

CAUSAL INFERENCE WITH SPATIO-TEMPORAL DATA: ESTIMATING THE EFFECTS OF AIRSTRIKES ON INSURGENT VIOLENCE IN IRAQ*

Georgia Papadogeorgou[†] Kosuke Imai[‡] Jason Lyall[§] Fan Li[¶]

April 1, 2020

Abstract

Although many causal processes have spatial and temporal dimensions, the classical causal inference framework is not directly applicable when the treatment and outcome variables are generated by spatio-temporal point processes. The methodological difficulty primarily arises from the existence of an infinite number of possible treatment and outcome event locations at each point in time. In this paper, we consider a setting where the spatial coordinates of the treatment and outcome events are observed at discrete time periods. We extend the potential outcomes framework by formulating the treatment point process as a stochastic intervention strategy. Our causal estimands include the expected number of outcome events that would occur in an area of interest under a particular stochastic treatment assignment strategy. We develop an estimation technique by applying the inverse probability of treatment weighting method to the spatially-smoothed outcome surfaces. We show that under a set of assumptions, the proposed estimator is consistent and asymptotically normal as the number of time periods goes to infinity. Our motivating application is the evaluation of the effects of American airstrikes on insurgent violence in Iraq from February 2007 to July 2008. We consider interventions that alter the intensity and target areas of airstrikes. We find that increasing the average number of airstrikes from 1 to 6 per day for seven consecutive days increases all types of insurgent violence.

Keywords: inverse probability of treatment weighting, point process, spatial propensity score, stochastic intervention, time-varying confounders

*Lyall gratefully acknowledges financial support from the Air Force Office of Scientific Research (Grant #FA9550-14-1-0072). The findings and conclusions reached here do not reflect the official views or policy of the United States Government or Air Force.

[†]Postdoctoral Associate, Department of Statistical Science, Duke University, Durham NC 27708. Email: gp118@duke.edu, URL: <https://gpapadogeorgou.netlify.com>

[‡]Professor, Department of Government and Department of Statistics, Harvard University. 1737 Cambridge Street, Institute for Quantitative Social Science, Cambridge MA, 02138. Email: imai@Harvard.Edu, URL: <https://imai.fas.harvard.edu>

[§]James Wright Chair in Transnational Studies and Associate Professor, Department of Government, Dartmouth College, Hanover, NH 03755. Email: jason.lyall@dartmouth.edu, URL: www.jasonlyall.com

[¶]Associate Professor, Department of Statistical Science, Duke University, Durham, NC 27708. Email: fl35@duke.edu, URL: <http://www2.stat.duke.edu/~fl35>

1 Introduction

Many causal processes involve both spatial and temporal dimensions. Examples include the environmental impact of newly constructed factories, the economic and social effects of refugee influx, and the various consequences of disease outbreaks. When the treatment and outcome variables are generated by spatio-temporal point processes, the primary methodological challenge arises from the fact that there exists an infinite number of possible treatment and outcome event locations at each point in time.

Unfortunately, the classical causal inference framework that dates back to [Neyman \(1923\)](#) and [Fisher \(1935\)](#) is not directly applicable to such settings. Indeed, the standard causal inference approaches assume that the number of units which can receive the treatment is finite (e.g., [Rubin, 1974](#); [Robins, 1997](#)). Although some studies develop a continuous time causal inference framework, they do not incorporate a spatial dimension (e.g., [Gill and Robins, 2001](#); [Zhang *et al.*, 2011](#)). In addition, a small number of researches have proposed causal inference methods for analyzing fMRI experiments, which have both spatial and temporal dimensions. For example, [Luo *et al.* \(2012\)](#) apply randomized inference, while [Sobel and Lindquist \(2014\)](#) employ structural modeling.

In this paper, we consider a setting where the treatment and outcome events are assumed to be generated by spatio-temporal point processes (Section 3). The proposed methodology is based on a single time series of spatial patterns of treatment and outcome variables, and builds upon three strands of the causal inference literature: interference, stochastic interventions, and time series. First, we address the possibility that treatments might affect outcomes at a future time period and at different locations in arbitrary ways. Although some scholars have considered unstructured interference, they assume non-spatial and cross-sectional settings (see [Basse and Airoidi, 2018](#); [Sävje *et al.*, 2019](#), and references therein). [Aronow *et al.* \(2019\)](#) study spatial randomized experiments in a cross-section setting under the assumption that the number of potential intervention locations is finite and

their spatial coordinates are known. In contrast, the proposed spatio-temporal causal inference framework allows for *temporally and spatially unstructured interference* over an infinite number of locations and across a finite number of time periods.

Second, instead of separately estimating the causal effects of treatment received at each location, we consider the impacts of different *treatment assignment strategies*, defined formally as the intervention distributions over treatment point patterns. Stochastic interventions have been used to deal with challenging causal inference problems (Díaz Muñoz and van der Laan, 2012), including the violation of positivity assumption (Kennedy, 2019), mediation analysis (Lok, 2016), and multiple treatments (Imai and Jiang, 2019). We show that this approach is also useful for dealing with spatio-temporal treatments and outcomes.

Lastly, we do not impose any restriction on the spatial patterns of interference. As such, our estimation does not assume the separation of units into minimally interacting sets (e.g., Tchetgen Tchetgen *et al.*, 2017). Instead, we view the region of interest as a single unit where outcome events are observed over time. We then develop estimation and inferential methods by building on the time-series approach in the causal inference literature (Bojinov and Shephard, 2019). We provide a sufficient set of assumptions that enable consistent estimation based on a spatially-smoothed inverse probability weighting estimator (Section 4). We conduct simulation studies to assess the finite sample performance of the proposed estimator (Section 5).

Our motivating application is the evaluation of the effects of American airstrikes on insurgent violence in Iraq from February 2007 to July 2008 (Section 2). We consider all airstrikes occurring within a day anywhere in Iraq as a *treatment pattern*, with bombed locations referred to as the *treatment-active locations*. Instead of focusing on the causal effects of each airstrike, we consider the impact of different *airstrike strategies*, defined formally as the distributions of airstrikes over the entire Iraq. The proposed methodology offers a data-driven approach to evaluate the effectiveness of airstrike strategies on the

complex spatial structure of insurgent violence across Iraq.

By using various specifications of stochastic interventions, we examine how different airstrike strategies affect subsequent insurgent attacks (Section 6). First, we show that a higher frequency of airstrikes, without modifying their spatial distribution, increases the number of insurgent attacks, especially near Baghdad, Mosul, and the roads between them. We also find that changing the focal point of airstrikes to Baghdad without modifying the overall frequency shifts insurgent attacks from Baghdad to Mosul and its environs. Finally, our analysis suggests that increasing the number of airstrikes for a few days may initially reduce attacks but ultimately increases them over the long run.

2 Motivating Application: Airstrikes and Insurgent Activities in Iraq

2.1 Background

Airstrikes have emerged as a principal tool for defeating insurgent and terrorist organizations in civil wars around the globe. In the past decade alone, the United States has conducted sustained air campaigns in at least six different countries, including Afghanistan, Iraq, and Syria. Although it has been shown that civilians have all-too-often borne the brunt of these airstrikes (Lyll, 2019b), we have few rigorous studies of airstrikes and their effects on subsequent insurgent violence. Even these few studies have largely reached opposite conclusions, with some claiming that airpower reduces insurgent attacks while others arguing they spark escalatory spirals of increased violence (e.g., Lyll, 2019a; Mir and Moore, 2019; Dell and Querubin, 2018; Kocher *et al.*, 2011). Regardless of conclusions, all of the existing studies discretize a continuous space into arbitrary geographical units and make simplifying assumptions about the patterns of spatial and temporal interference.

