function then allows the potential outcomes for intervention event i to depend on other intervention events treatment status only through the function $g_i(\boldsymbol{Z};\delta)$, in addition to its own treatment status Z_i. The potential outcomes in this interference framework are defined by $\Delta_i(\boldsymbol{Z}) = \Delta_{i;\delta}(Z_i, g_i(\boldsymbol{Z};\delta))$. For this geographic-based interference, there are now 3 potential outcomes:

$\Delta_{i;\delta}(1,1)$ =intervention event i is treated and experiences interference (overlap is present) - the wake of a control intervention event is within δ km of the wake of i's location.

$\Delta_{i;\delta}(1,0)$ =intervention event i is treated and does not experience interference (no overlap) - there are no control intervention event wakes within δ km of the wake of i's location.

$\Delta_{i;\delta}(0,0) = \Delta_i(0)$, intervention event i is control and $g_i(\boldsymbol{Z};\delta) = 0$ for all control intervention events.

The observed outcome for each wake can only take one of these values, and is represented as a combination of the assigned intervention event and interference indicator:

$\Delta_i = \Delta_{i;\delta}(1,1)$ if $Z_i = 1$ and $g_i(\boldsymbol{Z};\delta) = 1$

$\Delta_i = \Delta_{i;\delta}(1,0)$ if $Z_i = 1$ and $g_i(\boldsymbol{Z};\delta) = 0$

$\Delta_i = \Delta_i(0)$ if $Z_i = 0$

This notation allows for a causal estimand that is comparable to the average treatment effect (4.1) of the population as the *average treatment effect in the absence of interference*, also known as the "*interference-free effect*",

$$\tau_T = \mathbb{E}\{\Delta_{i;\delta}(1,0) - \Delta_i(0)\}$$

This comparison considers the two potential outcomes that would be observed under treatment and control intervention events in the absence of interference, and captures the average effect of treated intervention events when they are geographically separate from control intervention events and thus do not receive spillover. The "*interference effect*" capturing the *average effect of interference* or spillovers on treated units can also be defined,

$$\tau_S = \mathbb{E}\{\Delta_{i;\delta}(1,1) - \Delta_{i;\delta}(1,0)\}$$

This comparison quantifies the effect of treated intervention events who are geographically close (wake overlap) to a control intervention event in relation to those treated intervention events that are geographically separate from control intervention events. One can test for the presence of interference ($H_0 : \tau_S = 0$), since in the absence of interference $\Delta_{i;\delta}(1,1) = \Delta_{i;\delta}(1,0) = \Delta_i(1)$ for all i, implying that $\tau_S = 0$ (Keele and Titiunik, 2017).

Neither τ_T nor τ_S are identifiable under the current set of assumptions that hold for Scenario 1A. In order for both of these estimands to be identifiable, additional assumptions must be made to address the geographic nature of the problem, and deal with the consideration of one observed outcome for every i from three potential outcomes. Consistency and Assumption 2 still hold in this scenario, but another assumption must be made with regard to the interference areas, since SUTVA no longer holds.

Assumption 4 (As-if Random Geographic Location within Interference Areas (Keele

and Titiunik, 2017))

$$\begin{aligned}
\mathbb{E}\{\Delta_{i;\delta}(1,1)|\boldsymbol{X}, Z_i = 1, g_i(\boldsymbol{Z};\delta) = 1\} &= \mathbb{E}\{\Delta_{i;\delta}(1,1)|\boldsymbol{X}, Z_i = 1, g_i(\boldsymbol{Z};\delta) = 0\} \\
\mathbb{E}\{\Delta_{i;\delta}(1,0)|\boldsymbol{X}, Z_i = 1, g_i(\boldsymbol{Z};\delta) = 1\} &= \mathbb{E}\{\Delta_{i;\delta}(1,0)|\boldsymbol{X}, Z_i = 1, g_i(\boldsymbol{Z};\delta) = 0\}
\end{aligned}$$

This requires that falling in the interference region in the treatment intervention event wake is mean independent of potential outcomes, conditional on pretreatment covariates. This is weaker than assuming that each unit is randomly assigned to a geographic location in a treatment intervention event wake, so that whether or not they fall in the interference region is unrelated to their potential outcomes. With these assumptions established, both effects are identifiable,

$$\begin{aligned}
\tau_T(\boldsymbol{X}) &= \mathbb{E}\{\Delta_{i;\delta}(1,0) - \Delta_i(0)|\boldsymbol{X}\} \\
&= \mathbb{E}\{\Delta_{i;\delta}(1,0)|\boldsymbol{X}, Z_i = 1, g_i(\boldsymbol{Z};\delta) = 0\} - \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}, Z_i = 0\} \\
&= \mathbb{E}\{\Delta_i|\boldsymbol{X}, Z_i = 1, g_i(\boldsymbol{Z};\delta) = 0\} - \mathbb{E}\{\Delta_i|\boldsymbol{X}, Z_i = 0\}
\end{aligned} \tag{4.3}$$

$$\begin{aligned}
\tau_S(\boldsymbol{X}) &= \mathbb{E}\{\Delta_{i;\delta}(1,1) - \Delta_{i;\delta}(1,0)|\boldsymbol{X}\} \\
&= \mathbb{E}\Delta_{i;\delta}(1,1)|\boldsymbol{X}, Z_i = 1, g_i(\boldsymbol{Z};\delta) = 1\} - \mathbb{E}\{\Delta_{i;\delta}(1,0)|\boldsymbol{X}, Z_i = 1, g_i(\boldsymbol{Z};\delta) = 0\} \\
&= \mathbb{E}\{\Delta_i|\boldsymbol{X}, Z_i = 1, g_i(\boldsymbol{Z};\delta) = 1\} - \mathbb{E}\{\Delta_i|\boldsymbol{X}, Z_i = 1, g_i(\boldsymbol{Z};\delta) = 0\}
\end{aligned} \tag{4.4}$$

The estimators of these two parameters are now consistent for the DD design. Let $G_i \equiv g_i(\boldsymbol{Z};\delta)$. Then

$$\begin{aligned}
\Delta_i &= \gamma_0 + \gamma_1 n_{pre} + \tau_T Z_i + \tau_S G_i \cdot Z_i + \nu_i \\
&= \boldsymbol{X}_i^T \boldsymbol{\gamma} + \tau_T Z_i + \tau_S G_i \cdot Z_i + \nu_i
\end{aligned}$$

where Δ_i is the change in intensity of dependent events before and after an inter-

vention event, n_{pre} is the number of dependent events prior to the intervention event occuring, and

$$\begin{aligned}\nu_i = (1 - Z_i)\left[\Delta_i(0) - \mathbb{E}\{\Delta_i(0)|\boldsymbol{X}_i, Z_i = 0\}\right] + \\ Z_i \cdot (1 - G_i)\left[\Delta_{i;\delta}(1,0) - \mathbb{E}\{\Delta_{i;\delta}(1,0)|\boldsymbol{X}_i, Z_i = 1, G_i = 0\}\right] + \\ Z_i \cdot G_i\left[\Delta_{i;\delta}(1,1) - \mathbb{E}\{\Delta_{i;\delta}(1,1)|\boldsymbol{X}_i, Z_i = 1, G_i = 1\}\right] + \\ Z_i\left[\mathbb{E}\{\Delta_{i;\delta}(1,0) - \Delta_i(0)|\boldsymbol{X}\} - \mathbb{E}\{\Delta_{i;\delta}(1,0) - \Delta_i(0)\}\right] + \\ Z_i \cdot G_i\left[\mathbb{E}\{\Delta_{i;\delta}(1,1) - \Delta_{i;\delta}(1,0)|\boldsymbol{X}\} - \mathbb{E}\{\Delta_{i;\delta}(1,1) - \Delta_{i;\delta}(1,0)\}\right].\end{aligned}$$

Under consistency and Assumptions 2 and 4, $\mathbb{E}\left(\nu_i|\boldsymbol{X}, Z_i, G_i\right) = 0$ and the parameters can be consistently estimated.

4.3.3.1 Derivation of Spatio-temporal Overlap Where the Interference Effect Lives

We can easily derive the corresponding volume of the overlapping cylinders that experience this interference effect by returning to some principles of geometry. Consider overlapping cylinders with base circles $\boldsymbol{A}$ and $\boldsymbol{B}$, both with radius r. The absolute value of the total distance d between the centers of circles $\boldsymbol{A}$ and $\boldsymbol{B}$ must be less than or equal to two times the radius r, $|d| \leq 2r$. Figure 4.2 illustrates the two-dimensional version of this problem.

To simplify ideas, we work in the two-dimensional case using the bases of the cylinders, and then apply this to the overlapping volume of the cylinders as opposed to area. Two isosceles triangles, $\triangle ADE$ and $\triangle BDE$, are formed by the center of each circle and the two points of intersection. Drawing a line b from $\overline{DE}$ creates the shared base of the triangles. The distance from the center of each circle to b is $d/2$, and the angles formed by the radii connecting the center of each circle to the intersection points is represented by θ. The goal is to find the total area of the blue +

green shaded areas, which in turn leads to the volume with this corresponding base. This will allow us to represent the intensity of this overlapping region and calculate the interference effect.

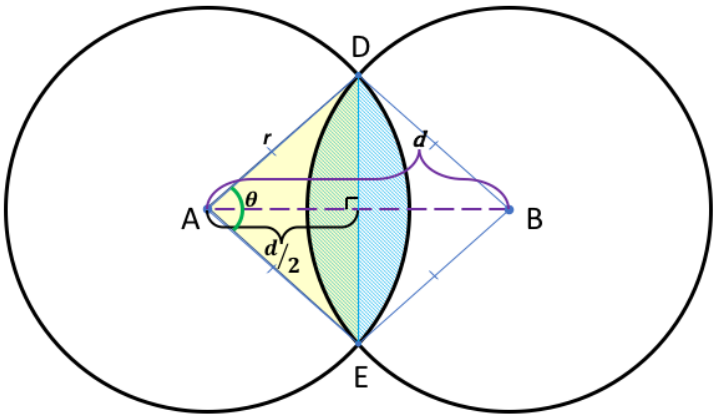


Figure 4.2: Area of overlapping circles **A** and **B** with equal radii r. Points D and E represent the intersection points of the two cylinders. The dashed horizontal line, length d, is the total distance between the centers of the two cylinders. The area of the two circles shaded in blue/green represents the base of the overlapping cylindrical area in which the interference effect lives.

First, the formula for the area of the sector (yellow+green+blue areas) is

$$A_{sector} = \frac{\theta}{2} r^2$$

where the central angle θ, measured in radians, can be found using the yellow isosceles triangle, $\triangle ADE$:

$$\begin{aligned} \cos\left(\tfrac{\theta}{2}\right) &= \tfrac{\frac{d}{2}}{r} \\ \cos\left(\tfrac{\theta}{2}\right) &= \tfrac{d}{2r} \\ \tfrac{\theta}{2} &= \cos^{-1}\left(\tfrac{d}{2r}\right) \\ \theta &= 2\cos^{-1}\left(\tfrac{d}{2r}\right) \end{aligned}$$

such that the area of the sector is represented as:

$$\begin{aligned} A_{sector} &= \tfrac{\theta}{2} r^2 \\ &= \tfrac{2\cos^{-1}\left(\frac{d}{2r}\right)}{2} r^2 \\ &= r^2 \cos^{-1}\left(\tfrac{d}{2r}\right) \end{aligned}$$

Using trigonometry, the area of the yellow isosceles triangle, $\triangle ADE$, is then:

$$\begin{aligned} \sin\left(\tfrac{\theta}{2}\right) &= \tfrac{\frac{b}{2}}{r} \\ b &= 2r \sin\left(\tfrac{\theta}{2}\right) \end{aligned}$$

$$\begin{aligned} \cos\left(\tfrac{\theta}{2}\right) &= \tfrac{h}{r} \\ h &= r \cos\left(\tfrac{\theta}{2}\right) \end{aligned}$$

$$\begin{aligned} A_{\triangle} &= \tfrac{1}{2} bh \\ &= \tfrac{1}{2} \left\{2r \sin\left(\tfrac{\theta}{2}\right)\right\} \left\{r \cos\left(\tfrac{\theta}{2}\right)\right\} \\ &= r^2 \sin\left(\tfrac{\theta}{2}\right) \cos\left(\tfrac{\theta}{2}\right) \\ &= \tfrac{1}{2} r^2 \sin\theta \end{aligned}$$

Or by way of the Pythagorean theorem,

$$\begin{aligned} \left(\tfrac{b}{2}\right)^2 + \left(\tfrac{d}{2}\right)^2 &= r^2 \\ \tfrac{b^2}{4} + \tfrac{d^2}{4} &= r^2 \\ b^2 + d^2 &= 4r^2 \\ b &= \sqrt{4r^2 - d^2} \end{aligned}$$

$$\begin{aligned} A_{\triangle} &= \tfrac{1}{2} bh \\ &= \tfrac{1}{2} \sqrt{4r^2 - d^2} \tfrac{d}{2} \\ &= \tfrac{d}{4} \sqrt{4r^2 - d^2} \end{aligned}$$

This leads to the area of the sector's segment (blue area), and thus the total over-

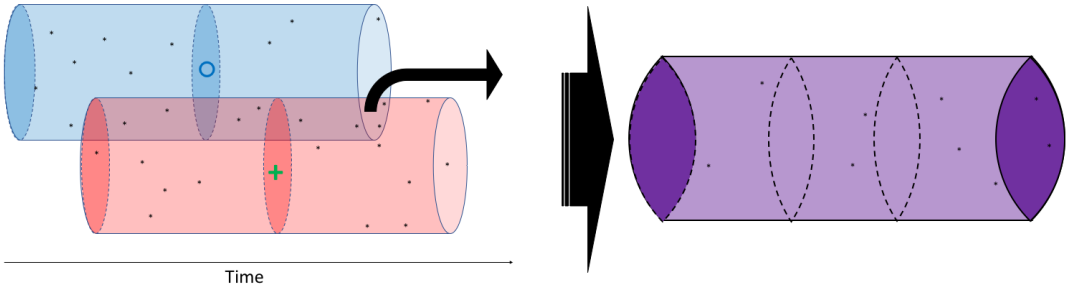
lapping section of the circles (green+blue areas), measured in radians (θ) or distance (d):

$$A_{overlap} = r^2\left(\theta - \sin\theta\right) \equiv 2r^2\cos^{-1}\left(\frac{d}{2r}\right) - \frac{d}{2}\sqrt{4r^2 - d^2}$$

Finally, given the formula for the volume of a cylinder, $V = \pi r^2 h$, the area of the overlapping segments $A_{segment}$ is substituted for the area of a circle πr^2 as the base of the overlapping cylinders to obtain the volume:

$$V = r^2 h\left(\theta - \sin\theta\right) \equiv 2r^2 h\cos^{-1}\left(\frac{d}{2r}\right) - \frac{d}{2}h\sqrt{4r^2 - d^2} \tag{4.5}$$

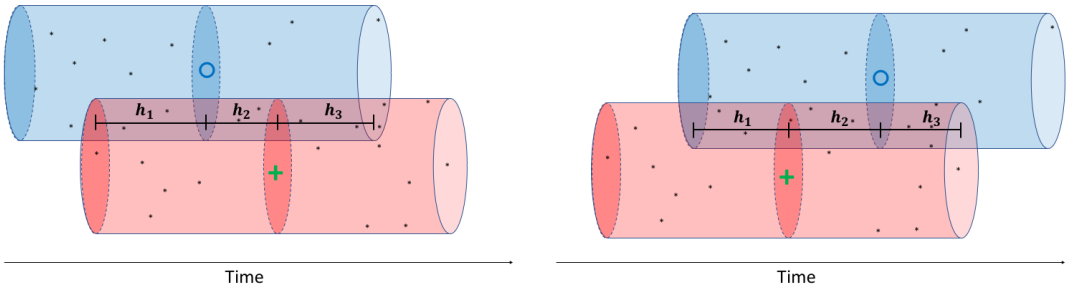
Knowledge of the analytical form of the overlapping volume of the wakes is important for the spatial point process theory that we consider, as it is a key element for defining the intensity in this overlapping portion. Specifically, it allows us to define the post-treatment intervention event intensity for unit i experiencing interference, $\lambda_{1i}(1,1)$, which differs from the post-treatment intervention event intensity for unit i not experiencing interference $\lambda_{1i}(1,0)$. The general scenario is illustrated in Figure 4.3.