We enter this debate by examining the American air campaign in Iraq. We use declassified US Air Force data on airstrikes and shows of force (simulated airstrikes where no weapons are released) for the February 2007 to July 2008 period. The period in question coincides with the “surge” of American forces and airpower designed to destroy multiple

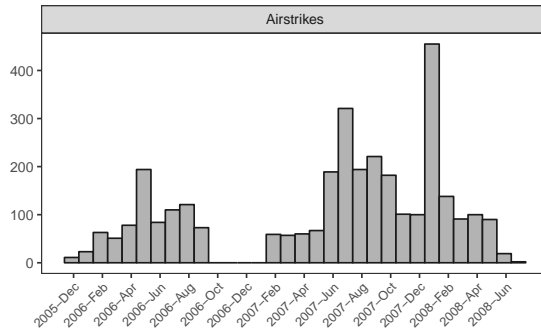
Sunni and Shia insurgent organizations in a bid to turn the war's tide.

Aircraft were assigned to bomb targets via two channels. First, airstrikes were authorized in response to American forces coming under insurgent attack. These close air support (CAS) missions represented the vast majority of airstrikes in 2007-08. Second, a small percentage (about 5%) of airstrikes were pre-planned against high-value targets, typically insurgent commanders, whose presence had been detected from intercepted communications or human intelligence. In each case, airstrikes were driven by insurgent attacks that were either ongoing or had occurred in the recent past in a given location. As a result, the models used later in this paper adjust for prior patterns of insurgent violence in a given location for several short-term windows.

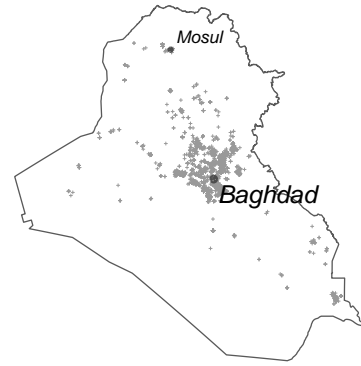
We also account for prior air operations, including shows of force, by American and allied aircraft. Insurgent violence in Iraq is also driven by settlement patterns and transportation networks. Our models therefore include population size and location of Iraqi villages and cities as well as proximity to road networks, where the majority of insurgent attacks were conducted against American convoys. Finally, prior reconstruction spending might also condition the location of airstrikes. Aid is often provided in tandem with airstrikes to drive out insurgents, while these same insurgents often attack aid sites to derail American hearts-and-minds strategies. Taken together, these four factors — recent insurgent attacks, the presence of American forces, settlement patterns, and prior aid spending — drove decisions about the location and severity of airstrikes.

2.2 Data

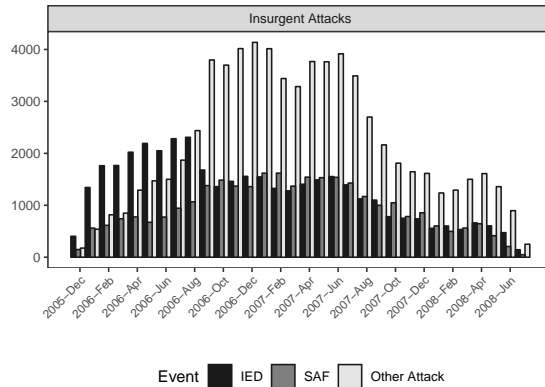
Figure 1 summarizes the spatial and temporal distributions of airstrikes (treatment variable) and insurgent violence (outcome variable). Figure 1a presents the temporal distribution of airstrikes recorded by the US Air Force each month. There were a total of 3,254 airstrikes during this period. Figure 1b plots the spatial density of these airstrikes across Iraq, with spatial clustering observed around Baghdad and the neighboring “Sunni Trian-



(a) Kinetic airstrikes over time



(b) Airstrikes over space



(c) Insurgent violence over time



(d) Insurgent violence over space

Figure 1: Distribution of the treatment and outcome point processes over time and space. Plots (b) and (d) show the locations of airstrikes and insurgent attacks during the time period January 1, 2006 to June 26, 2008. Insurgent attacks are sorted into one of three categories: Improvised Explosive Devices (IEDs), Small Arms Fire (SAF), and other attacks. Airstrikes from October 2006 to January 2007 are missing from Air Force records.

gle,” a hotspot of insurgency. Figure 1c plots the monthly distribution of insurgent attacks by type: Improvised Explosive Devices (IEDs), small arms fire (SAF), and other attacks. A total of 140,977 insurgent attacks recorded by the US Army’s CIDNE database during this time period. Finally, Figure 1d plots the locations of insurgent attacks across Iraq. Baghdad, the Sunni Triangle, and the highway leading north to Mosul are all starkly illustrated.

3 Causal Inference Framework for Spatio-temporal Data

In this section, we propose a causal inference framework for spatio-temporal point processes. We describe the setup, and define causal estimands using stochastic interventions.

3.1 The Setup

We represent the locations of airstrikes for each time period (e.g., day) as a spatial point pattern measured at time $t \in \mathcal{T} = \{1, 2, \dots, T\}$ where T is the total number of the discrete time periods. Let $W_t(s)$ denote the binary treatment variable at location s for time period t , indicating whether or not the site receives the treatment during the time period. We use W_t as a shorthand for $W_t(\Omega)$, which evaluates the binary treatment variable $W_t(s)$ for each element s of the set of all locations Ω that may receive the treatment. In addition, \mathcal{W} represents the set of all possible point patterns at each time period where, for simplicity, we assume that this set does not vary across time periods, i.e., $W_t \in \mathcal{W}$ for each t . Finally, the set of *treatment-active locations*, i.e., the sites that receive the treatment, at time t is denoted by $S_{W_t} = \{s \in \Omega : W_t(s) = 1\}$. We assume that the number of treatment-active locations is finite for each time period, i.e., $|S_{W_t}| < \infty$ for any t . In our study, the treatment-active locations correspond to the set of coordinates of airstrikes.

We use w_t to denote a realization of W_t and $\bar{\mathbf{w}}_t = (w_1, w_2, \dots, w_t)$ to denote the history of treatment point pattern realizations from time 1 through time t . Let $Y_t(\bar{\mathbf{w}}_t)$ represent the potential outcome at time $t \in \mathcal{T}$ for any given treatment sequence $\bar{\mathbf{w}}_t \in \mathcal{W}^t = \mathcal{W} \times \dots \times \mathcal{W}$, depending on *all* previous treatments. Similar to the treatment, $Y_t(\bar{\mathbf{w}}_t)$ represents a point pattern with locations $S_{Y_t(\bar{\mathbf{w}}_t)}$, which are referred to as the *outcome-active locations*. In our study, $S_{Y_t(\bar{\mathbf{w}}_t)}$ represents the locations of insurgent attacks if the patterns of airstrikes had been $\bar{\mathbf{w}}_t$. Let $\bar{\mathcal{Y}}_{t'} = \{Y_t(\bar{\mathbf{w}}_t) : \bar{\mathbf{w}}_t \in \mathcal{W}^t, t \leq t'\}$ denote the collection of potential outcomes up to time period t' for all treatment sequences.

Among all of these potential outcomes for time t , we only observe the one corre-

sponding to the observed treatment sequence \bar{w}_t^{obs} , denoted by $Y_t^{obs} = Y_t(\bar{w}_t^{obs})$. We use $\bar{Y}_t^{obs} = \{Y_1^{obs}, \dots, Y_t^{obs}\}$ to represent the collection of observed outcomes up to and including time period t . In addition, let \mathbf{X}_{t-1} be the set of possibly time-varying confounders that are realized prior to W_t but after W_{t-1} for $t \geq 1$, and $\bar{\mathcal{X}}_{t'} = \{\mathbf{X}_t(\bar{w}_t) : \bar{w}_t \in \mathcal{W}^t, t \leq t'\}$ be the set of potential values of \mathbf{X} under any possible treatment history \bar{w}_t . Also let $\bar{H}_{t-1} = \{\bar{W}_{t-1}, \bar{Y}_{t-1}^{obs}, \bar{\mathbf{X}}_{t-1}^{obs}\}$ denote all observed history up to time $t - 1$.

3.2 Causal Estimands under Stochastic Interventions

A notion central to our proposed causal inference framework is *stochastic interventions*. Instead of setting a treatment variable to a fixed value, a stochastic intervention specifies the probability distribution that generates the treatment. Although our framework accommodates a variety of intervention distributions, we consider intervention distributions based on Poisson point processes, which are fully characterized by an intensity function $h : \Omega \rightarrow [0, \infty)$. For example, a homogeneous Poisson point process with $h(s) = h$ for all $s \in \Omega$, implies that the number of treatment-active locations follows a $\text{Poisson}(h|\Omega|)$ distribution, with locations distributed independently and uniformly over Ω . In general, the specification of the stochastic intervention should be motivated by policy or scientific objectives. Such examples in the context of our study are given in Section 6.1.