(a) (b)

Figure 4.3: An illustration of the volume of overlapping wakes.

The pre-treatment intervention event intensity is the same for unit i regardless of interference indicator, that is, $\lambda_{0i}(1,1) = \lambda_{0i}(1,0)$. There are two situations that can arise for the post-intervention event intensity (Figure 4.4).



(a) (b)

Figure 4.4: An illustration of overlapping wakes that could occur for calculating the interfered-with treatment wake intensities. Scenario (a) represents the overlap that occurs if the control intervention wake occurs first in time, and scenario (b) represents the overlap that occurs if the treatment intervention wake occurs first in time.

The intensity for unit i after experiencing a treatment intervention event and interference is determined by which intervention event occurred first temporally. Either a portion of the time post-treatment intervention event or the full length of time post-intervention event is subjected to this interference and thus contributes differing portions to the intensity $\lambda_{1i}(1,1)$. Thus, we can represent $\lambda_{1i}(1,1)$ in terms of $\lambda_{1i}(1,0)$, $\lambda_{1j}(0)$, and $\lambda_{0j}(0)$ for $i \neq j$:

$$\lambda_1(1,1) = \begin{cases} \left[(\lambda_1(1,0)\cdot v) + \left\{\lambda_1(0)\cdot\left(2r^2 h_3 \cos\left(\frac{d}{2r}\right) - \right.\right.\right. & \\ \left.\left.\left. \frac{d}{2}h_3\sqrt{4r^2-d^2}\right)\right\}\right]/v, & \text{if date(Control) } < \text{date(Treatment)} \\ \left[(\lambda_1(1,0)\cdot v) + \left\{\lambda_1(0)\cdot\left(2r^2 (h_2+h_3) \cos\left(\frac{d}{2r}\right) - \right.\right.\right. & \\ \left.\left.\left. \frac{d}{2}(h_2+h_3)\sqrt{4r^2-d^2}\right)\right\}\right]/v, & \text{if date(Control) } > \text{date(Treatment)} \end{cases}$$

4.3.4 Scenario 2A: Proposing the Spatio-temporal Wake Radius

This scenario is fleshed out in Chapter 3, focusing on the impact of mis-specifying the radius of wakes and the potential bias that is introduced. We do not re-discuss this scenario here, but refer the reader back to Chapter 3.

4.3.5 Scenario 2B: How to Account for Potential Overlap as Interference with a Proposed Spatio-temporal Wake Radius

This scenario combines the challenges of misspecified wake radius with potentially overlapping wakes. The underlying foundation mirrors that of Scenario 1B when considering the interference effect and interference-free effect. However, when applying this method to a real-world problem, the true radius of influence for the wakes is unknown and must be chosen. This scenario is a step towards a more realistic setup to use for a data analysis, and is illustrated in the numerical studies below.

4.3.5.1 Further Considerations Regarding Wake Overlap

In theory, the configuration of wakes can be accounted for in 4 different schemes. Figure 4.5 presents a visual representation of these 4 possibilities, where it is possible for no overlap (a), overlap of two intervention events of the same treatment assignment ($Z_i = Z_j = 1$ or $Z_i = Z_j = 0$) (b and c), or one intervention event of each treatment assignment occurring (d). Taking the conflict data that is the underlying driver of this method in the prior literature, we limit this work to no overlap (Figure 4.5 (a)) and the type of overlap pictured in Figure 4.5 (d). Both Figure 4.5 (b) and (c) illustrate *compounded treatment*, which is outside the scope of what we consider here. The overlapping area between wakes in Figure 4.5 (b) experiences an *enhanced effect*, while the overlapping area between wakes in Figure 4.5 (d) experiences a *diluted effect*.

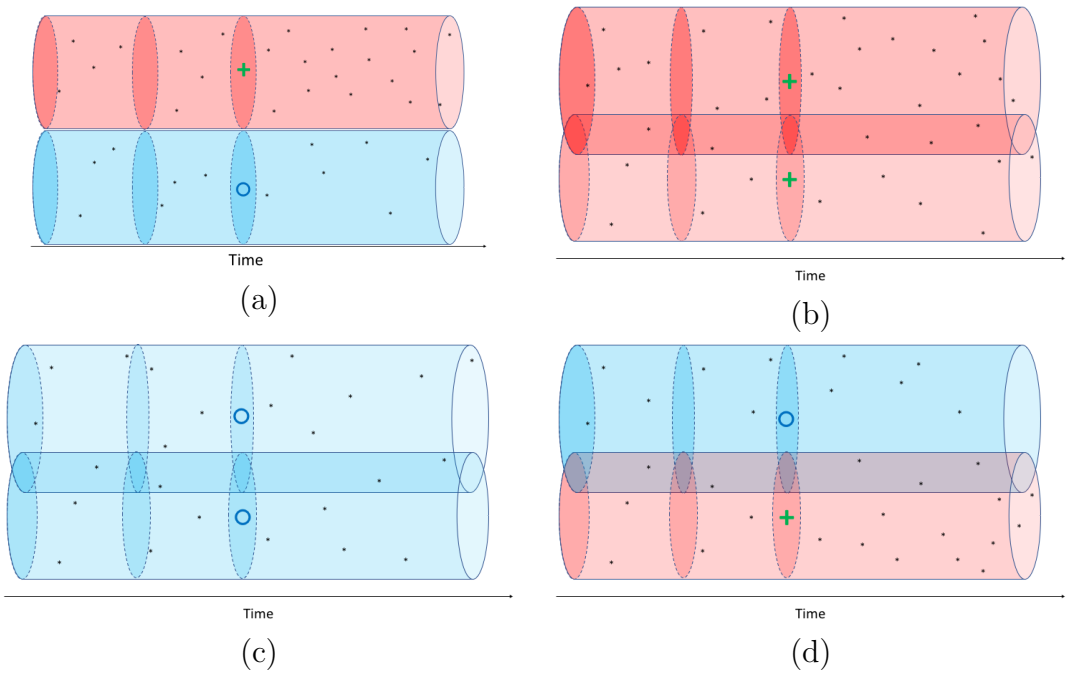


Figure 4.5: The 4 intervention event wake configurations, where red represents a treatment intervention event (indiscriminate insurgent violence with civilian casualties) and blue represents a control intervention event (selective insurgent violence with no civilian casualties).

4.4 Numerical Studies

4.4.1 Simulations

Simulations were conducted mimicking the conflict data from Schutte and Donnay (2014) and the underlying setup is similar to that of Chapter 3. A total of 150 treatment intervention events and 300 control intervention events were used in order to facilitate large enough sample sizes for each of the interference and interference-free effect estimates. The end date of treatment and control events, as well as their associated dependent events, was shortened from one year to approximately 9 months. The total area that is considered is reduced to facilitate increases of overlapping wakes to illustrate this method (approximately 111 km by 111 km). We also consider much larger wake radii, from about 11.11 km to 17.78 km, and a larger temporal window of 10 days, again to more clearly illustrate the proposed method. Background dependent events were generated following a Poisson distribution with mean intensity 0.0002 times the volume of the whole cube minus the summation of the volume of the cylindrical wakes. The background dependent events were set to avoid overlapping placement with dependent events occurring within wakes of interest and avoid skewing estimation of intensities. A random Poisson distribution with mean equal to intensity times volume of the cylinder was used to generate the number of dependent events pre- and post- intervention event. We assume a pre-intervention event intensity of 0.1, with an equivalent post-intervention event intensity for controls and an increased post-intervention event intensity for treatment of 0.1×1.75, or 0.175. Ten iterations were performed for each spatial window size and compared against radii ranging from 10 to 20 km as the spatial influence of the wakes.

The goal of this preliminary simulation work is to quantify the amount of overlap that occurs and how the intensity patterns differ from those in Chapter 3. Ultimately, for a robust analysis of the interference and interference-free effects, one would need

to simulate wakes that do not overlap in the same space and time as wakes that do overlap. These wakes that do overlap would need to be placed in a way that the volume of overlap is uniform. Thus, setting the true interference and interference-free effects would be possible and the comparison of the estimates of these effects to the truth would be quantifiable. We leave this piece for future work in this area.

Figures 4.6, 4.7, and 4.8 represent one scenario of this simulation, with the true radius set at 15.56 km. The pieces of the interference effect are illustrated in Figure 4.6, where a noisier version of the patterns in Chapter 3 can be seen. There is a plateau in the average estimated values - both for treatment intervention events that experience interference and those that do not - up to the true radius, followed by a decline in these values for larger spatial windows than the truth. The averages of $\widehat{\lambda}_1(1,1)$ and $\widehat{\lambda}_0(1,1)$ seem to be moving closer to one another the larger the spatial radius becomes, resulting in less clearly-defined curves pre- and post-intervention event when there is interference. The middle plot of this figure indicates that there is potentially some separation between estimated intensities before and after treatment intervention events that do not experience interference, albeit on a much smaller scale than for treatment intervention events that do experience interference. The pattern within the average estimated interference effects is harder to determine given the variability and smaller number of iterations.

The middle plot from Figure 4.6 is also seen as the left-hand plot in Figure 4.7, used to calculate the interference-free effect. In comparison, the middle plot of Figure 4.7 illustrates the average $\widehat{\lambda}_1(0)$ and $\widehat{\lambda}_0(0)$ for each iteration overlapping for the entirety of the curve, which is what we would expect for control intervention wakes. This follows the pattern from the previous chapter. The average estimated interference-free effects also show more noise; however, there seems to be a more-defined curved lower bound to the estimates.

A larger disparity pre- to post-intervention event is seen within the total 150 treat-

ment intervention events if one were to ignore the overlap and treat the problem as in Chapter 3 (Figure 4.8). The converging of the pre- and post-treatment intervention event curves near the tail for larger spatial windows is again seen here, most likely influenced by the pattern in the left-hand plot of Figure 4.6. The middle figure for control intervention wakes is the same as that from Figure 4.7. It is interesting to note that the interference and interference-free effects tend to have similar makeups in terms of the spread and magnitudes (righthand plot).

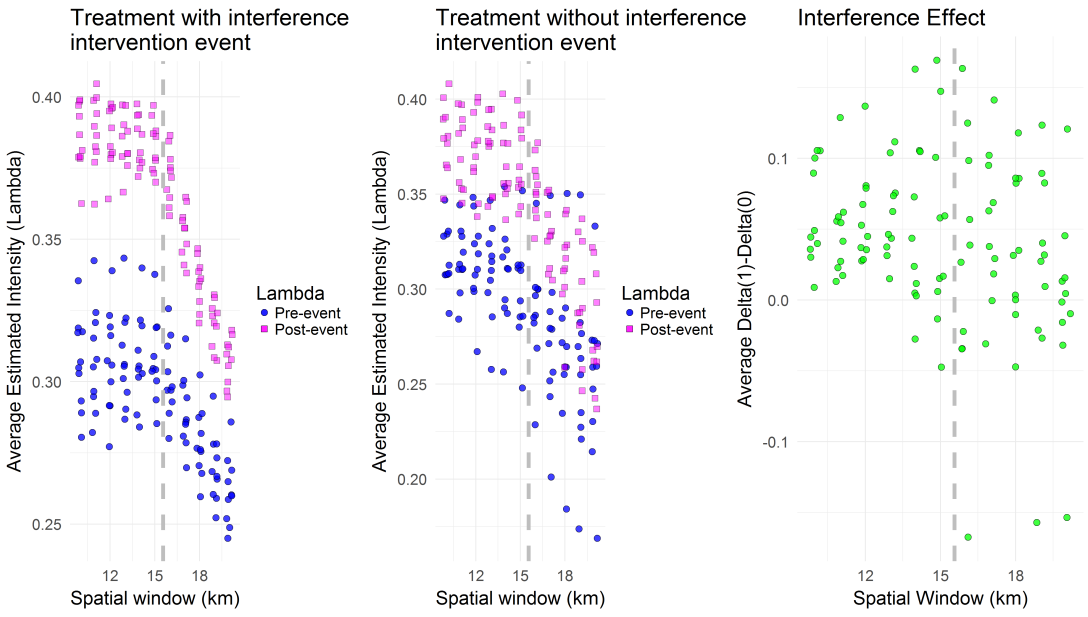


Figure 4.6: Illustrations of the average $\widehat{\lambda}_1(1,1)$ and $\widehat{\lambda}_0(1,1)$ (left panel), average $\widehat{\lambda}_1(1,0)$ and $\widehat{\lambda}_0(1,0)$ (middle panel) with 150 total treatment intervention events, some of which receive interference the others do not. The average difference in $\widehat{\Delta}(1,1)$ and $\widehat{\Delta}(1,0)$ is plotted in the right hand panel. The dotted grey line lines represent the true spatial window size.

These results indicate that the interference effect may be more pronounced than the interference-free effect by the more distinct separation of pre- to post- intervention intensity estimates. While this could be an artifact of the setup of our simulation, further study into this phenomenon is warranted. However, these preliminary results show that our method is capable of separating out and estimating the interference and interference-free effects when a large amount of overlap between units is present

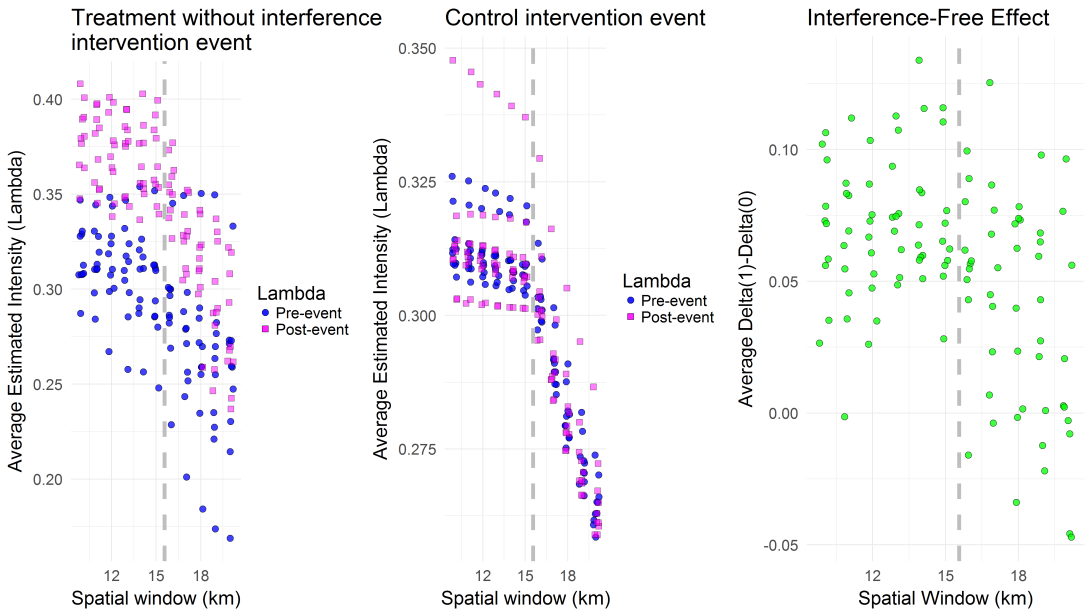


Figure 4.7: Illustrations of the average $\widehat{\lambda}_1(1,0)$ and $\widehat{\lambda}_0(1,0)$ (left panel) from the non-interference treatment intervention events out of 150 total treatment intervention events, and average $\widehat{\lambda}_1(0)$ and $\widehat{\lambda}_0(0)$ (middle panel) with 300 control intervention events. The average difference in $\widehat{\Delta}(1,0)$ and $\widehat{\Delta}(0)$ is plotted in the right hand panel. The dotted grey line lines represent the true spatial window size.

in the data.