Our causal estimands are the expected number of (potential) outcome-active locations under a specific stochastic intervention of interest, and the comparison of such quantities under different intervention distributions. We begin by defining the causal estimands for a stochastic intervention taking place over a single time period. Let F_h denote the distribution of a spatial point process with intensity h . Also, let $N_B(\cdot)$ denote a counting measure on a region $B \subset \Omega$. Then, we can define the expected number of outcome-active locations for a region B at time t as

$$\bar{N}_{Bt}(F_h) = \int_{\mathcal{W}} N_B(Y_t(\bar{w}_{t-1}^{obs}, w_t)) dF_h(w_t) = \int_{\mathcal{W}} |S_{Y_t(\bar{w}_{t-1}^{obs}, w_t)} \cap B| dF_h(w_t). \quad (1)$$

In our application, this quantity represents the expected number of insurgent attacks within a region of Iraq B if the airstrikes at time t were to follow the point process specified by F_h , given the observed history of airstrikes up to time $t - 1$.

We now extend the above estimand to an intervention taking place over M consecutive time periods. Consider an intervention, denoted by $F_h = F_{h_1} \times \dots \times F_{h_M}$, under which the treatment at time t is assigned according to F_{h_1} , at time $t - 1$ according to F_{h_2} , continuing until time period $t - M + 1$ for which treatment is assigned according to F_{h_M} . A treatment path based on this intervention is displayed in Figure 2(a). Then, we define a general estimand as

$$\begin{aligned} \bar{N}_{Bt}(F_h) &= \int_{\mathcal{W}^M} N_B(Y_t(\bar{w}_{t-M}^{obs}, w_{t-M+1}, \dots, w_t)) dF_{h_M}(w_{t-M+1}) \dots dF_{h_1}(w_t) \\ &= \int_{\mathcal{W}^M} |S_{Y_t(\bar{w}_{t-M}^{obs}, w_{t-M+1}, \dots, w_t)} \cap B| dF_{h_M}(w_{t-M+1}) \dots dF_{h_1}(w_t). \end{aligned} \quad (2)$$

This quantity represents the expected number of outcome events within region B and at time t if the treatment point pattern during the previous M time periods were to follow the stochastic intervention with distribution F_h . Treatments during the initial $t - M$ time periods were the same as observed. A special case of F_h assumes that treatments during the M time periods are independent and identically distributed draws from the same distribution

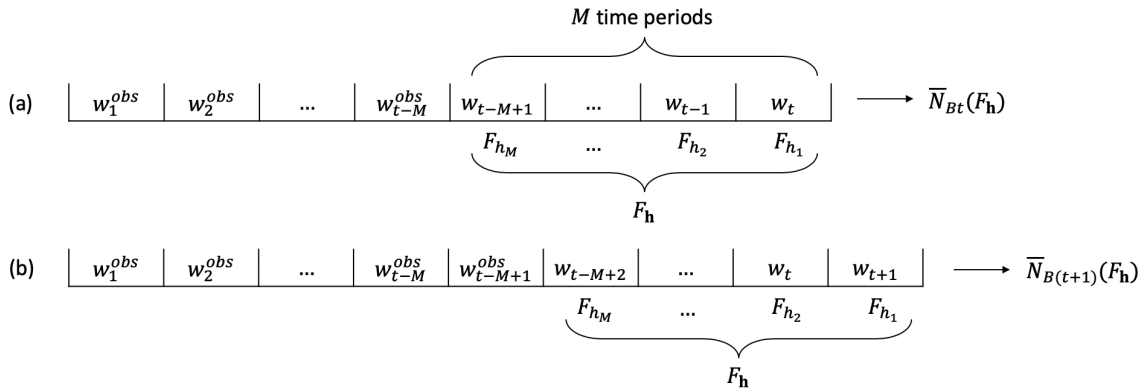


Figure 2: Graphical Illustration of Stochastic Intervention over Multiple Time Periods for Time Period t and $t + 1$. Under intervention F_h , treatments during time periods $t - M + 1, \dots, t - 1, t$ are assigned according to distributions $F_{h_M}, \dots, F_{h_2}, F_{h_1}$.

F_h , which we denote by $F_{\mathbf{h}} = F_h^M$.

Given the above setup, we define the average treatment effect of stochastic intervention $F_{\mathbf{h}'}$ versus $F_{\mathbf{h}''}$ for a region B at time t as

$$\tau_{Bt}(F_{\mathbf{h}'}, F_{\mathbf{h}''}) = \overline{N}_{Bt}(F_{\mathbf{h}''}) - \overline{N}_{Bt}(F_{\mathbf{h}'}), \quad (3)$$

where $\mathbf{h}' = (h'_1, h'_2, \dots, h'_M)$ represents a collection of treatment intensities over M consecutive time periods (similarly for \mathbf{h}'').

We further define the average, over time periods $t = M, M + 1, \dots, T$, of the expected potential outcome for region B at each time period if treatments during the M proceeding time periods arose from $F_{\mathbf{h}}$ as

$$\overline{N}_B(F_{\mathbf{h}}) = \frac{1}{T - M + 1} \sum_{t=M}^T \overline{N}_{Bt}(F_{\mathbf{h}}). \quad (4)$$

Figure 2 shows two of the terms averaged in (4), $\overline{N}_{Bt}(F_{\mathbf{h}})$ and $\overline{N}_{B(t+1)}(F_{\mathbf{h}})$. For $\overline{N}_{Bt}(F_{\mathbf{h}})$, treatments up to $t - M$ are set to their observed values, and treatments at time periods $t - M + 1, \dots, t$ are drawn from $F_{\mathbf{h}}$. The same definition applies to $\overline{N}_{B(t+1)}(F_{\mathbf{h}})$, but intervention time periods are shifted by 1: treatments up to $t - M + 1$ are set to their observed values, and treatments during time periods $t - M + 2, \dots, t + 1$ are drawn from $F_{\mathbf{h}}$. In (4), the summation starts at $t = M$ since the quantity $\overline{N}_{Bt}(F_{\mathbf{h}})$ assumes that there exist M prior time periods during which treatments are intervened on.

Based on $\overline{N}_B(F_{\mathbf{h}})$, we define the causal effect of intervention $F_{\mathbf{h}'}$ versus $F_{\mathbf{h}''}$ as

$$\tau_B(F_{\mathbf{h}'}, F_{\mathbf{h}''}) = \overline{N}_B(F_{\mathbf{h}''}) - \overline{N}_B(F_{\mathbf{h}'}) = \frac{1}{T - M + 1} \sum_{t=M}^T \tau_{Bt}(F_{\mathbf{h}'}, F_{\mathbf{h}''}). \quad (5)$$

This estimand represents the average, over time periods $t = M, M + 1, \dots, T$, of the expected change in the number of points at each time period when the observed treatment

path \bar{w}_T^{obs} was followed until $t-M$ with subsequent treatments $w_{(t-M+1):t}$ arising according to F_h . The effect size of a point pattern treatment would depend on M , and a greater value of M would be more useful for studying slow-responding outcome processes. Finally, specifying $F_{h'}$ and $F_{h''}$ such that they are identical except at M time periods ago yields the lagged effect of a treatment change, as we will discuss in more detail in Section 6.1.

4 Estimation

In this section, we first describe a set of assumptions sufficient for the identification of the causal quantities presented in the previous section. We then propose an estimator based on the inverse probability of treatment weighting and derive its asymptotic properties. All proofs are given in Appendix B.

4.1 The Assumptions

Similar to the classic causal inference settings, variants of the unconfoundedness and overlap assumptions are required for the current settings based on stochastic interventions. For simplicity, we focus on stochastic interventions with identical and independent distribution over M periods, $F_h = F_h^M$, and intensity h . Our theoretical results, however, extend straightforwardly to stochastic interventions with non-i.i.d. treatment patterns.

Assumption 1 (Unconfoundedness). *The treatment assignment at time t is independent of all, past or future, potential outcomes and potential confounders conditional on the observed history of treatments, confounders and outcomes up to time $t-1$: $W_t \perp\!\!\!\perp \{\bar{\mathcal{Y}}_T, \bar{\mathcal{X}}_T\} \mid \bar{H}_{t-1}$.*

Assumption 1 resembles the sequential ignorability assumption in longitudinal treatment settings (Robins, 1999; Robins *et al.*, 2000). The difference is that Assumption 1 requires the treatment assignment to be conditionally independent of all potential values of the time-varying confounders as well as those of the outcome variable. Unlike the typical longitudinal settings, under which a large number of units are assumed to be independent of one another, the current setting has a *single* time series of maps. Our assumption is similar to

the non-anticipating treatment assumption of [Bojinov and Shephard \(2019\)](#), who consider a setting with a single unit being exposed to a binary treatment over time. We extend their assumption by explicitly incorporating time-varying confounders.

The next assumption weakens the standard overlap assumption (also known as positivity). We define the probability *density* of treatment realization w at time t given the history, $e_t(w) = f(W_t = w \mid \overline{H}_{t-1})$, as the propensity score at time period t . Also, let f_h denote the probability density function of the stochastic intervention F_h .

Assumption 2 (Bounded relative overlap). *There exists a constant $\delta_W > 0$ such that $e_t(w) > \delta_W \cdot f_h(w)$ for all $w \in \mathcal{W}$.*

Assumption 2 ensures that all the treatment patterns which are possible under the stochastic intervention of interest can also be observed. If the density f_h of the stochastic intervention is bounded, Assumption 2 is automatically satisfied so long as the standard overlap assumption holds, i.e., there exists $\delta_W > 0$ such that $e_t(w) > \delta_W$ for all $w \in \mathcal{W}$.

4.2 The Propensity Score for Point Process Treatments

As shown below, the propensity score plays an important role in our estimation. Here, we show that the propensity score for point process treatments has two properties analogous to those of the standard propensity score ([Rosenbaum and Rubin, 1983](#)). That is, the propensity score is a balancing score, and under Assumption 1 the treatment assignment is unconfounded conditional on the propensity score.

Proposition 1. *The propensity score $e_t(w)$ is a balancing score in that it satisfies: $f(W_t = w \mid e_t(w), \overline{H}_{t-1}) = f(W_t = w \mid e_t(w))$, for all t .*

Proposition 1 allows us to empirically assess the propensity score model specification by checking the predictive power of covariates in \overline{H}_{t-1} in a model for the treatment W_t , with and without adjusting for the propensity score. For example, if a covariate significantly improves prediction in a point process model for W_t that adjusts for the estimated propensity score, then the covariate is not balanced and propensity score misspecification is likely.

Proposition 2. *Under Assumption 1, the treatment assignment at time t is unconfounded given the propensity score at time t , that is, given $W_t \perp\!\!\!\perp \{\overline{\mathcal{Y}}_T, \overline{\mathcal{X}}_T\} \mid \overline{H}_{t-1}$, we have $W_t \perp\!\!\!\perp \{\overline{\mathcal{Y}}_T, \overline{\mathcal{X}}_T\} \mid e_t(W_t)$.*

Proposition 2 shows that the potentially high-dimensional set \overline{H}_{t-1} can be reduced to one dimensional propensity score function $e_t(w)$, which is a sufficient conditioning set for identifying the causal effect of W_t .

4.3 The Estimators

We propose a class of estimators for the causal estimands defined in Section 3, which combines the inverse probability of treatment weighting (IPW) with the kernel smoothing of spatial point patterns. The estimator proceeds in two steps. First, at each time period t , the surface of outcome-active locations is spatially smoothed according to a chosen kernel. Then, this surface is weighted by the relative density of the observed treatment pattern under the stochastic intervention of interest and under the actual data generating process.

Formally, consider a univariate kernel $K : [0, \infty) \rightarrow [0, \infty)$ satisfying $\int K(u)du = 1$, and bandwidth parameter b . Let K_b denote the scaled kernel defined as $K_b(u) = b^{-1}K(u/b)$. We define $\widehat{Y}_t(F_h^M) : \Omega \rightarrow \mathbb{R}^+$ as

$$\widehat{Y}_t(F_h^M; \omega) = \prod_{j=t-M+1}^t \frac{f_h(w_j^{obs})}{e_j(w_j^{obs})} \left[\sum_{s \in S_{Y_t^{obs}}} K_b(\|\omega - s\|) \right], \quad (6)$$

where $\|\cdot\|$ denotes the Euclidean norm. The summation on the right hand side represents the spatially-smoothed version of the outcome point pattern at time period t . The product of ratios represents a weight similar to the weights in the marginal structural models, but in accordance with the stochastic intervention F_h^M : each of the M terms represents the likelihood ratio of observing the treatment w_j^{obs} in the counterfactual world of the intervention F_h versus the actual world with the observed data at a specific time period.

Assuming that the kernel K is continuous, the estimator in (6) defines a continuous surface over Ω . The continuity of $\widehat{Y}_t(F_h^M)$ allows us to conceptualize causal quantities by

viewing $\widehat{Y}_t(F_h^M)$ itself as an intensity function. This leads to the following estimator for the expected number of outcome-active locations in any region B at time t , defined in (2),

$$\widehat{N}_{Bt}(F_h^M) = \int_B \widehat{Y}_t(F_h^M; \omega) d\omega. \quad (7)$$

We use this estimator for the temporally-expected average potential outcome in (4) as

$$\widehat{N}_B(F_h^M) = \frac{1}{T - M + 1} \sum_{t=M}^T \widehat{N}_{Bt}(F_h^M), \quad (8)$$

and for the causal contrast between two interventions $F_{h_1}^M$ and $F_{h_2}^M$ defined in (5) as

$$\widehat{\tau}_B(F_{h_1}^M, F_{h_2}^M) = \widehat{N}_B(F_{h_2}^M) - \widehat{N}_B(F_{h_1}^M). \quad (9)$$

For a specific intervention F_h^M , the estimator in (6) can be calculated once and it can be used to estimate the temporally-expected effects defined in Section 3 for any $B \subset \Omega$.

4.4 Asymptotic Properties of the Proposed Estimators

Below, we establish the asymptotic properties of the proposed estimators. Our results differ from most asymptotic normality results in the causal inference literature in two ways. First, they are based on a single observation (namely, one time series of point patterns that are potentially interconnected over both space and time). Second, we employ a kernel-smoothed version of the outcome. We introduce a set of mild regularity conditions.

Assumption 3 (Regularity conditions). *The following three conditions hold.*

- (a) *There exists $\delta_Y > 0$ such that $|S_{Y_t(\bar{\mathbf{w}}_t)}| < \delta_Y$ for all $t \in \mathcal{T}$ and $\bar{\mathbf{w}}_t \in \mathcal{W}^T$.*
- (b) *Let $\bar{H}_t^* = \{\bar{W}_t, \bar{Y}_T, \bar{X}_T\} \supset \bar{H}_t$, and $v_t = \text{Var}[\widehat{N}_{Bt}(F_h^M) | \bar{H}_{t-M}^*]$ for $t \geq M$. Then, there exists $v \in \mathbb{R}^+$ such that $(T - M + 1)^{-1} \sum_{t=M}^T v_t \xrightarrow{p} v$ as $T \rightarrow \infty$.*
- (c) *There exists a neighborhood of set B 's boundary over which outcome active locations are observed during at most T^{1-Q^*} time periods, for some $Q^* \in (1/2, 1)$ and as $T \rightarrow \infty$.*

Assumption 3(a) states that there is an upper limit on the number of outcome-active locations at any time period and under any treatment path. In our application, it is reasonable to assume that the number of insurgent attacks occurring during any day is bounded. In Assumption 3(b), \overline{H}_t^* represents the expanded history preceding W_{t+1} , including previous treatments, all potential outcomes, and all potential confounders. Given the assumptions of bounded relative positivity and bounded number of outcome-active locations, $\widehat{N}_{Bt}(F_h^M)$ is also bounded, so Assumption 3(b) is a weak addition. Lastly, a more rigorous statement of Assumption 3(c) requires additional notation, which is presented in Appendix B. Since the size of the boundary's neighborhood can be arbitrarily small, this assumption is also reasonable. We now present the main theoretical results.