We were also interested in quantifying the amount of overlap that occurred between wakes used in estimating the interference and interference-free effects to determine the impact it may have. Summary measures of this, along with the estimates themselves, for the preliminary simulation studies are given in Tables 4.2 and 4.3. In comparing the average estimated interference (Table 4.2) and interference-free (Table 4.3) effects, the estimated interference-free effects tend to be significant more often for smaller spatial windows in comparison to the associated estimated interference effects. This leads one down the path of questioning whether the causal effects without taking into account interference (effects under SPP framework from Chapter 3) would still be found to be significant as well for the iterations that had non-significant interference effects. This would be a natural next step in providing further evidence of the benefits of this framework. Additionally, we see that there is a much larger

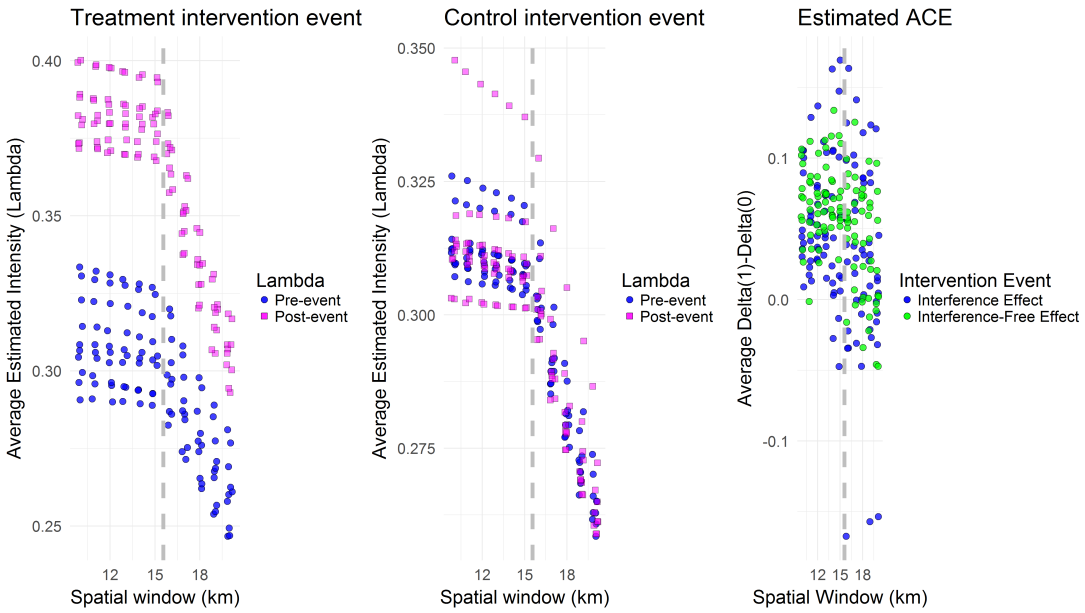


Figure 4.8: Illustrations of the average $\widehat{\lambda}_1(1)$ (all $\widehat{\lambda}_1(1,1)$ and $\widehat{\lambda}_1(1,0)$) and $\widehat{\lambda}_0(1)$ (all $\widehat{\lambda}_0(1,1)$ and $\widehat{\lambda}_0(1,0)$) (left panel) with 150 treatment intervention events, and average $\widehat{\lambda}_1(0)$ and $\widehat{\lambda}_0(0)$ (middle panel) with 300 control intervention events. Both the average difference in $\widehat{\Delta}(1,1)$ and $\widehat{\Delta}(1,0)$ and the average difference in $\widehat{\Delta}(1,0)$ and $\widehat{\Delta}(0)$ are plotted in the right hand panel, for comparison. The dotted grey line lines represent the true spatial window size.

percentage of wakes that experience interference (overlap) in Table 4.2 prior to matching, and still a fair amount of overlap after matching. On the other hand, there is absolutely no overlap between wakes used in estimating the interference-free effect (by definition), and a small amount of overlapping wakes that are introduced after matching. We would expect a larger percentage of wakes with overlap to be used in the interference effect 'bucket', by definition. The percentages are all under 17%, in which case one could potentially further adjust for previous control and/or treatment events to further mitigate the effects of violating SUTVA. By identifying two causal effects that consider slightly different groups of wakes, the issues of large percentages of overlapping wakes are addressed appropriately. More extensive simulations and measures to adequately and fully assess the full benefits and advantages of this framework are needed in the future.

		Results					% Overlap Before Matching				% Overlap After Matching			
Time	Space	Average Estimate	S.D.	Minimum Estimate	Maximum Estimate	N(p < 0.05)	Average	Minimum	Maximum	S.D.	Average	Minimum	Maximum	S.D.
	10	0.061	0.036	0.009	0.106	4	62.73	53.33	77.33	7.03	13.80	4.67	20.67	7.03
	11	0.048	0.034	0.013	0.129	1	69.27	61.33	80.67	6.31	12.33	6.00	19.33	6.31
	12	0.059	0.034	0.028	0.137	3	74.20	66.00	84.00	5.86	12.27	4.67	19.33	5.86
	13	0.060	0.031	0.015	0.112	3	78.53	68.67	86.00	5.21	11.27	6.00	18.67	5.21
	14	0.051	0.059	-0.027	0.163	3	82.13	73.33	86.67	4.62	9.40	4.00	15.33	4.62
10	15	0.051	0.070	-0.047	0.169	4	84.67	77.33	89.33	3.69	7.93	4.00	13.33	3.69
	16	0.025	0.095	-0.167	0.164	3	86.93	82.00	90.67	3.10	6.53	2.67	12.67	3.10
	17	0.052	0.052	-0.031	0.141	2	89.27	84.00	92.00	2.42	5.93	3.33	8.67	2.42
	18	0.039	0.053	-0.047	0.118	2	90.53	86.00	92.67	2.17	5.27	3.33	8.00	2.17
	19	0.020	0.078	-0.157	0.124	1	92.00	88.00	95.33	2.20	4.07	2.00	6.00	2.20
	20	-0.001	0.068	-0.154	0.121	1	93.40	91.33	95.33	1.39	3.67	2.00	6.00	1.39

Table 4.2: Summary measures are given for preliminary simulations of interference effects. Averages as well as the minimum, maximum, and standard deviation across iterations are given for the effect estimates. N($p < 0.05$) is a count of the number of iterations at that temporal and spatial window that had significant effects. The average, minimum, maximum, and standard deviation of the percentage of overlapping wakes is also calculated, both pre- and post-matching. 'Overlap' is defined as wakes for which the center intervention events are two different types (one control and one treatment) and are located less than $2r$ apart.

		Results					% Overlap Before Matching				% Overlap After Matching			
Time	Space	Average Estimate	S.D.	Minimum Estimate	Maximum Estimate	N(p < 0.05)	Average	Minimum	Maximum	S.D.	Average	Minimum	Maximum	S.D.
	10	0.070	0.027	0.026	0.106	8	0	0	0	0	11.20	4.00	16.67	3.99
	11	0.063	0.032	-0.001	0.112	7	0	0	0	0	10.33	4.67	14.67	3.65
	12	0.063	0.023	0.026	0.103	7	0	0	0	0	10.73	4.00	19.33	4.39
	13	0.077	0.022	0.049	0.113	9	0	0	0	0	9.73	5.33	16.67	4.05
	14	0.078	0.027	0.051	0.134	7	0	0	0	0	8.13	3.33	15.33	4.56
10	15	0.070	0.026	0.028	0.116	6	0	0	0	0	6.80	2.67	13.33	3.87
	16	0.058	0.032	-0.016	0.099	4	0	0	0	0	5.33	2.00	10.67	2.90
	17	0.052	0.039	-0.004	0.125	4	0	0	0	0	5.27	3.33	8.67	2.14
	18	0.039	0.040	-0.034	0.078	3	0	0	0	0	4.60	1.33	8.00	1.87
	19	0.035	0.039	-0.022	0.098	1	0	0	0	0	3.40	0.67	6.00	1.68
	20	0.015	0.048	-0.047	0.096	1	0	0	0	0	3.00	0.67	6.00	1.76

Table 4.3: Summary measures are given for preliminary simulations of interference-free effects. Averages as well as the minimum, maximum, and standard deviation across iterations are given for the effect estimates. N($p < 0.05$) is a count of the number of iterations at that temporal and spatial window that had significant effects. The average, minimum, maximum, and standard deviation of the percentage of overlapping wakes is also calculated, both pre- and post-matching. 'Overlap' is defined as wakes for which the center intervention events are two different types (one control and one treatment) and are located less than $2r$ apart.

4.4.2 Conflict Data: Civilian Collaboration in Afghanistan

The data are described in detail in Chapter 3. Restated, the original hypothesis is:

> H_A: Indiscriminate insurgent violence (treatment) using IEDs increases civilian handover of unexploded ordnances to U.S. troops compared to selective insurgent violence (control) using IEDs.

The estimated average causal effect associated with this hypothesis is the change in intensity of dependent events, represented as instances of civilian turn-in/assistance, over time following an intervention event that experienced indiscriminate insurgent violence (treatment) versus the change in intensity of civilian turn-ins if the intervention event experienced selective insurgent violence (control). Here, we consider the estimation of an interference effect and interference-free effect, assessing whether this framework is more appropriate given the potential overlap of wakes. The hypothesis above is then transformed into the following two hypotheses, for the interference effect:

> $H_{A,S}$: Indiscriminate insurgent violence (treatment) using IEDs has a differing effect on civilian handover of unexploded ordnances to U.S. troops when in close proximity to selective insurgent violence (control) using IEDs than when further away.

and for the interference-free effect:

> $H_{A,\mathcal{T}}$: Indiscriminate insurgent violence (treatment) using IEDs increases civilian handover of unexploded ordnances to U.S. troops when geographically far from any selective insurgent violence (control) using IEDs, and thus do not receive any spillover influence.

Given the results of the analysis in Chapter 3, we do not test for a separate interference and interference-free effect in this particular problem. This is due to the

relatively weak signal and low intensities that were found. Particularly with strongest signals in the near vicinity (1-3 km) of violent events, consideration would need to be given to whether this interference setup brings further insights to this particular problem. We set up the hypotheses for this problem merely for illustration of the proposed approach and as an example for future applications.

4.5 Discussion

The work in this chapter continues to address the need for a spatially explicit causal framework for determining how spatial effects impact treatment variation and treatment effects. A major challenge in spatial and spatio-temporal casual relationships is the underlying complicated processes that exist in these settings. In this chapter, we focus on spatial interaction between units that can violate the Stable Unit Treatment Value Assumption (SUTVA), a key assumption underlying the counterfactual framework.

Continuing under the spatial point process framework in Chapter 3, connections are made to geographically-defined effects that take interactions and spillover of units into consideration. By allowing for an interference effect and interference-free effect, rather than one overall causal effect, we are able to re-define the causal relationships without violating SUTVA. That is, we are able change the question and goals to better align with the problem setting at hand. Keele and Titiunik (2017) also provide that the interference effect may be tested, allowing for the appropriate use of the original, overall causal effect. This flexibility provides a more robust process in identifying when the violation of SUTVA meaningfully changes results to a degree that it should be accounted for. This builds upon the suggestions of Schutte and Donnay (2014) in appropriately accounting for the overlap of spatio-temporal wakes when estimating the causal effect.

We begin to setup the simulation models necessary to understand the impact of the interference and interference-free effects in the face of overlapping wakes, representing a violation of SUTVA. These results show an indication in a positive direction of the benefits and advantages this framework can provide when interference may be present within a problem. Continued modeling to build up to a setting that reflects the full complexities of this scenario is needed. To appropriately and accurately generate the moving pieces involved with the spatial effects, interactions, and treatment effects is complex. This work would greatly benefit from a larger-scale simulation study to extend and further validate the framework laid out here. This work further confirms that research in the area of causal inference can greatly benefit from this spatial perspective, especially in multidisciplinary applications.

Chapter 5

Conclusions

5.1 Spatial-Causal Setting

While interest in causal inference continues to grow across multiple disciplines, there are areas that still need major contributions and methodological advances, particularly when it comes to spatial data and geographical elements. The fundamental ideas of causal inference in the counterfactual framework rely on key assumptions that can be difficult to justify when applied to spatial data and geographic questions. When treatments are defined by geographic features (e.g. via local policy decisions), understanding the effects of these treatments through observational studies can be challenging. The impact of spatial effects on aspects of the counterfactual framework must be considered, falling into three broad categories: 1) how the treatment is assigned or chosen, 2) potential sources of variation in treatment variables, and 3) what effects are being estimated (Kolak, 2017).

5.1.1 Spatially Defined Treatment

How the treatment or intervention being investigated is defined is an important determining factor in what methods may be appropriate in analysis. We limit the scope of this dissertation to the discussion of two treatment settings - treatments that are defined by regions, and treatments that are defined by distance from an event point in space and time. When considering geography in a causal context, *local* treatments are of most interest - individuals *here* are considered treated, compared to individuals *there* that are not. This spatial definition of treatment manifests itself in three ways - assignment by some spatial pattern, proximity to intervention that affects likelihood of being treated or not, and places with certain characteristics more likely to be treated (Kolak, 2017). In any scenario, leveraging the geographical nature of treatment assignment to more appropriately compare units is the common goal of interest. Considering a spatially-defined treatment also leads to needing to correctly account

for potential issues of selection bias (geographic and otherwise), spatial dependence in the form of interference or spillover, and/or spatial heterogeneity.

5.1.2 Methods for Proposed Specific Questions

There have been a number of studies and methodological extensions from the traditional causal literature focusing on spatial extensions. These cover methods from propensity scores and matching methods, to difference-in-difference frameworks, regression discontinuity, and instrumental variables. Kolak (2017) provides a great summary of these in Figure 1 (p.34). Specific extensions include spatial propensity score matching (Chagas et al., 2012), distance-propensity score matching algorithm (DAPSm) (Papadogeorgou et al., 2016), geographic regression discontinuity (GRD) (Keele and Titiunik, 2014; Keele et al., 2015; Keele and Titiunik, 2016), and matched wake analysis (MWA) (Schutte and Donnay, 2014), to name a few.