Theorem 1. *If Assumptions 1–3 hold and the bandwidth $b_T \rightarrow 0$, we have,*

$$\sqrt{T}(\widehat{N}_B(F_h^M) - \overline{N}_B(F_h^M)) \xrightarrow{d} \mathcal{N}(0, v),$$

as $T \rightarrow \infty$, where v is the quantity defined in Assumption 3(b).

The key idea is to separate the estimation error arising due to the treatment assignment, $[W_t \mid \overline{H}_{t-1}^*]$, from the error due to spatial smoothing. Using martingale theory similar to Bojinov and Shephard (2019), we show the former to be \sqrt{T} -asymptotically normal. The latter is shown to converge to zero at a rate faster than $1/\sqrt{T}$.

Knowing v would allow us to estimate the temporally-expected potential outcome and make inference based on the estimator's asymptotic distribution, since $\widehat{N}_B(F_h^M) \overset{d}{\approx} \mathcal{N}(\overline{N}_B(F_h^M), v/T)$ for large T . The variance v is the converging point of $(T - M + 1)^{-1} \sum_{t=M}^M v_t$ in Assumption 3(b), i.e., the average over the time t estimator's variance. Because for each time period t we only observe one treatment path, one cannot directly estimate v_t and thus v without additional assumptions. We circumvent this problem by using an upper bound of v , a quantity, which we can consistently estimate. Specifically, let $v_t^* = \mathbb{E} \left\{ [\widehat{N}_{Bt}(F_h^M)]^2 \mid \overline{H}_{t-M}^* \right\}$. Then, it is clear that $v \leq (T - M + 1)^{-1} \sum_{t=M}^T v_t^* \xrightarrow{P} v^*$, and an α -level confidence interval for $\overline{N}_B(F_h^M)$ which uses the asymptotic variance bound

v^*/T will achieve the nominal coverage. Even though v^* cannot be directly calculated either, there exists a consistent estimator of the asymptotic variance's upper bound, as stated in the following lemma:

Lemma 1. *It holds that $\frac{1}{T-M+1} \sum_{t=M}^T \left[\hat{N}_{Bt}(F_h^M)^2 - v_t^* \right] \xrightarrow{p} 0$.*

The results presented above for the estimator $\hat{N}_B(F_h^M)$ can be easily extended to the estimator $\hat{\tau}_B(F_{h_1}^M, F_{h_2}^M)$:

Theorem 2. *If Assumptions 1–3 hold,*

$$\frac{1}{T-M+1} \sum_{t=M}^T \text{Var} \left[\hat{\tau}_{Bt}(F_{h_1}^M, F_{h_2}^M) \mid \overline{H}_{t-M}^* \right] \xrightarrow{p} \eta,$$

for some $\eta > 0$, and the bandwidth $b_T \rightarrow 0$, we have

$$\sqrt{T}(\hat{\tau}_B(F_{h_1}^M, F_{h_2}^M) - \tau_B(F_{h_1}^M, F_{h_2}^M)) \xrightarrow{d} N(0, \eta),$$

as $T \rightarrow \infty$. Furthermore, an upper bound of the asymptotic variance v can be consistently estimated by

$$\frac{1}{T-M+1} \sum_{t=M}^T \left[\hat{\tau}_{Bt}(F_{h_1}^M, F_{h_2}^M) \right]^2 \xrightarrow{p} v^* \geq v.$$

All of the theoretical results presented above have been established for the correct and known propensity score $e_t(w)$. However, in practice, the propensity score is unknown and must be estimated. When the propensity score is estimated from a correctly-specified model, it has been shown in various contexts that the resulting estimator maintains the desirable properties described above while having a lower asymptotic variance than the estimator utilizing the true propensity score (e.g., [Hirano et al., 2003](#)). Therefore, even for the estimated propensity score, we make inference based on the upper bound of the asymptotic variance derived above.

5 Simulation Studies

We conduct simulation studies to evaluate (a) the performance of the proposed estimator as the number of time periods increases, (b) the accuracy of the asymptotic approximation, (c) the difference between the theoretical variance bound and the actual variance, (d) the performance of the inferential approach based on the estimated asymptotic variance bound, and (e) the balancing properties of the estimated propensity score. We use the `spatstat` R package (Baddeley *et al.*, 2015) to generate point patterns from Poisson processes and fitting Poisson process models to the simulated data.

5.1 The Simulation Design

We consider a time series of point patterns of length $T \in \{200, 400, 500\}$ on the unit square, $\Omega = [0, 1] \times [0, 1]$. For each time series length T , 200 data sets are generated. We now describe our simulation design in detail.

Time-varying and time-invariant confounders. Our simulation study includes two time-invariant confounders. For the first time-invariant confounder, we construct a hypothetical road network on Ω using lines and arcs, which is highlighted by bright white lines in Fig-

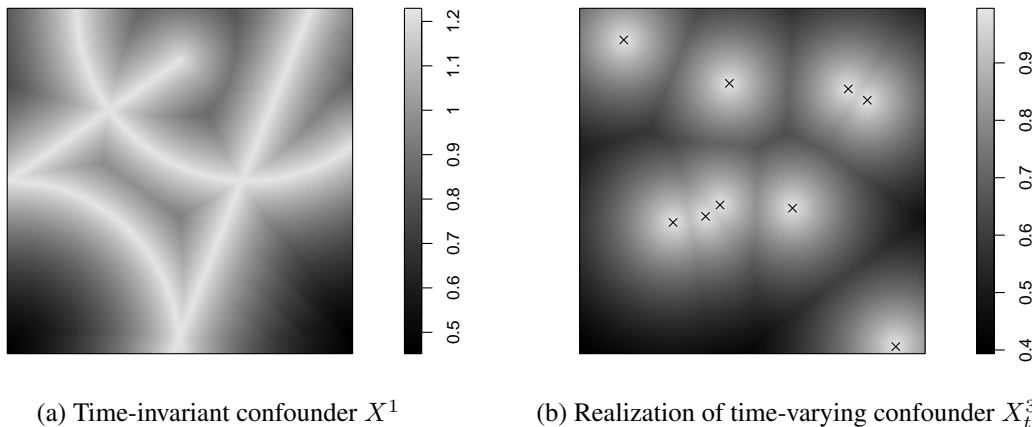


Figure 3: Simulated Confounders. Panel (a) shows one of the two time-invariant confounders representing the exponential decay of distance to the road network. Panel (b) shows one realization for one of the time-varying confounders, which is generated as a function of the exponential decay of distance to the locations indicated by \times .

ure 3a. Then, we define $X^1 = 1.2 \exp(-2D_1)$ where D_1 is the distance to the closest line. The second time-invariant covariate is constructed similarly, as $X^2 = \exp(-3D_2)$ where D_2 is the distance to the closest arc. In addition, we generate two time-varying confounders, X_t^3 and X_t^4 , based on the exponential decay of distance to the closest point; these points are generated according to non-homogeneous Poisson point processes with the following intensity function

$$\lambda_t^{X^j}(\omega) = \exp\{\rho_0^j + \rho_1^j X^1(\omega)\}, \quad j = 3, 4,$$

where $\rho_1^3 = 1$, and $\rho_1^4 = 1.5$. Figure 3b shows one realization of X_t^3 .

Spatio-temporal point processes for treatment and outcome variables. For each time period $t \in \{1, 2, \dots, T\}$, we generate W_t from a non-homogeneous Poisson process that depends on all confounders $\mathbf{X}_t(\omega)^T = (X^1, X^2, X_t^3, X_t^4)(\omega)$, and the previous treatment and outcome realizations, W_{t-1} and Y_{t-1} . The intensity of this process is given by

$$\lambda_t^W(\omega) = \exp\{\alpha_0 + \mathbf{X}_t(\omega)^T \alpha_X + \alpha_W W_{t-1}^*(\omega) + \alpha_Y Y_{t-1}^*(\omega)\}, \quad (10)$$

where $W_{t-1}^*(\omega) = \exp(-2D_W(\omega))$ and $Y_{t-1}^*(\omega) = \exp(-2D_Y(\omega))$ with $D_W(\omega)$ and $D_Y(\omega)$ being the minimum distance of ω to the points in $S_{W_{t-1}}$ and $S_{Y_{t-1}}$, respectively.

Similarly, we generate Y_t from a non-homogeneous Poisson process with intensity

$$\lambda_t^Y(\omega) = \exp\{\gamma_0 + \mathbf{X}_t(\omega)^T \gamma_X + \gamma_2 X_{t-1}^2(\omega) + \gamma_W W_{(t-3):t}^*(\omega) + \gamma_Y Y_{t-1}^*(\omega)\}, \quad (11)$$

where $W_{(t-3):t}^*(\omega) = \exp(-2D_W^*(\omega))$ is the minimum distance of ω to the points in $\bigcup_{j=t-3}^t S_{W_j}$. This specification imposes a lag-four dependence of the outcome on the lagged treatment process. The model gives rise to the average of 5 observed treatment-active locations and 21 observed outcome-active locations within each time period.

Stochastic interventions. We consider interventions of the form F_h^M based on a homogeneous Poisson process with intensity h that is constant over Ω and ranges from 3 to 7. We consider various lengths of each intervention by setting $M \in \{1, 3, 7, 30\}$. The second intervention we consider is defined over the three time periods, i.e., $F_h = F_{h_3} \times F_{h_2} \times F_{h_1}$ with $M = 3$. The intervention for the first time period F_{h_3} is a homogeneous Poisson process with intensity h_3 ranging from 3 to 7, whereas $F_{h_2} = F_{h_1}$ is a homogeneous Poisson process with intensity equal to 5. For each stochastic intervention, we consider the region of interest, denoted by set B , of three different sizes: $B = \Omega = [0, 1] \times [0, 1]$, $B = [0, 0.5] \times [0, 0.5]$, and $B = [0.75, 1] \times [0.75, 1]$.

Approximating the true values of the estimands. From (11), it is clear that the potential outcomes depend on the realized treatments during the last four time points as well as the realized outcomes from the previous time period. This implies that the estimands for all interventions, even for $M > 4$, depend on the observed treatment and outcome paths and are therefore not constant across simulated data sets. Therefore, we approximate the true values of the estimands in the following manner. For each time period t , and each $r = 1, \dots, R$ repetition, we generate realizations $w_{t-M+1}^{(r)}, \dots, w_{t-1}^{(r)}, w_t^{(r)}$ from the intervention distribution F_h . Based on the treatment path $(\bar{w}_{t-M}^{obs}, w_{t-M+1}^{(r)}, \dots, w_t^{(r)})$, we generate outcomes $Y_{t-M+1}^{(r)}, \dots, Y_t^{(r)}$ using (11). This yields $S_{Y_t^{(r)}}$, which contains the outcome-active locations based on one realization from the stochastic intervention. Repeating this process R times and calculating the average number of points that lie within B provides a Monte Carlo approximation of $\bar{N}_{Bt}(F_h)$, and further averaging these over time gives an approximation of $\bar{N}_B(F_h)$.

Estimation. We estimate the expected number of points $\bar{N}_B(F_h)$ and the effect of a change in the intervention on this quantity $\tau_B(F_{h'}, F_{h''})$ using the following estimators: (a) the proposed estimators defined in (8) and (9) with the true propensity scores; (b) the same proposed estimators but with the estimated propensity scores based on the correctly-speci-

fied model; (c) the above two estimators but with the Hájek-type standardization where the weight of time t is divided by the mean of weights across all time periods (see Appendix C); and (d) the unadjusted estimator based on a propensity score model which is specified as a homogeneous Poisson process without confounders.

All estimators utilize the smoothed outcome point pattern. Spatial smoothing is performed using Gaussian kernels with standard deviation equal to $10T^{-2/3}$, which is decreasing in T . We choose this bandwidth such that for $T = 500$ (the longest time series in our simulation scenario) the bandwidth is approximately equal to 0.16, smaller than the size of the smallest B (which is equal to $[0.75, 1]^2$). A discussion on choosing the bandwidth is given in Section 6.3.

Theoretical variance and its upper bound. Theorems 1 and 2 provide the expressions for the asymptotic variances of the proposed estimators. We compute Monte Carlo approximations to these theoretical variances and their upper bounds. Specifically, for each time period t and each replication r , the computation proceeds as follows: 1) We generate treatment and outcome paths $w_{t-M+1}^{(r)}, y_{t-M+1}^{(r)}, \dots, w_t^{(r)}, y_t^{(r)}$ using the distributions specified in (10) and (11), 2) Using the data $(w_{t-M+1}^{(r)}, \dots, w_t^{(r)})$ and the outcome $y_t^{(r)}$, we compute the estimator according to (6) and (7), and finally 3) we calculate the variance and the second moment of these estimates over R replications, which can be used to compute the asymptotic variance and variance bound of interest. The averages of these estimates over time gives the desired Monte Carlo approximations. We use a similar procedure to approximate the theoretical variance and variance bound of $\hat{\tau}_B(F_{\mathbf{h}'}, F_{\mathbf{h}''})$.

Estimating the variance bound and the resulting inference. We use Lemma 1 to estimate the variance bound. This estimated variance bound is then used to compute the confidence intervals and conduct a statistical test of whether the causal effect is zero. Inference based on the Hájek-type adjustment is discussed in Appendix C.

Balancing property of the propensity score. Using the correctly specified model, we estimate the propensity score at each time period t . The inverse of the estimated propensity score is then used as the weight in the weighted Poisson process model for W_t with the intensity specified in (10). We compare the statistical significance of the predictors between the weighted and unweighted model fits. Large p -values under the weighted model would suggest that the propensity score adequately balances the confounding variables.