The spatial propensity score of Chagas et al. (2012) relaxes spatial effects and considers the spatial dimension to be latent. This spatial propensity score is defined as a spatial autocorrelation, spatial autoregression, or spatial error model, since a logit/probit model cannot be used. It thus can confirm conditional likelihood or spatial dependence. DAPSm (Papadogeorgou et al., 2016), while also following a propensity score model for matching, differs in the model specification. The goal of DAPSm is to unite spatially-indexed data with propensity score matching in order to account for unmeasured spatial confounding by adjusting for observed confounders in addition to unmeasured spatial confounders. A propensity score is produced as the weighted average of the propensity score difference and a measure of the distance between treated-control pairs as a measure of similarity.

The general idea of regression discontinuity is to determine whether a behavior deviates from the norm at the point of discontinuity, assuming that the population is similar on either side of the discontinuity or any differences can be controlled

for. This deviation would suggest a direct impact resulting from the phenomenon that is being studied. Imbens and Lemieux (2008) and Lee and Lemieux (2010) both provide thorough reviews of regression discontinuity. This idea can be extended spatially in various directions. Keele and Titiunik (2014); Keele et al. (2015) formalize spatial effects through their geographic regression discontinuity design by specifying geographic boundaries as regression discontinuities. This is explored in more depth later (Section 5.3). Boundary-continuity design using a set of spatial weights to account for observables and unobservables may be appropriate as a spatial regression discontinuity design (Kolak, 2017).

Also explored in more detail in section 5.3 is MWA. This method uses statistical matching with sliding spatio-temporal windows in an effort to address selection bias and the modifiable areal unit problem. It incorporates a difference-in-differences design to establish treatment effects for the matched sample (Schutte and Donnay, 2014). An interpretable causal approach that is computationally efficient for conflict event studies is one of the goals of the MWA.

5.2 Example Applications

5.2.1 Voting

In the works of Keele and others, voter turnout is a topic of interest that is explored from many different facets involving geographical components. One specific instance is studying whether campaign ads increase voter turnout by using the exogenous variation in the volume of TV ads created by media market boundaries (Keele et al., 2015). Here, the treated and control units are defined by the media market boundaries, which tend to be identical to county boundaries. The specific element of geography introduces some subtle, yet highly important, differences in this problem and its handling within an adapted standard regression discontinuity design. First, it is possible for

compound treatments to occur, where multiple treatments that affect the outcome of interest occur simultaneously and individuals may receive different combinations of exposures to treatments, depending on their location. Second, different measures of distance from the defined boundary may require different identification assumptions. Finally, spatial correlation must also be accounted for when making inferences on such data.

Another voting problem discussed by Keele and Titiunik (2017) looks at the effect of all-mail voting, the only convenience voting method that eliminates precinct-place voting, on voter turnout. Here, the question is examined in the context of a geographic natural experiment in Colorado, where a key element of the design in Keele and Titiunik (2017) is fulfilled by focusing on a small geographic area around the boundary separating two counties. While this ensures that individuals close to the county border on either side are similar in terms of their pretreatment covariates, it also introduces the issue of *interference* by the fact that treated and control voters live in close proximity. This problem of interference is more likely to be an issue in designs that exploit the geographic proximity, as compared to those that do not. By treating the geographic locations of the voters in the sample as random and the boundary between treated and control areas as fixed, Keele and Titiunik (2017) are able to investigate interference for a geographical natural experiment that focuses on a narrow band around a boundary by considering two estimands of interest.

The basic concepts of the work in Chapters 3 and 4 of this dissertation still apply for this problem since the *latent* spatial point process underneath would give distributional results for each aggregation of the data. As scan statistics (e.g., SaTScan (Kulldorff, 1997)) do, the method would jump to the next area of aggregation if considering non-contiguous aggregations, such as administrative boundaries, that are not a continuous space.

5.2.2 Alcohol Sales

The laws regarding sales of alcohol and related establishments may differ from county to county. Waller and Chang are currently studying the effects of bar closing times and driving under the influence (DUI) arrests. If one county changes their law to close bars and stop alcohol sales at 2:00am, compared to the previous cutoff time of 3:00am that was shared with their neighboring counties, one could measure the effect of this law-change and the amount of activity that is seen spilling over to the neighboring counties. (This would be a very interesting area of future research!)

5.2.3 Air Pollution

The area of air pollution also brings up many hypotheses of a spatial-causal nature. In particular, many of the applications that fall into this category also run into the issue of interference. The effect of air quality regulation at a given location on air pollution at other locations (Zigler et al., 2012) is one such example, i.e., increased regulation *here* impacts ambient pollution downwind. Another example is in considering how ambient ozone concentrations are effected by power plants with and without selective catalytic or selective non-catalytic (SCR/SNCR) NO_x technologies (Papadogeorgou et al., 2016).

To set up a similar framework to what we propose in this dissertation with this example of air pollution, we start by considering power plants with exposures 1 and 0 (present and not present). Each power plant would then have a corresponding impact zone in space and over time. The idea would be to quantify the effect of the power plant with exposure=1 compared to power plant with exposure=0 over time. It would still be necessary to determine the correct radius of influence, as in Chapter 3. The underlying spatial point process would be better represented by a *heterogeneous Poisson process* in this context, which would introduce further potential SUTVA violation issues; thus, expanding on the work in Chapter 4 would need to

be explored. With this framework, one would expect to see some sort of continuous drop off for the exposure effect, given the difficulty in choosing a 'hard line' space of influence (air pollution travels, though is not as strong further from the point of emission). Incorporating the concept of multiple points in time (some days power plants emit a lot of pollution, other days not very much), either within the definition of the effect of interest or as a covariate to adjust for, would also need to be considered.

5.2.4 Conflict Analysis

In research exploring causes and effects of human conflict, the drivers of local levels of violence can be thought of and described in three main parts. There are *a priori* exposures of any location to violence that can be driven by exogenous geographic conditions. Additionally, levels of violence tend to vary over time, which can be described as the momentum of a conflict at a given time. It is imperative to be able to tease apart the effects of the true exposure and the momentum in order to understand and analyze the causal effect of specific interventions also contributing to levels of violence.

Schutte and Donnay (2014) investigate causal relationships exploring whether civilians are more likely to deny insurgents access to explosives in response to local occurrences of indiscriminate violence. To our knowledge, from an epidemiological or statistical perspective, little has been done in the current literature in regards to this setting, considering both spatial and temporal elements in a potential outcomes framework for causal inference.

5.3 Review of Current Methods for Spatial-Causal Inference

In this section, we provide a brief but thorough review of the methods mentioned in the preceding chapters and outline areas of future work in each.

5.3.1 Matched Wake Analysis

As detailed in the previous chapters, MWA (Schutte and Donnay, 2014) is a novel method for quantifying causal relationships in spatio-temporal event data, specifically when one is interested in how different interventions affect subsequent levels of reactive events in nearby areas. By combining spatial and causal methods, the authors address a number of challenges that arise with spatio-temporal event data with a particular emphasis on conflict research. The process and details of the inner workings of the method are discussed in detail in Chapter 3.

Advantages of MWA include its ability to find patterns in event data, the use of coarsened exact matching (CEM) (Iacus et al., 2012) as an efficient, automated approach to balanced matching, and the incorporation of sliding spatio-temporal windows. The geographic relevance of events as treatments or interventions and quantifying these causally in the social sciences has become increasingly popular, and this method brings additional insights and challenges to the problem. The traditional definition of a unit as an individual that receives or experiences some level of exposure does not quite fit this setup. Instead, units are themselves defined by the intervention-type of an event and are defined as an associated area (volume) of interest. Schutte and Donnay (2014) define the associated volume in space and time as a cylindrical wake centered around the intervention event, which is either treatment or control. Both units and treatment receipt are geographically defined. The spatio-temporal nature of units has potential consequences related to the modifiable areal unit prob-

lem (MAUP) (Openshaw and Taylor, 1979) and spatial interference and/or spillover, violating the stable unit treatment value assumption (SUTVA) (Rubin, 1980).

MAUP refers to the fact that the selection of artificial cell sizes drives spatial inference, and expresses the idea that results of an analysis on a neighborhood or region (or any arbitrarily defined space) does not yield the same results on a different measure of space, say ZIP code, or on the individuals in that area. Conclusions based on data aggregated in one way may change if the data are aggregated in another way. The remedy in MWA to address this issue is to use a sliding spatial and temporal window design similar to SaTScan (Kulldorff, 1997). SaTScan was originally applied to epidemiological problems for testing whether there may be an elevated risk of disease in a certain region but that the extent of the risk elevation is unknown. In general, the method reveals whether a cluster of events is consistent with chance allocation under the null hypothesis. The impact of violating SUTVA with overlapping units in space and time is assessed, and Schutte and Donnay (2014) determined that unbiased estimation is still feasible if there is a mild to moderate ($\sim$ 20% or less) amount of overlap. They also provide three potential solutions to the violation of this key assumption:

1. One could check the data for overlaps of treatment and control events since these spatio-temporal overlaps are easily identified in empirical data. Selecting a subset not affected by this problem would avoid drawing false inferences, but limit the scope of the time period.

2. To retain interventions in the post-matching sample with similar histories, one could match on the number of previous treatment and control events. This would avoid the result where the causal effect attributed to the intervention is actually due to the product of a mixture of interventions. This also accounts for the scenario where different intervention types occur prior to the actual intervention of interest, affecting subsequent levels of dependent events. This

leads to a matched sample with fewer overlapping events, decreasing the overall balance between treatment and control groups with regard to exposure.

3. Remove overlapping observations from the sample, which biases estimates from such non-random deletion.

5.3.2 Geographic Regression Discontinuity

The GRD design (Keele and Titiunik, 2014, 2016, 2017; Keele et al., 2015) gives rise to a matching framework that incorporates geographic proximity and observed covariates in a flexible manner. This approach is useful when units are sorted along some border, receiving or not receiving treatment based on their location in space. While one can condition on a set of observed pretreatment covariates, there is still a potential for unmeasured confounders (Keele et al., 2015; Keele and Titiunik, 2017). GRD proposes to combine this conditioning on observed covariates with the exploitation of geographic proximity, detailed in Keele and Titiunik (2014); Keele et al. (2015). The idea behind this work is to use the regression discontinuity (RD) framework, where counterfactuals are the treated and control groups near the boundary, making local treatment effects are identifiable.

This GRD design deals with the issue of strong self-selection around some border of interest. The proposed solution is to assume that after conditioning on pretreatment covariates $\boldsymbol{X}$, treatment assignment is as-if randomized for those who live near the boundary. Geographic distance between treated and control observations to the boundary is minimized while enforcing balance in pretreatment covariates. The conditional geographic treatment ignorability in local neighborhood assumption (Keele and Titiunik, 2017) states that a small neighborhood exists where potential outcomes and treatment assignment are conditionally independent given pretreatment covariates.

The work of Keele and Titiunik (2017) generalizes and expands that of Lee (2008) and Cattaneo et al. (2015) to consider a small geographic neighborhood around the

boundary separating treatment and control areas, such that the condition of as-if randomization, or independence, holds *after* conditioning on pretreatment covariates. It is noted that the emphasis on *after* conditioning on pretreatment covariates is important, since this as-if randomization (independence) may not hold unconditionally for the small geographic neighborhood around the boundary separating treatment and control areas. This assumption allows for the identification of a treatment effect at the boundary points and points included in the geographic neighborhood where it holds. The form of this assumption makes sense to use when treatment assignment is based on geography. Additionally, it is weaker than the assumption when only conditioning on observables, as it requires conditional independence for a subset of the population. Furthermore, the impact of interference is considered by defining an estimand that specifically quantifies the interference effect.

This extension takes into consideration the nature of spatial data and provides a solution to SUTVA violations. Keele and Titiunik (2017) modify identification assumptions appropriately to account for this, and focus on spatially proximate units that may introduce spillovers or interference. Given that the neighborhood around the boundary separating treated and non-treated units is small enough and enough pretreatment covariates are available, one can consider a treatment effect on the treated as an interference-free effect, and the average effect of interference on treated units as the interference effect.

5.4 Comparing and Contrasting Approaches to Spatial Causal Inference

Both the MWA and GRD allow for the estimation of causal effects in a spatial setting. Here, we compare the three recent papers that most embody these methods (Table 5.1 provides a summary). There are some notable threads that connect these two

seemingly distinct approaches. First, we note that MWA is most appropriate for treatment and control point-event data and the GRD design has been considered in the context of area-level observations. While Keele et al. (2015) and Schutte and Donnay (2014) focus on an average effect of treatment on an outcome of interest, Keele and Titiunik (2017) focuses on quantifying and testing the hypothesis of an interference effect, rather than just assuming an "interference-free" effect as in the other two papers.

Both space and time are addressed in the causal question, definitions, and spatial challenges considered in MWA, as opposed to space only in the GRD literature. A substantial difference between MWA and GRD also lies in the definition of units and defined treatment and control boundaries. MWA considers the units, defined as treated or control wakes by the space and time local to the corresponding intervention event, with created or user-defined boundaries; GRD, on the other hand, considers a strictly contiguous boundary between the treated and control areas that is fixed.

When considering the underlying assumptions necessary for causal inference under each method, there are a number of differences. Both MWA and GRD Keele and Titiunik (2017) invoke SUTVA in order to estimate causal effects; however, that is the only similarity to be found. Assumptions for MWA stem from CEM, where for each observation, the potential outcome under treatment is always observed and the potential outcome under control is always estimated, and analysis is restricted to common empirical support for matches. Both GRD approaches rely on a geographic-variation of the "no unmeasured confounders" assumption, or strong ignorability, specifically within a local neighborhood. Keele et al. (2015) use a 2-dimensional score to accomplish this, and Keele and Titiunik (2017) focuses on extending the local neighborhood strong ignorability to within interference areas. The inherent unit of study for each of these approaches also differs.

MWA	**GRD (Keele and Titiunik (2017))**	**GRD (Keele et al. (2015))**
	Key Question	
Effect of IED explosions with civilian casualties on rates of information turn-in over time after intervention event	Effect of convenience voting - all-mail election- compared to traditional in-person voting on election day on voter turnout - Does interference alter inferences about the effect of all-mail voting?	Effect of ballot initiatives on voter turnout
	Key Steps	
1. Define spatio-temporal wakes around intervention events 2. Count rates over time of dependent events before, after intervention event 3. Match control and treatment wakes on risk factors/ covariates/confounders, including trend in the dependent variable before interventions 4. DD regression to obtain estimate of causal effect	1. Define "local" area for comparison 2. Define causal estimand under assumption of no interference, then estimate 3. Define causal estimands allowing for interference, estimate 4. Compare - if $\tau_T \approx \tau$, τ_S not significantly different from 0 $\implies$ no interference	1. Define "local" area for comparison 2. Optimal subset matching with integer programming
	Data Format	
Space and Time Disaggregated point/count event data	Space Areal individual-level data	Space Areal individual-level data
	Units	
Treatment/Control *wakes* defined by treatment/ control intervention *events*	Treatment/control *individuals* defined by treatment/ control (fixed, contiguous) *areas*	Treatment/control *individuals* defined by treatment/ control (fixed, contiguous) *areas*
	Assumptions Made	
1. (From CEM) 'no omitted bias', $T_i \perp\!\!\!\perp \{Y_i(0), Y_i(1)\} \mid \boldsymbol{X}_i$ $\implies$ implicitly, SUTVA 2. (From CEM) For each observation, $Y_i(1)$ is always observed, $Y_i(0)$ always estimated 3. (From CEM) Restriction to common emprical support for matches	1. No interference 2. Conditional mean independence in local neighborhood "Revoke" Assumption 1. to continue under assumption of interference 3. As-if random geographic location within interference areas *Note*: A1 + A2, or A2 + A3 made together	1. Conditional geographic treatment ignorability 2. Continuity in 2-dimensional score "Revoke" Assumption 1 3. Conditional geographic treatment ignorability in local neighborhood *Note*: A1+ A2, or A2+A3 made together; A3 is weaker than A1
	Causal Challenges Addressed	
Interference	Geography in ignorability assumption Interference Division into treatment/control areas occur in as-if random fashion (naturally or by assumptions)	Geography in ignorability assumption Division into treatment/control areas occur in as-if random fashion (naturally or by assumptions)
	Spatial Challenges Addressed: Boundaries	
Unit boundaries created/estimated	Boundary naturally defined, fixed	Boundary naturally defined, fixed

Table 5.1: Comparison of the methodology and problem of interest in Schutte and Donnay (2014), Keele et al. (2015), and Keele and Titiunik (2017).

5.4.1 Complications of Spatial Setting for Definitions of Elements of Standard Causal Framework

While MWA provides a novel approach to evaluating and estimating causal relationships of conflict events that could be used for other types of georeferenced event data, there are some limitations to the approach. Namely, the spatio-temporal cylinders defining the units that receive treated or control intervention events can potentially overlap with one another, violating SUTVA. Violating this assumption could lead to biased effect estimates. Two scenarios could result from MWA - (1) multiple units of the same intervention event value overlap in space and time, biasing estimates upward (assuming a positive treatment effect); or (2) a unit of each of a treatment intervention and control intervention could overlap and bias the estimated treatment effect downward.

In the context of the polling problem of interest, Keele and others consider the possibility for interference, resulting in a violation of SUTVA. However, for this application there has been shown to be little evidence of treatment spillovers even within households (Sinclair et al., 2012). Thus, SUTVA violation in the form of voters in control area encouraging neighbors in treated area to vote, or the likelihood of interference across voters even with adjacent households, is slim. Regardless, if such a violation occurred it would bias the effect towards 0, and positive effects would be conservative estimates.

The necessity of SUTVA is twofold: first, it ensures that there exists as many potential outcomes as the number of values the treatment can take; and second, it says that we observe one of the potential outcomes for each unit unaffected by the treatment assignment of other units. For any given question, SUTVA violations will differ depending on the context. If exposure is considered to be an area, then interference deals with units close to the boundary of the treated and non-treated units interacting with one another and how to chose the appropriate boundary between

the treated and non-treated units. If exposure is considered to be events, then the violation comes in the form of multiple units overlapping within a spatial or spatio-temporal cluster. This spillover effect, or indirect exposure effect, characterizes what occurs when the effect of an intervention carries over to the defined control area or control units (or vice versa).

5.5 Future Work

It is of interest to consider more complex scenarios for which this spatial point process view may be beneficial and additional challenges that may be addressed. A number of reasonable extensions of this framework can be considered, from the underlying spatial point process to the assumptions regarding the types of interference and overlap that occur. Below we give details regarding the following areas of extensions to this spatial point process causal framework: (1) underlying spatial point process, (2) shape of the impact zone, (3) alternative adjustment methods for confounding and outcome model designs for generalizability, (4) addressing multiple comparisons, and (5) alternative overlap assumptions and interference types.

First, we use a piecewise homogeneous Poisson point process to characterize the underlying spatial point process. It may be of interest to consider a heterogeneous Poisson point process, or a continuous exponential or Gaussian spatial point process to define the intensity within wakes and/or outside of the wakes. The implications of choosing a heterogeneous spatial point process and properties of another type of spatial point process on causal assumptions such as SUTVA and consistency would need to be thoroughly studied. Along these lines, considering the problem on a plane that takes population density into consideration may be a potential extension, given the kind of problem one is interested in addressing. Transforming space through the use of a cartogram (Inoue and Shimizu, 2006) rather than a standard spatial plane

may allow adjustment for population density without it being a necessary factor in the causal adjustment steps. However, one must give careful consideration to how this changes the question that is being answered.

Second, it is reasonable to consider expansions to the shape of the impact zone. For example, assuming that the span of influence from an intervention event follows a cone-shape rather than a cylindrical shape. This may be a more accurate depiction of the patterns following an intervention event in the particular case of Schutte and Donnay (2014), given the spatio-temporal lag seen in the intensity of civilians reporting events. There is also the situation where treatment and control intervention events differ in the shape of the wakes; say, treatment intervention event wakes follow a cone-shape, while control intervention event wakes have a cylindrical shape. This could help characterize either a ramping up or dispersion of the effect, again considering the temporal behavior of dependent events leading up to or following an intervention event. It would also account for the consistency in intensity over time for control intervention wakes.

Third, broadening the use of this framework could also include incorporating other adjustment methods. Here, coarsened exact matching (CEM) is used to estimate causal effects. Other matching algorithms and the use of propensity scores in matching or inverse weighting may also work well in this setting, with recent methods such as entropy weighting (Hainmueller et al., 2012) and the covariate balancing propensity score (CBPS) (Imai and Ratkovic, 2014). This framework is not limited to a difference-in-differences design, and lends itself to intuitive extensions in regards to specific causal mechanisms for controlling pre-treatment confounders and effect estimation. These may lend themselves to broadening the impact to other public health applications.

Fourth, there is the issue of multiple comparisons when performing multiple tests over the sliding spatio-temporal windows. By being able to determine the correct

window size, we are able to mitigate this issue in some respect. However, the interpretation of secondary or further clusters can still be complicated by multiple comparisons as seen in the data analysis in Chapter 3. How to deal with multiple comparisons and the effect on false discovery rate, considering all distance and time lags is still a concern. Spatial scan statistics, in particular SaTScan (Kulldorff, 1997), are able to circumvent this issue by testing the maximum score of 'unusualness' of a potential cluster. That is, data are simulated under the null hypothesis to obtain a distribution of the maximum statistic, so that the test statistic value from the actual data can then be compared to this null distribution and one can determine the unusualness of the cluster that is seen in the data. While the matched wake analysis approach is similar to SaTScan, it does not result in an overall p-value for all spatio-temporal windows in comparison to a null distribution. Taking the sliding spatio-temporal windows and SatScan-like approach a step further would be able to incorporate the issue of multiple comparisons in a clever way; however, it would not necessarily address the issue at hand in our motivating example.

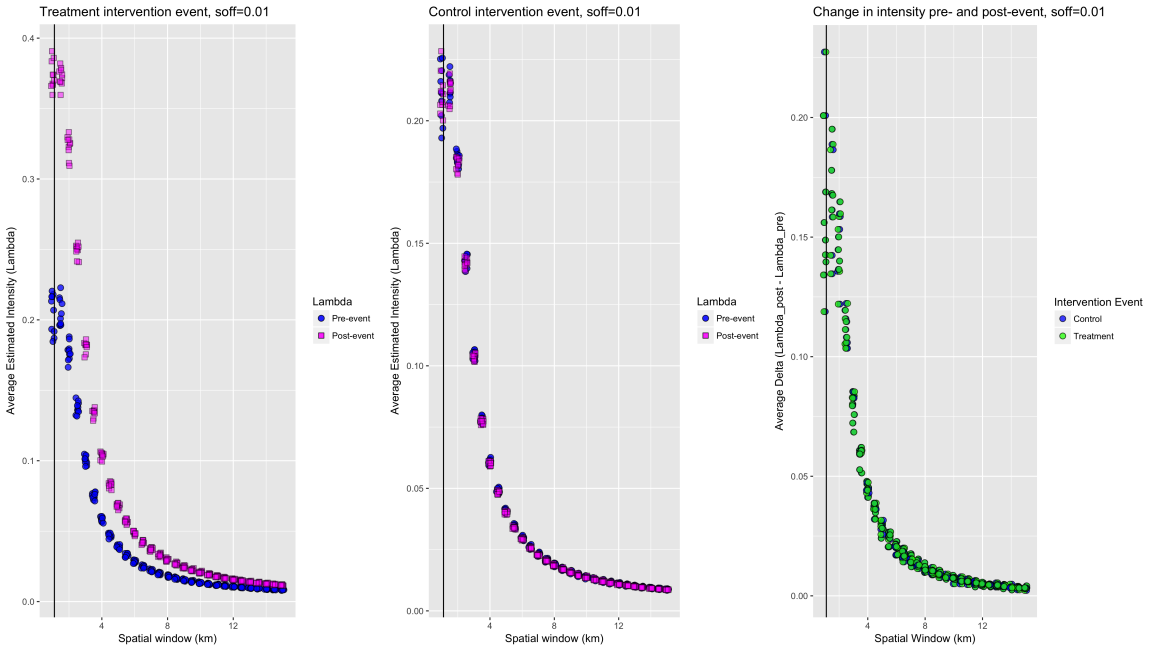
Fifth, the incorporation of different types of interference in defining the alternative hypothesis is an area of potential extension. In Chapter 4 one-sided interference with controls influencing treated is assumed. However, this work could be extended to a four-level overlap structure. Distinguishing between the idea that interference between multiple treated intervention event wakes has a different effect than interference between a treatment and control intervention wake with 2-sided interference has a different effect than no overlap (or overlap between control intervention event wakes) is an important step in considering the types of effects one is estimating with the substantive problem at hand. What makes sense in one context, may not in another. We do not assume that compounded treatment can occur in our scenarios; however, in the context of the conflict problem we have considered, this may provide more insight than controlling for this overlapping event.

Applying MWA with this spatial point process perspective could benefit many other fields beyond conflict analysis when incorporating spatial features into causal problems. Many times it is expensive and impractical to implement an intervention in public health, health-policy, and disease surveillance modeling, and measure the effect over time. Being able to use already-available data and shedding light on results across time and space with a principled framework is appealing. This work would also be able to incorporate SUTVA violations that occur frequently when answering causal questions with geographic data. Overall, the contribution of a solid theoretical framework behind MWA in combination with spatial point process theory presents an important step forward in considering causal frameworks in spatial and spatio-temporal settings.

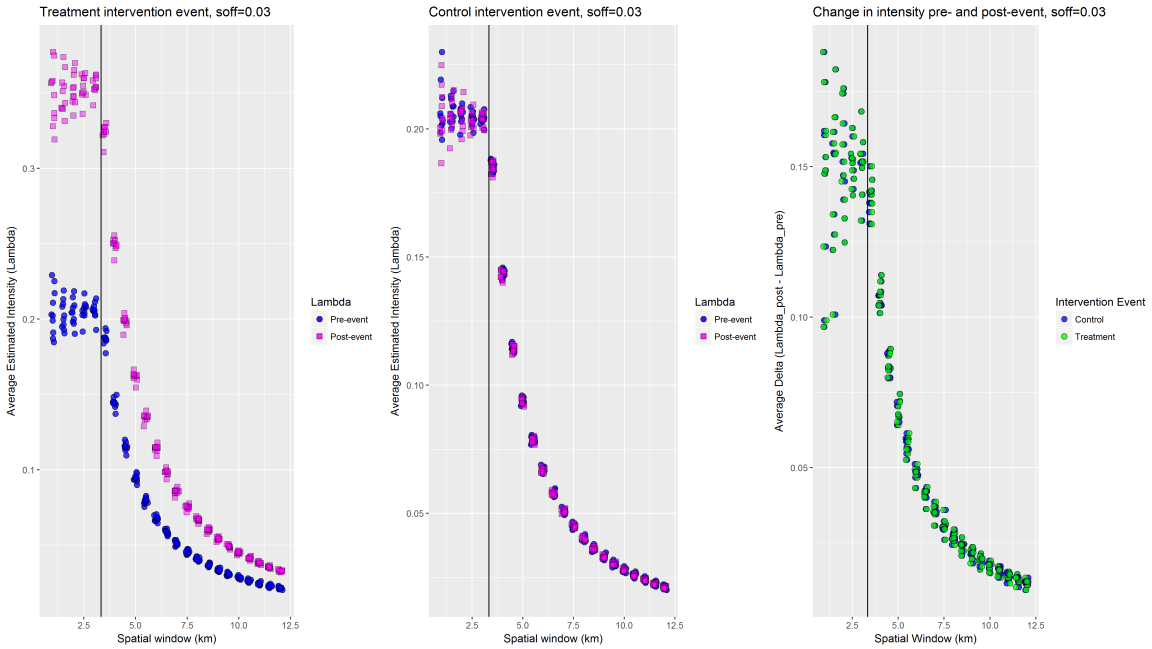
Appendix A

Appendix A: Linking Spatial Point Process Theory and Causal Inference

Below are the additional simulation scenarios for all spatial windows examined that were not included in Chapter 3. We also give other viewpoints of the point event data from the data analysis in both Chapters 3 and 4. Figures A.4, A.5, A.6, and A.7 show patterns of point event locations over time within Afghanistan.

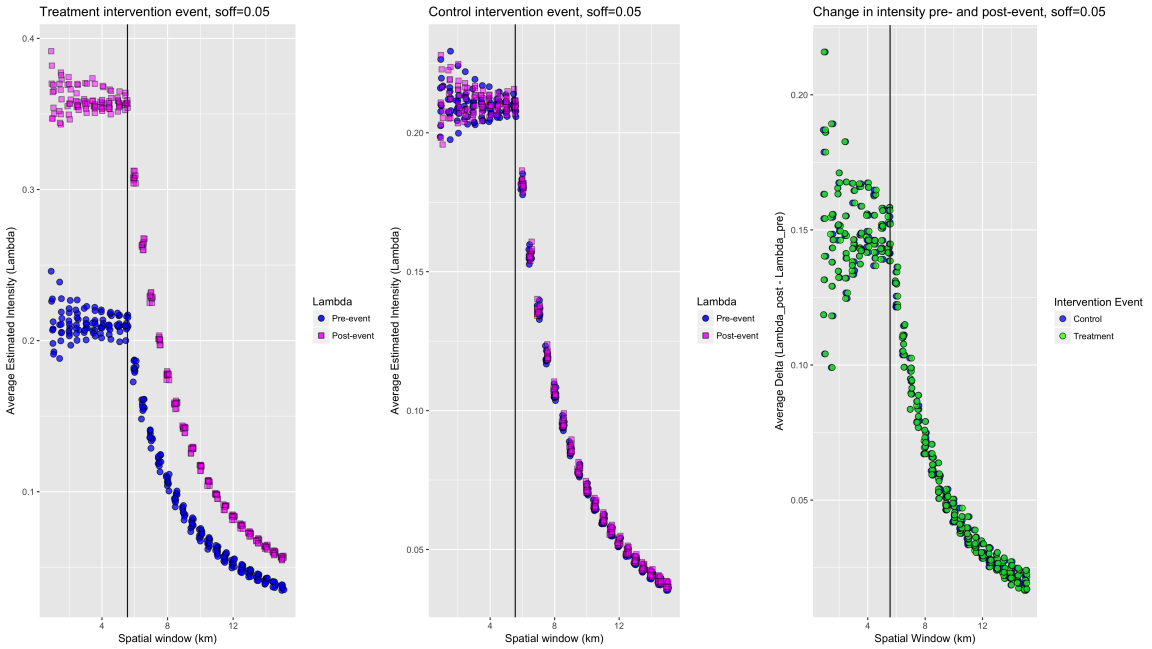


(a)