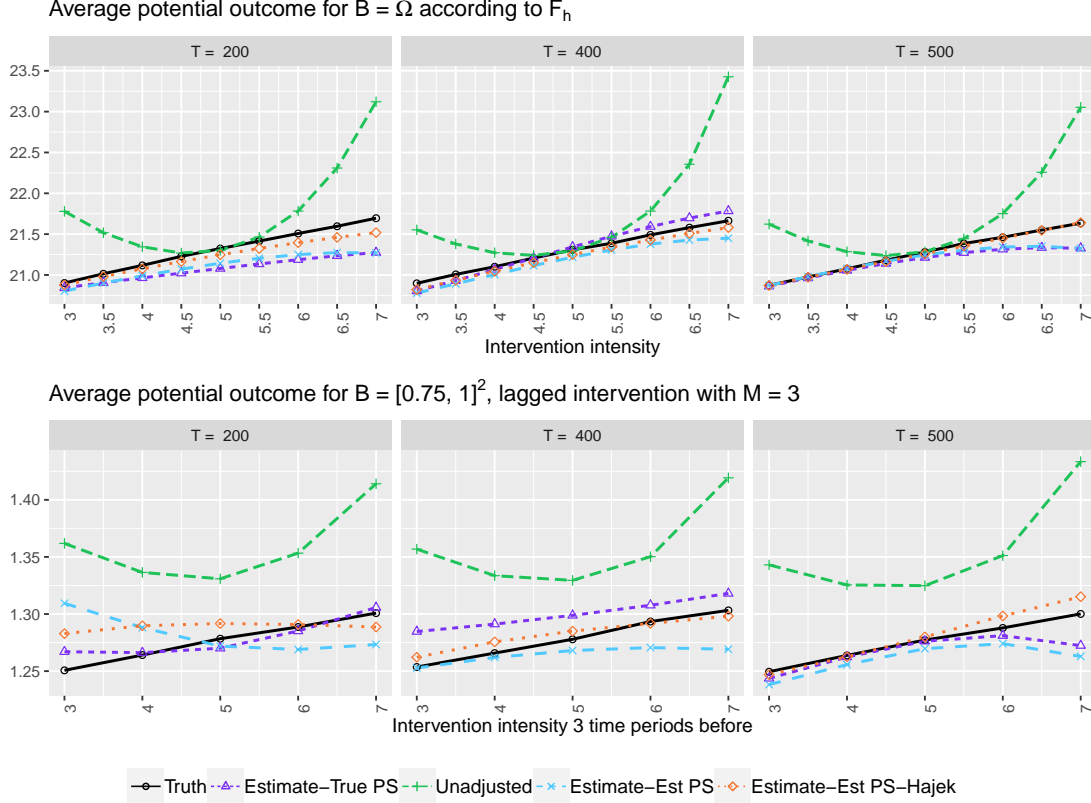


Figure 4: Simulation Results for the True and Estimated Average Potential Outcomes. In the top panel, we present the true and estimated average potential outcomes in the entire region $B = \Omega$ under single-time interventions with the varying intensity (horizontal axis). In the bottom panel, we consider the average potential outcome in the sub-region $B = [0.75, 1]^2$ for the intervention F_h , with $M = 3$, the varying intensity of F_{h_3} (horizontal axis), and F_{h_1}, F_{h_2} intensity set to 4. The black lines with solid circles represent the truths, while the other dotted or dashed lines represent the estimates; the estimator based on the true propensity score (purple triangles), the unadjusted estimator (green crosses), the estimator based on the estimated propensity score (blue x's), the Hájek estimator based on the estimated propensity score (orange rhombuses).

5.2 Simulation Results

Figure 4 presents the results for the two simulation scenarios. The top panel shows how the (true and estimated) average potential outcomes in the whole region ($B = \Omega$) change as the intensity varies under the single time period interventions. The bottom panel shows how the true and estimated average potential outcomes in the sub-region $[0.75, 1]^2$ change under the three time period interventions when the intensity at three time periods ago ranges from 3 to 7. For both simulation scenarios, we vary the length of the time series from 200 (left plots) to 500 (right plots).

As expected, the unadjusted estimates (green crosses) are far away from the true average potential outcome (black solid circles) across all simulation scenarios. In contrast, and consistent with the results of Theorems 1 and 2, the accuracy of the proposed estimator (purple triangles based on the true propensity score, blue x's based on the estimated propensity score) improves as the number of time periods increases. We note that the convergence is slower when $M = 3$ than $M = 1$. This is expected because the uncertainty of the treatment assignment is greater for a stochastic intervention with a longer time period.

We find that the Hájek estimator performs well across all simulation scenarios even when T is small and M is large, whereas the IPW estimator tends to suffer from extreme weights because the weights are multiplied over the intervention time periods as shown in (6). Appendix D.1 presents additional simulation results for the IPW and Hájek estimators under the interventions F_h^M with $M = 3, 7, 30$, and the intensity h varying from 3 to 7. These results indicate a deteriorating performance of the IPW estimator as the value of M increases. In contrast, the standardization of weights used in the Hájek estimator appears to partially alleviate this issue, with its effect estimates much closer to their true values.

Next, we compare the true theoretical variance, v/T , with the variance bound v^*/T and its consistent estimate (see Lemma 1). Since part of the goal here is to assess the conservativeness of the theoretical variance bound to the theoretical variance, we focus on

the proposed estimators with the true propensity score. Figure 5 shows the results of an intervention F_h^M for $M = 3$, and for region $B = [0, 0.5]^2$. First, we focus on the theoretical variance and variance bound (blue line with open circles, and orange dotted lines with open triangles, respectively). As expected, the true variance decreases as the total number of time periods increases. Moreover, the proposed theoretical variance bound tightly follows the theoretical variance, reflected by the fact that the two lines are visually indistinguishable. These results indicate that the theoretical variance bound is not overly conservative compared to the true theoretical variance. Indeed, the variance bound is substantially larger than the true variance only in the low-variance scenarios (the interventions over a single time period and resembling the actual data generating mechanism, as shown in Appendix D.2).

Second, we compare the theoretical variance bound with the estimated variance bound (green dashed lines with open rhombuses). As the length of time series increases, the estimated variance bound more closely approximates its theoretical value (consistent with Lemma 1). Furthermore, the estimated variance bound is close to its theoretical value under low uncertainty scenarios and when the intervention intensity closely resembles that of the actual data generating process. However, we find that the estimated variance bound underestimates the true variance bound in high uncertainty scenarios, and convergence to its true value is slower for larger values of M (see Appendix D.2).

The results on the asymptotic variance and variance bound in Figure 5 lead to similar conclusions with respect to the coverage of 95% confidence intervals of the IPW estimator. In Appendix D.3, we find that except when $M = 30$, the confidence interval for the IPW estimator based on either the true asymptotic variance or the true variance bound has a coverage of about 80% or higher. This implies that the asymptotic normality established in Theorem 1 provides an adequate approximation to the estimator's sampling distribution for small or moderate values of M . However, for $M = 30$, the confidence interval for the IPW estimator is anti-conservative due to the fact that the weights, which equal the product of

Standard deviation and bound of the average potential outcome estimator for $B = [0, 0.5]^2$

Intervention over 3 time points

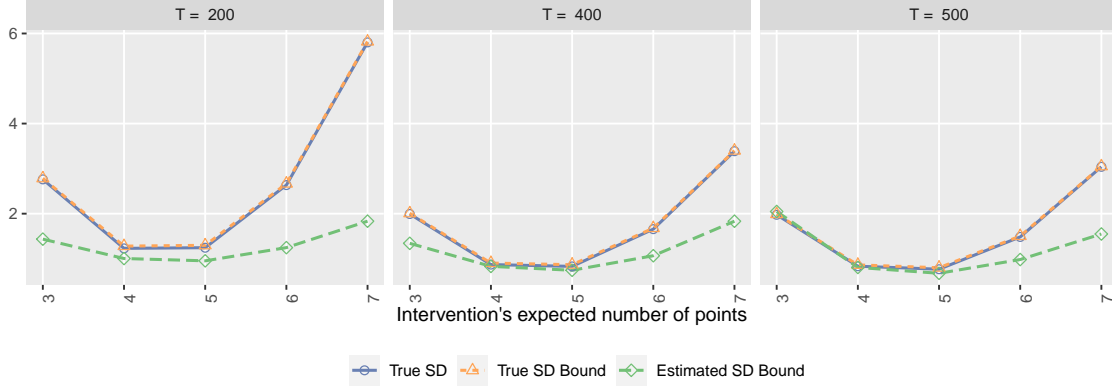


Figure 5: Comparison of the Asymptotic Standard Deviation with the True and Estimated Asymptotic Standard Deviation Bound for the Average Potential Outcome Estimator. The comparison is based on the varying number of expected points (horizontal axis) under a homogeneous Poisson process intervention taking place over $M = 3$ time periods. Each plot corresponds to a simulation setting with a different time series length.

ratios across many time periods, become extremely small. In addition, the underestimation of the variance bound in high uncertainty scenarios found in Figure 5 leads to the under-coverage of the confidence intervals based on the IPW estimator and using the estimated variance bound, especially when the interventions take place over long time periods.

In our simulations, we find that the confidence interval for the Hájek estimator discussed in Appendix C has a better coverage probability even for the interventions over long time periods (see Appendix D.3). Appendix D.4 also shows that the estimated standard deviation for the Hájek estimator outperforms that for the IPW estimator under many simulation scenarios. Partly based on these findings, we use the Hájek estimator and its associated confidence interval in our application (see Section 6).