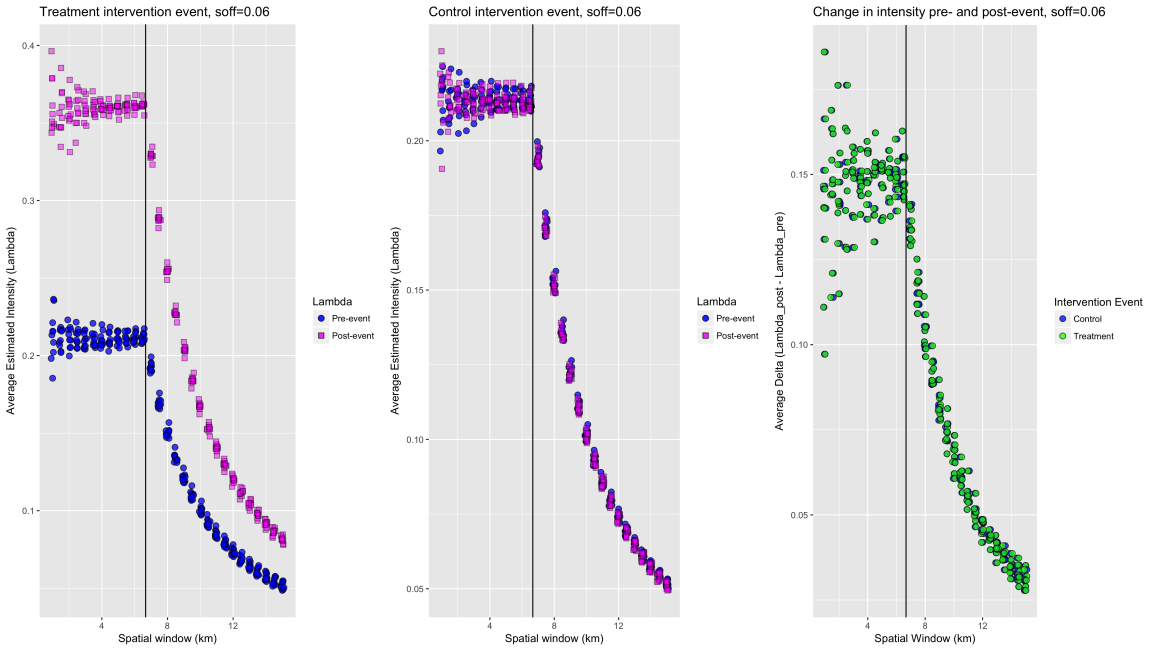


(b)

Figure A.1: Illustrations of the average $\widehat{\lambda}_1(1)$ and $\widehat{\lambda}_0(1)$ with 100 treatment intervention events (left panels), average $\widehat{\lambda}_1(0)$ and $\widehat{\lambda}_0(0)$ with 200 control intervention events (middle panels), and average $\widehat{\Delta}(1)$ and $\widehat{\Delta}(0)$ (right panels). The vertical black lines represent the true spatial window size.

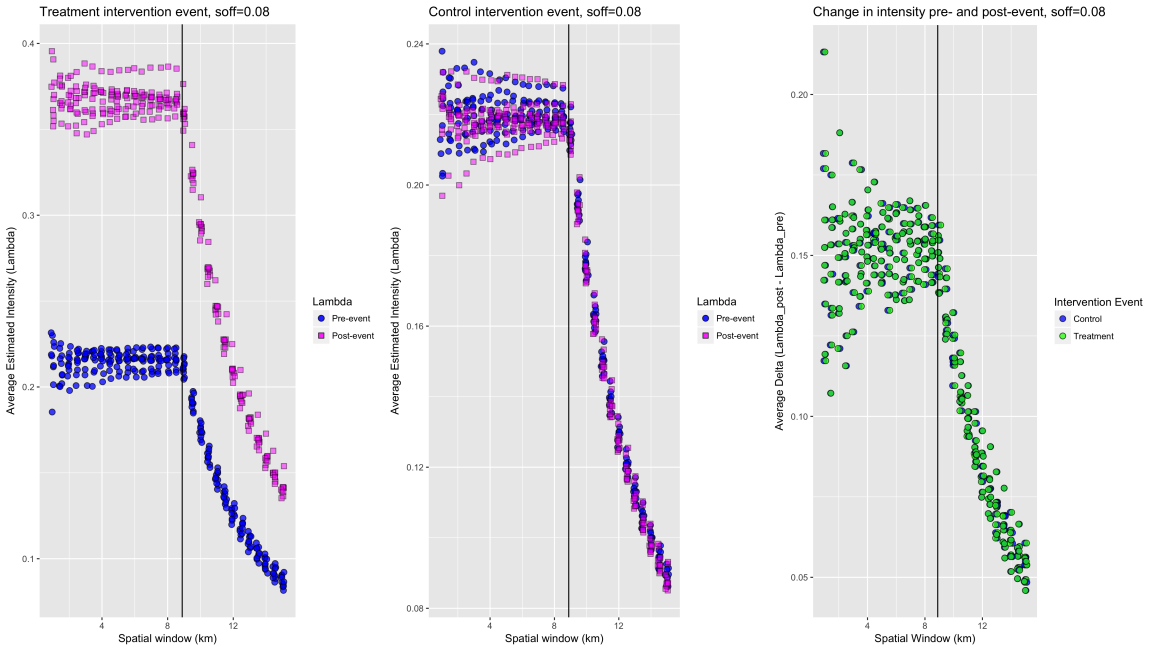


(c)

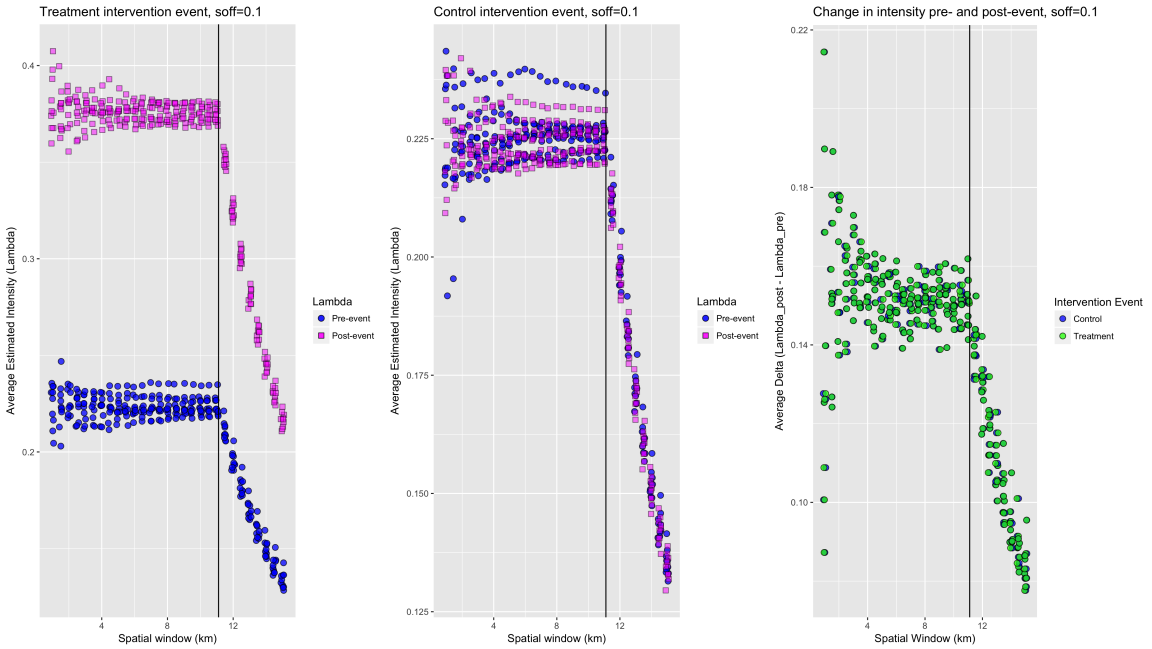


(d)

Figure A.2: Illustrations of the average $\widehat{\lambda}_1(1)$ and $\widehat{\lambda}_0(1)$ with 100 treatment intervention events (left panels), average $\widehat{\lambda}_1(0)$ and $\widehat{\lambda}_0(0)$ with 200 control intervention events (middle panels), and average $\widehat{\Delta}(1)$ and $\widehat{\Delta}(0)$ (right panels). The vertical black lines represent the true spatial window size.



(e)



(f)

Figure A.3: Illustrations of the average $\widehat{\lambda}_1(1)$ and $\widehat{\lambda}_0(1)$ with 100 treatment intervention events (left panels), average $\widehat{\lambda}_1(0)$ and $\widehat{\lambda}_0(0)$ with 200 control intervention events (middle panels), and average $\widehat{\Delta}(1)$ and $\widehat{\Delta}(0)$ (right panels). The vertical black lines represent the true spatial window size.

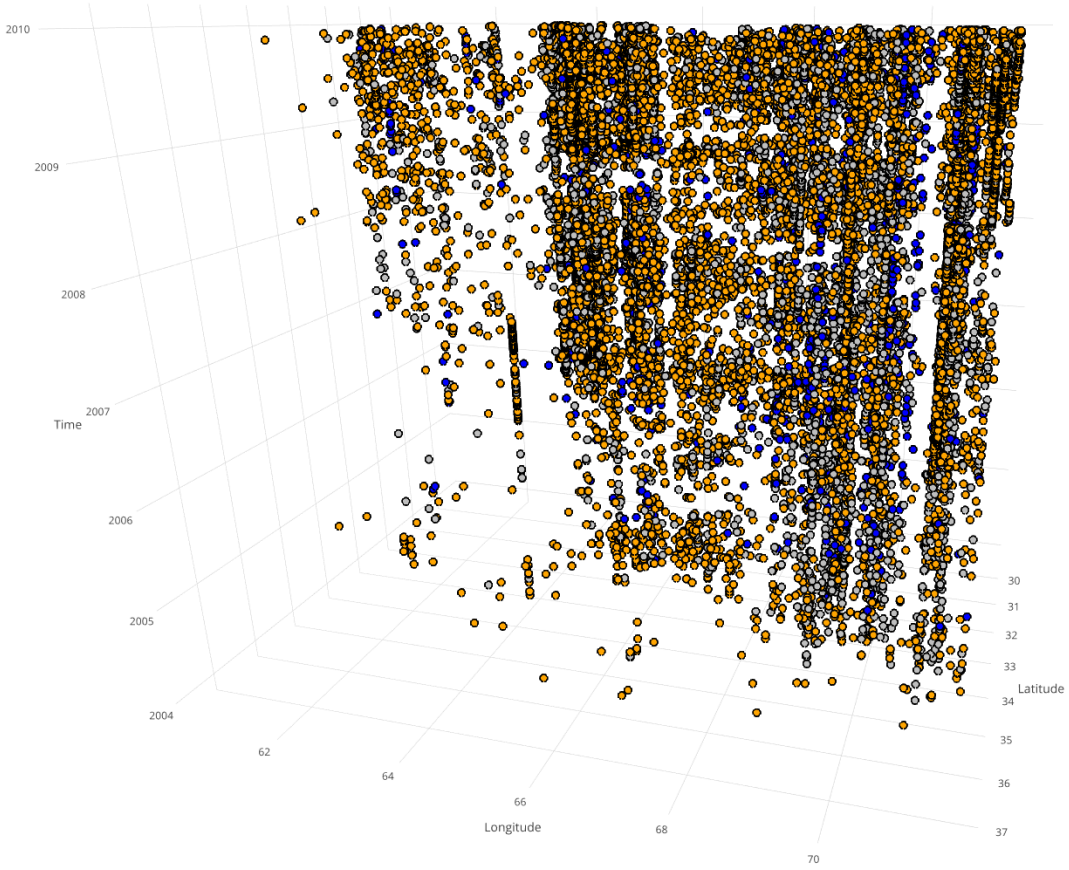


Figure A.4: A 3-dimensional rendering of all data points - treatment and control intervention events and dependent events - from the data set. Points in blue are the dependent events (instances of civilian assistance), points in orange are control intervention events (selective insurgent violent events), and points in gray are treatment intervention events (indiscriminate insurgent violent events).

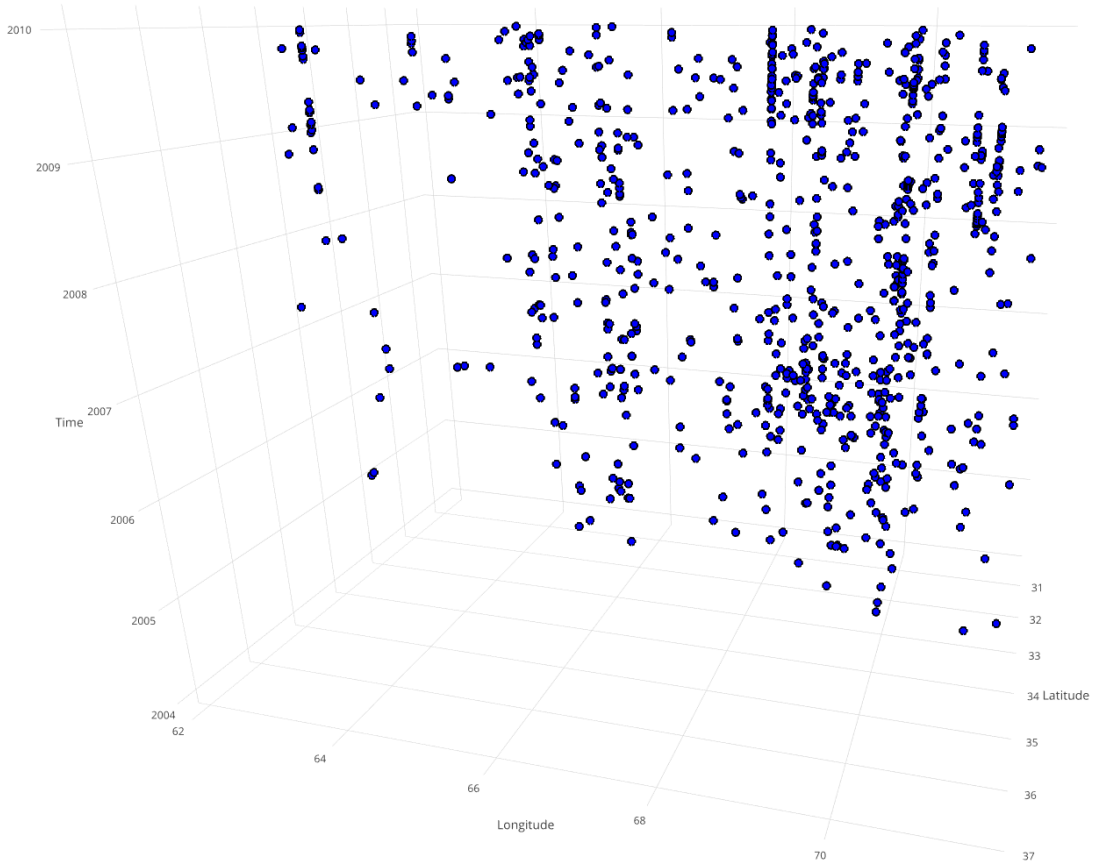


Figure A.5: Illustration of where all dependent events occur in space and time.

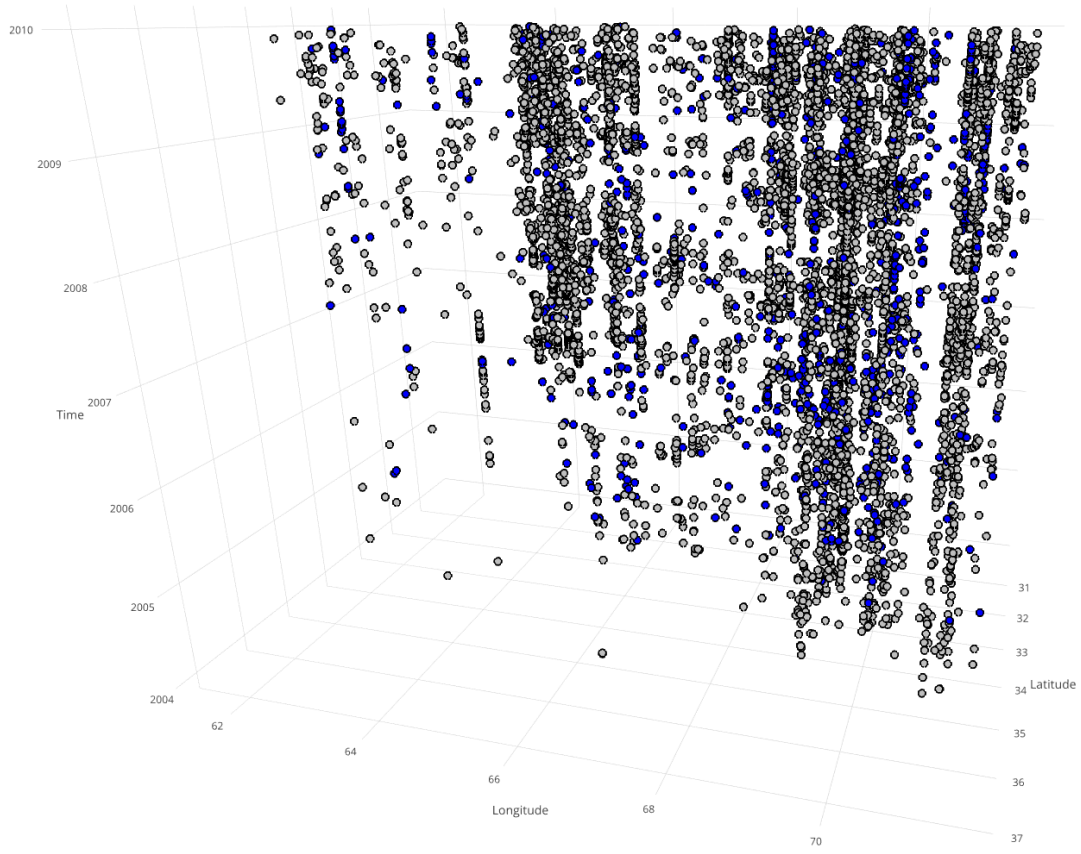


Figure A.6: Comparison of locations in space and time of dependent events and treatment intervention events. The dependent events are more sparse than treatment intervention events. Heavier areas of events over the entirety of the time period are located around larger cities in Afghanistan.

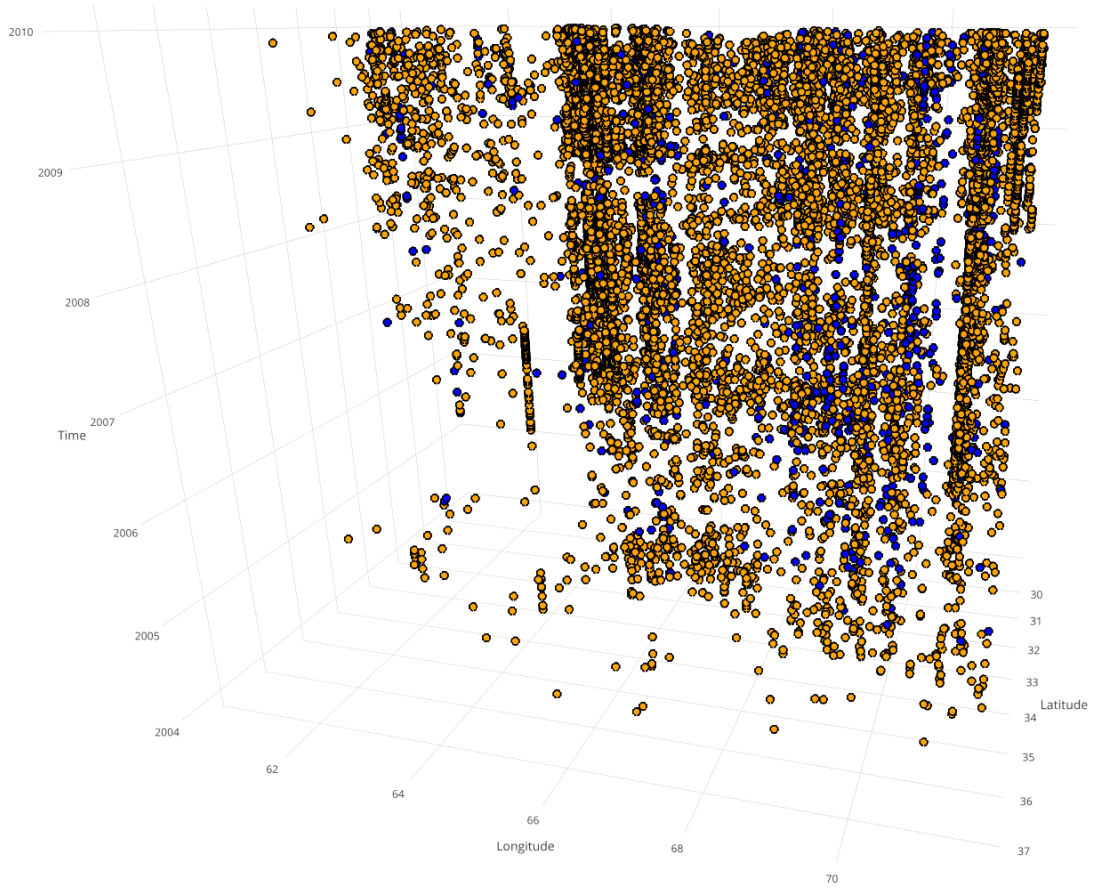


Figure A.7: Comparison of locations in space and time of dependent events and control intervention events. The dependent events are more sparse than control intervention events. Heavier areas of events over the entirety of the time period are located around larger cities in Afghanistan.

Appendix B

Appendix B: Defining the Interference Effect With Spatial Point Process Theory

Below are the additional simulation scenarios for all spatial windows examined that were not included in Chapter 4.

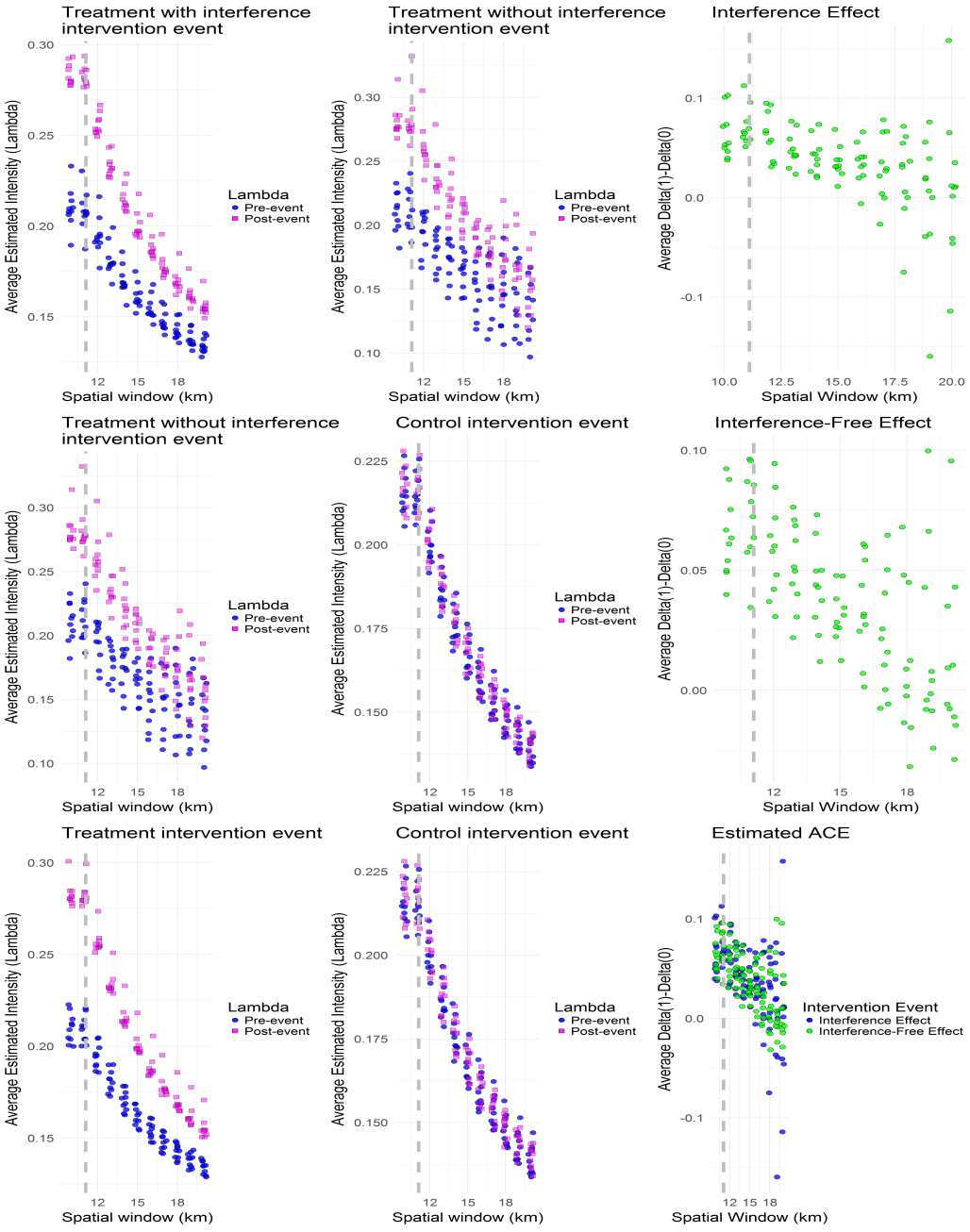


Figure B.1: The top panel includes the average $\widehat{\lambda}_1(1,1)$ and $\widehat{\lambda}_0(1,1)$ on the left, average $\widehat{\lambda}_1(1,0)$ and $\widehat{\lambda}_0(1,0)$ in the middle, and average $\widehat{\Delta}(1,1)$ - $\widehat{\Delta}(1,0)$ on the right. The 100 treatment intervention events make up those with and without interference to illustrate the interference effect. The middle panel illustrates components of the interference-free effect, with average $\widehat{\lambda}_1(1,0)$ and $\widehat{\lambda}_0(1,0)$ on the left, average $\widehat{\lambda}_1(0)$ and $\widehat{\lambda}_0(0)$ in the middle, and average $\widehat{\Delta}(1,0) - \widehat{\Delta}(0)$ on the right, with 200 control intervention events and the non-interference treatment intervention events.The overall average $\widehat{\lambda}_1(1)$, $\widehat{\lambda}_0(1)$, $\widehat{\lambda}_1(0)$, and $\widehat{\lambda}_0(0)$ are shown in the bottom left and middle, respectively. Both effects are plotted in the bottom right. The dotted gray lines represent the true spatial window size in all panels and plots.

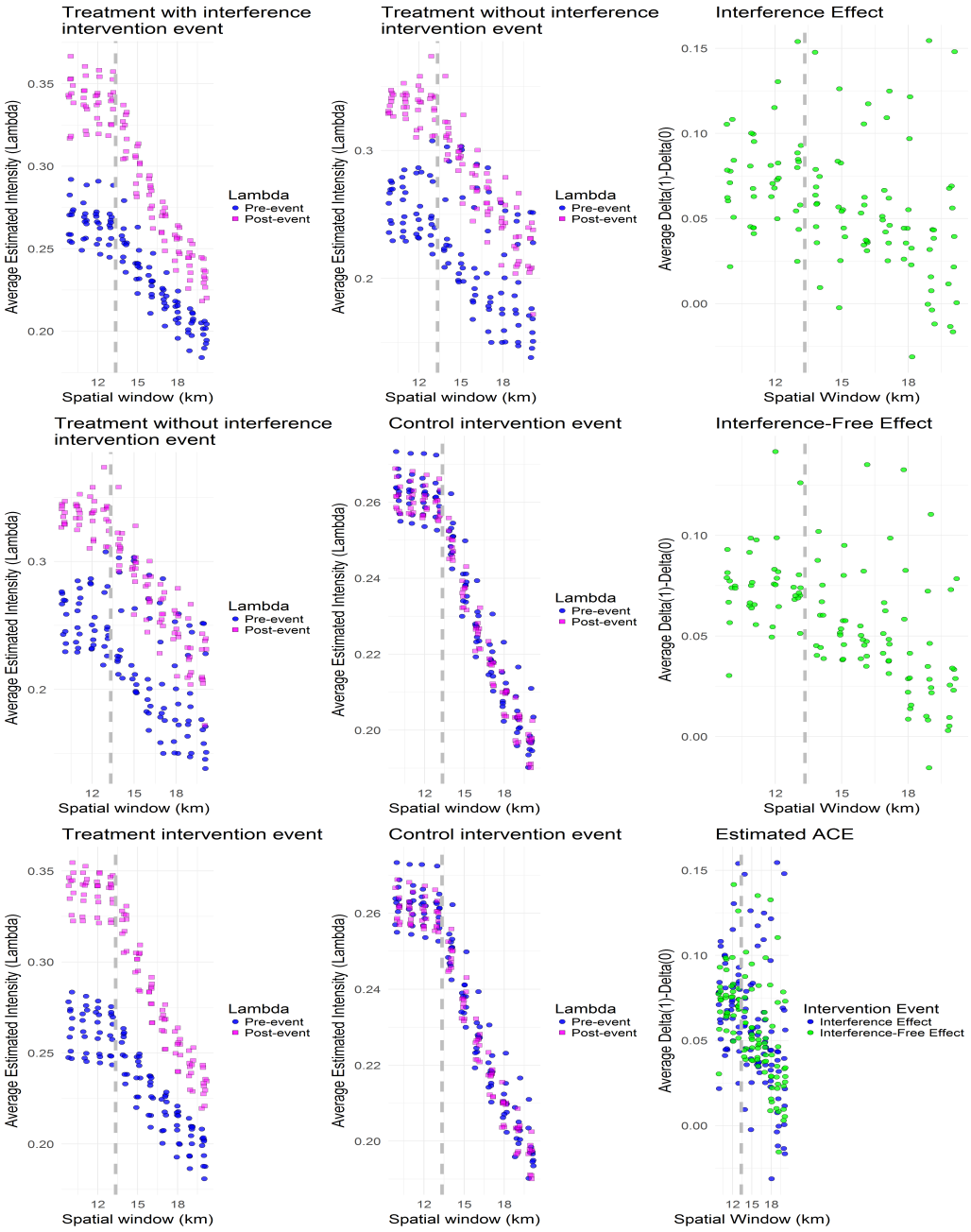


Figure B.2: The top panel includes the average $\widehat{\lambda}_1(1,1)$ and $\widehat{\lambda}_0(1,1)$ on the left, average $\widehat{\lambda}_1(1,0)$ and $\widehat{\lambda}_0(1,0)$ in the middle, and average $\widehat{\Delta}(1,1)$ - $\widehat{\Delta}(1,0)$ on the right. The 100 treatment intervention events make up those with and without interference to illustrate the interference effect. The middle panel illustrates components of the interference-free effect, with average $\widehat{\lambda}_1(1,0)$ and $\widehat{\lambda}_0(1,0)$ on the left, average $\widehat{\lambda}_1(0)$ and $\widehat{\lambda}_0(0)$ in the middle, and average $\widehat{\Delta}(1,0) - \widehat{\Delta}(0)$ on the right, with 200 control intervention events and the non-interference treatment intervention events. The overall average $\widehat{\lambda}_1(1)$, $\widehat{\lambda}_0(1)$, $\widehat{\lambda}_1(0)$, and $\widehat{\lambda}_0(0)$ are shown in the bottom left and middle, respectively. Both effects are plotted in the bottom right. The dotted gray lines represent the true spatial window size in all panels and plots.

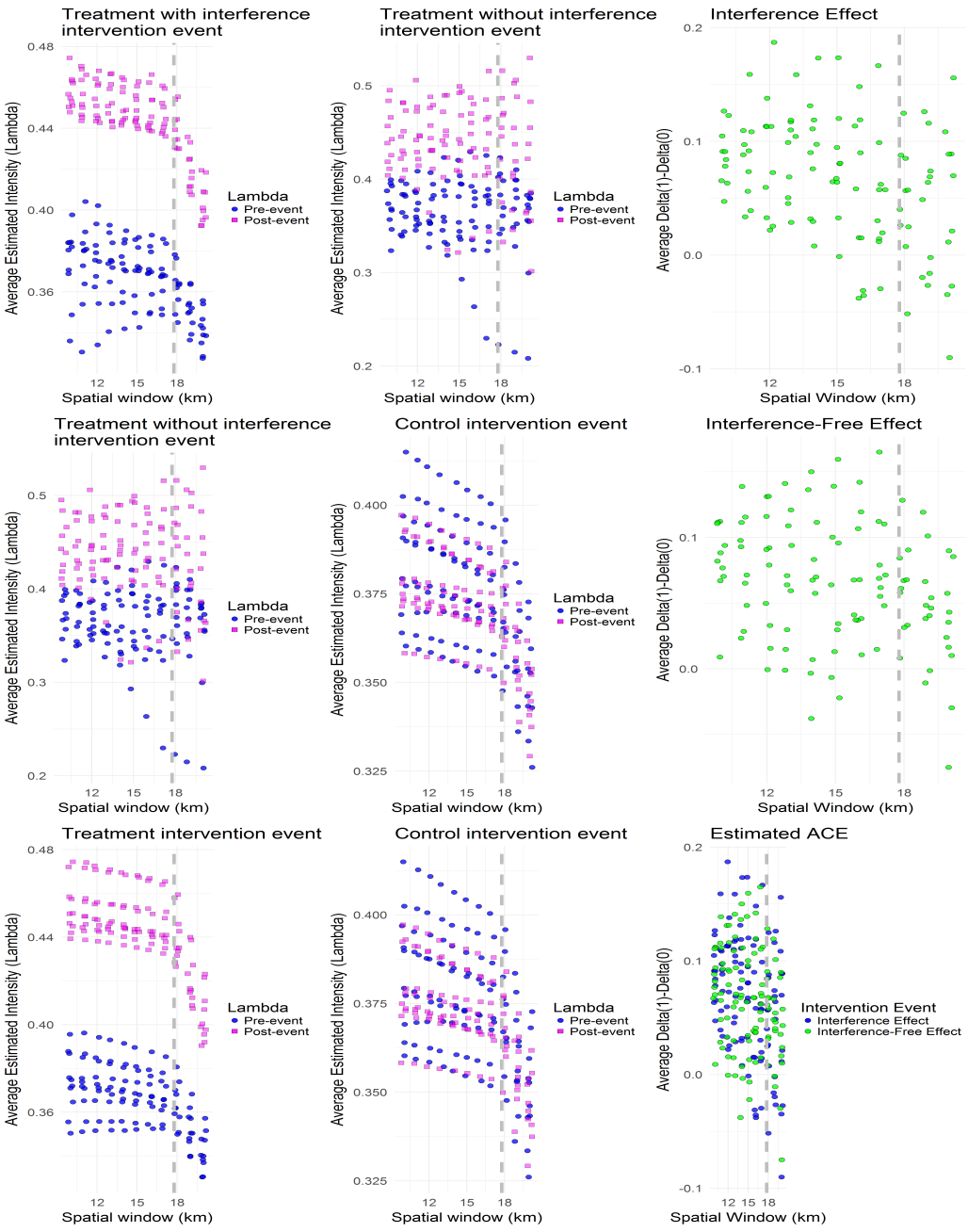


Figure B.3: The top panel includes the average $\widehat{\lambda}_1(1,1)$ and $\widehat{\lambda}_0(1,1)$ on the left, average $\widehat{\lambda}_1(1,0)$ and $\widehat{\lambda}_0(1,0)$ in the middle, and average $\widehat{\Delta}(1,1)$ - $\widehat{\Delta}(1,0)$ on the right. The 100 treatment intervention events make up those with and without interference to illustrate the interference effect. The middle panel illustrates components of the interference-free effect, with average $\widehat{\lambda}_1(1,0)$ and $\widehat{\lambda}_0(1,0)$ on the left, average $\widehat{\lambda}_1(0)$ and $\widehat{\lambda}_0(0)$ in the middle, and average $\widehat{\Delta}(1,0) - \widehat{\Delta}(0)$ on the right, with 200 control intervention events and the non-interference treatment intervention events. Lastly, the overall average $\widehat{\lambda}_1(1)$, $\widehat{\lambda}_0(1)$, $\widehat{\lambda}_1(0)$, and $\widehat{\lambda}_0(0)$ are shown in the bottom left and middle, respectively. Both the interference and interference-free effects are plotted in the bottom right. The dotted gray lines represent the true spatial window size in all panels and plots.

Bibliography

Angrist, J. D. and Pischke, J.-S. (2009), Mostly Harmless Econometrics, Princeton University Press, Princeton, N.J.

Austin, P. C. (2011), 'An introduction to propensity score methods for reducing the effects of confounding in observational studies', Multivariate Behavioral Research **46**(3), 399–424.

Braithwaite, A. and Johnson, S. D. (2012), 'Space-time modeling of insurgency and counterinsurgency in Iraq', Journal of Quantitative Criminology **28**(1), 31–48.

Cattaneo, M. D., Frandsen, B. R. and Titiunik, R. (2015), 'Randomization inference in the regression discontinuity design: An application to party advantages in the us senate', Journal of Causal Inference **3**(1), 1–24.

Chagas, A. L. S., Toneto, R. and Azzoni, C. R. (2012), 'A spatial propensity score matching evaluation of the social impacts of sugarcane growing on municipalities in Brazil', International Regional Science Review **35**(1), 48–69.

CIESIN and CIAT (2005), 'Gridded population of the world version 3 (gpwv3): Population density grids', Palisades, NY: Socioeconomic Data and Applications Center (SEDAC), Columbia University . Available at `http://sedac.ciesin.columbia.edu/gpw`.

Cole, S. R. and Frangakis, C. E. (2009), 'The consistency statement in causal inference: A definition or an assumption?', Epidemiology **20**(1), 3–5.

Cressie, N. A. (1993), Statistics for Spatial Data, Wiley Online Library.

Diggle, P. J. (1983), Statistical Analysis of Spatial Point Patterns, Academic Press.

Downes, A. B. (2007), 'Draining the sea by filling the graves: Investigating the effectiveness of indiscriminate violence as a counterinsurgency strategy', Civil Wars **9**(4), 420–444.

Ellsberg, D. (1970), 'Revolutionary judo'.

Frangakis, C. E. and Rubin, D. B. (2002), 'Principal stratification in causal inference', Biometrics **58**(1), 21–29.

Galles, D. and Pearl, J. (1998), 'An axiomatic characterization of causal counterfactuals', Foundations of Science **3**(1), 151–182.

Gesch, D. B., Verdin, K. L. and Greenlee, S. K. (1999), 'New land surface digital elevation model covers the earth', EOS, Transactions American Geophysical Union **80**(6), 69–70.

Greenland, S. and Robins, J. M. (1986), 'Identifiability, exchangeability, and epidemiological confounding', International Journal of Epidemiology **15**(3), 413–419.

Hainmueller, J. et al. (2012), 'Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies', Political Analysis **20**(01), 25–46.

Holland, P. W. (1986), 'Statistics and causal inference', Journal of the American Statistical Association **81**(396), 945–960.

Iacus, S. M., King, G. and Porro, G. (2012), 'Causal inference without balance checking: Coarsened exact matching', Political Analysis **20**(1), 1–24.

Imai, K. and Ratkovic, M. (2014), 'Covariate balancing propensity score', Journal of the Royal Statistical Society: Series B (Statistical Methodology) **76**(1), 243–263.

Imai, K. and Van Dyk, D. A. (2004), 'Causal inference with general treatment regimes: Generalizing the propensity score', Journal of the American Statistical Association **99**(467), 854–866.

Imbens, G. W. (2000), 'The role of the propensity score in estimating dose-response functions', Biometrika **87**(3), 706–710.

Imbens, G. W. and Lemieux, T. (2008), 'Regression discontinuity designs: A guide to practice', Journal of Econometrics **142**(2), 615–635.

Inoue, R. and Shimizu, E. (2006), 'A new algorithm for continuous area cartogram construction with triangulation of regions and restriction on bearing changes of edges', Cartography and Geographic Information Science **33**(2), 115–125.

Kalyvas, S. N. (2006), The Logic of Violence in Civil War, Cambridge University Press.

Kalyvas, S. N. and Kocher, M. A. (2009), 'The dynamics of vviolence in Vietnam: An analysis of the hamlet evaluation system (HES)', Journal of Peace Research **46**(3), 335–355.

Kao, E. K. (2017), 'Causal inference under network interference: A framework for experiments on social networks', arXiv preprint arXiv:1708.08522 .

Keele, L. J. and Titiunik, R. (2014), 'Geographic boundaries as regression discontinuities', Political Analysis **23**(1), 127–155.

Keele, L. and Titiunik, R. (2016), 'Natural experiments based on geography', Political Science Research and Methods **4**(1), 65–95.

Keele, L. and Titiunik, R. (2017), 'Geographic natural experiments with interference: The effect of all-mail voting on turnout in Colorado'.

Keele, L., Titiunik, R. and Zubizarreta, J. R. (2015), 'Enhancing a geographic regression discontinuity design through matching to estimate the effect of ballot initiatives on voter turnout', Journal of the Royal Statistical Society: Series A (Statistics in Society) **178**(1), 223–239.

Kolak, M. A. (2017), Policy and Place: A Spatial Data Science Framework for Research and Decision-Making, PhD thesis, Arizona State University.

Kulldorff, M. (1997), 'A spatial scan statistic', Communications in Statistics-Theory and Methods **26**(6), 1481–1496.

Lee, D. S. (2008), 'Randomized experiments from non-random selection in us house elections', Journal of Econometrics **142**(2), 675–697.

Lee, D. S. and Lemieux, T. (2010), 'Regression discontinuity designs in economics', Journal of Economic Literature **48**(2), 281–355.

Linke, A. M., Witmer, F. D. and O'Loughlin, J. (2012), 'Space-time granger analysis of the war in iraq: A study of coalition and insurgent action-reaction', International Interactions **38**(4), 402–425.

Lyall, J. (2009), 'Does indiscriminate violence incite insurgent attacks? evidence from Chechnya', Journal of Conflict Resolution **53**(3), 331–362.

Nordhaus, W., Azam, Q., Corderi, D., Hood, K., Victor, N. M., Mohammed, M., Miltner, A. and Weiss, J. (2006), 'The g-econ database on gridded output: Methods and data', Yale University, New Haven **6**.

O'Loughlin, J., Witmer, F. D., Linke, A. M. and Thorwardson, N. (2010), 'Peering into the fog of war: The geography of the Wikileaks Afghanistan War logs, 2004-2009', Eurasian Geography and Economics **51**(4), 472–495.

Openshaw, S. and Taylor, P. (1979), 'A million or so correlation coefficients', Statistical Methods in the Spatial Sciences pp. 127–144.

Papadogeorgou, G., Choirat, C. and Zigler, C. (2016), 'Adjusting for unmeasured spatial confounding with distance adjusted propensity score matching', arXiv preprint arXiv:1610.07583 .

Ripley, B. D. (1991), Statistical Inference for Spatial Processes, Cambridge University Press.

Robins, J. (1987), 'A graphical approach to the identification and estimation of causal parameters in mortality studies with sustained exposure periods', Journal of Chronic Diseases **40**, 139S–161S.

Rosenbaum, P. R. and Rubin, D. B. (1983), 'The central role of the propensity score in observational studies for causal effects', Biometrika **70**(1), 41–55.

Rubin, D. B. (1974), 'Estimating causal effects of treatments in randomized and nonrandomized studies', Journal of Educational Psychology **66**(5), 688.

Rubin, D. B. (1980), 'Discussion of "randomization analysis of experimental data in the fisher randomization test", by Basu', Journal of the American Statistical Association **75**, 591–593.

Rubin, D. B. (1986), 'Comment: Which ifs have causal answers', Journal of the American Statistical Association **81**(396), 961–962.

Schutte, S. (2017), 'Violence and civilian loyalties: Evidence from Afghanistan', Journal of Conflict Resolution **61**(8), 1595–1625.

Schutte, S. and Donnay, K. (2014), 'Matched wake analysis: Finding causal relationships in spatiotemporal event data', Political Geography **41**, 1–10.

Selvin, H. C. (1958), 'Durkheim's suicide and problems of empirical research', American Journal of Sociology **63**(6), 607–619.

Sheehan, N. (1998), A Bright Shining Lie: John Paul Vann and America in Vietnam, Random House.

SIGACT (2010), 'Wikileaks: The Iraq War logs', Available online at `http://wikileaks.org`. Accessed 2017-08-30.

Simpson, E. H. (1951), 'The interpretation of interaction in contingency tables', Journal of the Royal Statistical Society. Series B (Methodological) pp. 238–241.

Sinclair, B., McConnell, M. and Green, D. P. (2012), 'Detecting spillover effects: Design and analysis of multilevel experiments', American Journal of Political Science **56**(4), 1055–1069.

Stoyan, D., Kendall, W. S. and Mecke, J. (1995), Stochastic Geometry and Its Applications, John Wiley & Sons.

Stuart, E. A. (2010), 'Matching methods for causal inference: A review and a look forward', Statistical Science: A Review Journal of the Institute of Mathematical Statistics **25**(1), 1.

Toulis, P. and Kao, E. (2013), Estimation of causal peer influence effects, in 'International Conference on Machine Learning', pp. 1489–1497.

VanderWeele, T. J. (2011), 'Principal stratification-uses and limitations', The International Journal of Biostatistics **7**(1), 1–14.

Verbitsky-Savitz, N. and Raudenbush, S. W. (2012), 'Causal inference under interference in spatial settings: A case study evaluating community policing program in Chicago', Epidemiologic Methods **1**(1), 107–130.

Waller, L. A. and Gotway, C. A. (2004), Applied Spatial Statistics for Public Health Data, Vol. 368, John Wiley & Sons.

Zigler, C. M., Dominici, F. and Wang, Y. (2012), 'Estimating causal effects of air quality regulations using principal stratification for spatially correlated multivariate intermediate outcomes', Biostatistics **13**(2), 289–302.