Finally, we evaluate the performance of the propensity score as a balancing score, as established in Proposition 1. In Appendix D.5, we show that the p-values of the previous outcome-active locations variable (Y_{t-1}^* in (10)) are substantially greater in the weighted propensity score model than in the unweighted model, where the weights are given by the inverse of the estimated propensity score. We also find the estimated coefficients in the weighted model are centered around zero. These findings are consistent with the balancing

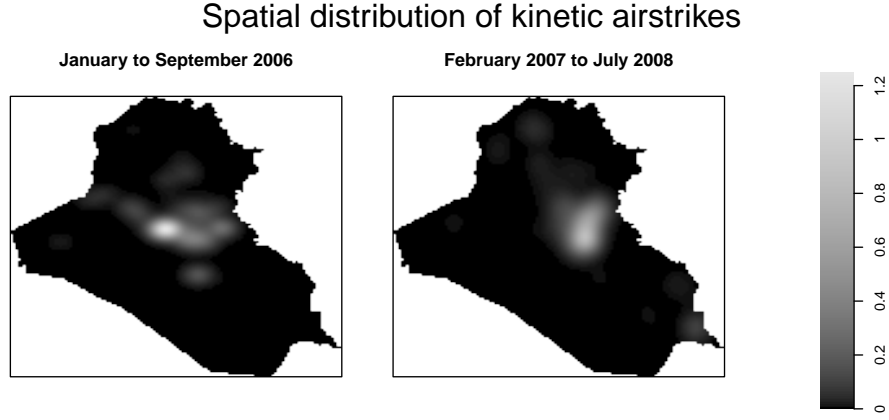


Figure 6: Spatial Density Estimate of Airstrike Locations during January 1 – September 24, 2006 (left) and the Entire Study Period February 2007 – July 2008 (right).

property of the propensity score.

6 Empirical Analyses

In this section, we present our empirical analyses of the data sets introduced in Section 2.

We first describe the airstrike strategies of interest and then discuss the causal effect estimates obtained under those strategies.

6.1 Airstrike Strategies and Causal Effects of Interest

For the purpose of this application, we consider hypothesized stochastic interventions that generate airstrike locations based on a simple non-homogeneous Poisson point process with finite and non-atomic intensity $h : \Omega \rightarrow [0, \infty)$. We first specify a baseline probability density ϕ_0 over Ω . To make this baseline density realistic and increase the credibility of the overlap assumption, we use the airstrike data during January 1 – September 24, 2006 to estimate the baseline distribution ϕ_0 for our stochastic interventions. This subset of the data is not used in the subsequent analysis. The left plot of Figure 6 shows the estimated baseline density using kernel-smoothing of airstrikes, which is more diffused than the estimated spatial distribution of the entire study period (shown in the right plot of the figure).

We consider the following three questions: (1) How does an increase in the number of airstrikes affect insurgent violence? (2) How does the shift in the prioritization of certain

locations for airstrikes change the spatial pattern of insurgent attacks? (3) How long does it take for the effects of change in these airstrike strategies to be realized? The last question examines how quickly the insurgents respond to the change in airstrike strategy.

We address the first question by considering stochastic interventions that have the same spatial distribution but vary in the expected number of airstrikes. We represent such strategies using intensities $h(\omega) = c\phi_0(\omega)$ with different values of $c > 0$. Since $\int_{\Omega} h(\omega)d\omega$ represents the expected number of points from a Poisson point process, these interventions have the same spatial distribution ϕ_0 , but the number of airstrikes is monotonically increasing as a function of c . In our analysis, we consider $\{1, 2, \dots, 6\}$ as the range of c .

For the second question, we fix the intensity but vary the focal locations of airstrikes. To do this, we specify a distribution over Ω with power-density $d_{\alpha}(\omega) = d(\omega)^{\alpha} / (\int_{\Omega} d(\omega)^{\alpha})$ and modes located at $s_f \in \Omega$. Based on d_{α} , we specify $h_{\alpha}(\omega) = c_{\alpha}\phi_0(\omega)d_{\alpha}(\omega)$ where c_{α} satisfies the constraint $\int_{\Omega} h_{\alpha}(\omega)d\omega = c$, so that the overall expected number of airstrikes remains constant. Locations in s_f are prioritized under h_{α} as an increasing function of α . For our analysis, we choose the center of Baghdad to be the focal point s_f and d_{α} to be the normal distribution centered at s_f with precision α . We set the expected number of airstrikes per day c to be 3, and vary the precision parameter α from 0 to 3. The visualization of the spatial distributions in h_{α} for the different values of α is shown in Figure A.8.

As discussed in Section 3, for both of these questions, we can specify airstrike strategies of interest taking place over a number of time periods, M , by specifying the stochastic interventions as $F_h = F_h^M$. We assume that insurgent attacks at day t do not affect airstrikes on the same day, and airstrikes at day t can only affect attacks during subsequent time periods. Thus, causal quantities for interventions taking place over M time periods refer to insurgent attacks occurring on the following day.

In addition, we may also be interested in the lagged effects of airstrike strategies as illustrated by the third question mentioned above. We specify lagged intervention to be the

one which differs only for the M time periods ago, i.e., $F_h = F_{h_0}^{M-1} \times F_{h_1}$, where $h_0 = \phi_0$ represents the baseline intensity (with $c = 1$), and $h_1 = c\phi_0$ is the increased intensity with different values of c ranging from 1 to 6. For our analysis, we consider M to be equal 1 day, 3 days, 1 week, and 1 month.

Finally, in Appendix E.3, we also introduce local interventions, which change the airstrike strategy in one region while leaving that for the rest of the country unchanged.

6.2 The Specification and Diagnostics of the Propensity Score Model

Our propensity score model is a non-homogeneous Poisson point process with intensity $\lambda_t(\omega), \omega \in \Omega$, where $\log \lambda_t(\omega) = \beta^T \mathbf{X}_t(\omega)$, and \mathbf{X} includes an intercept and all the covariates. The two main drivers of military decisions over airstrikes are the prior number and locations of observed insurgent attacks and airstrikes, conditional on which unconfoundedness in Assumption 1 is expected to approximately hold. Therefore, our model includes the observed airstrikes and attacks during the last day, week, and month. For example, the airstrike history of time t during the previous week is specified as $\overline{W}_{t-1}^*(\omega) = \sum_{j=1}^7 \sum_{s \in S_{w_{t-j}^{obs}}} \exp\{-\text{dist}(s, \omega)\}$, which represents a continuous surface on Ω with locations closer to the airstrikes in the previous week having greater positive values than more distant locations.

Our propensity score model also includes additional important covariates that might affect both airstrikes and insurgent attacks. For example, we adjust for shows-of-force (i.e., simulated bombing raids designed to deter insurgents) that occurred one day, one week, and one month before each airstrike. Patterns of US aid spending might also affect the location and intensity of insurgent violence and airstrikes, as we discussed in Section 2. We therefore include the amount of aid spent (in \$US dollars) in each Iraqi district in the past month as a time-varying covariate. Finally, we also incorporate temporal splines and several time-invariant spatial covariates, including the airstrike's distance from major cities, road networks, rivers, local settlements, and the population (logged, measured in 2003) of

the governorate in which the airstrike took place.

We evaluate the covariate balance by comparing the p -values of estimated coefficients in the propensity score model to the p -values in the weighted version of the same model, where each time period is inversely weighted by its truncated propensity score estimate (truncated above at the 90th quantile). Although 13 out of 35 estimated coefficients had p -values smaller than 0.05 in the fitted propensity score model, all the p -values in the weighted propensity score model are close to 1, suggesting that the estimated propensity score adequately balances these confounders.

6.3 The Choice of the Bandwidth Parameter for the Spatial Kernel Smoother

The kernel smoothing part of our estimator is not necessary for estimating the number of points within any set $B \subset \Omega$ since we can simply use an IPW estimator based on the observed number of points within B . However, kernel smoothing is useful for visualizing the estimated intensities of insurgent attacks under an intervention of interest over the entire country. One can also use it to acquire estimates of the expected number of insurgent attacks under the intervention for any region of Iraq by considering the intensity's integral over the region. Theorem 1 shows that, for any set $B \subset \Omega$, kernel smoothing does not affect the estimator's asymptotic normality as long as the bandwidth converges to zero. In practice, the choice of the bandwidth should be partly driven by the size of the sets B .

In our analysis, we estimate the causal quantities for the entire country and the Baghdad administrative unit (see Figure A.9). We choose an adaptive bandwidth separately for each outcome using the `spatstat` package in R. We consider all observed outcome event locations during our study period, and use Scott's criterion for choosing an optimal, constant bandwidth parameter for an isotropic kernel estimation (Scott, 1992). Using the estimated density as the pilot density, we calculate the optimal adaptive bandwidth surface according to Abramson's inverse-square-root rule (Abramson, 1982). This procedure yields a value of the bandwidth used for kernel smoothing at each outcome event location.

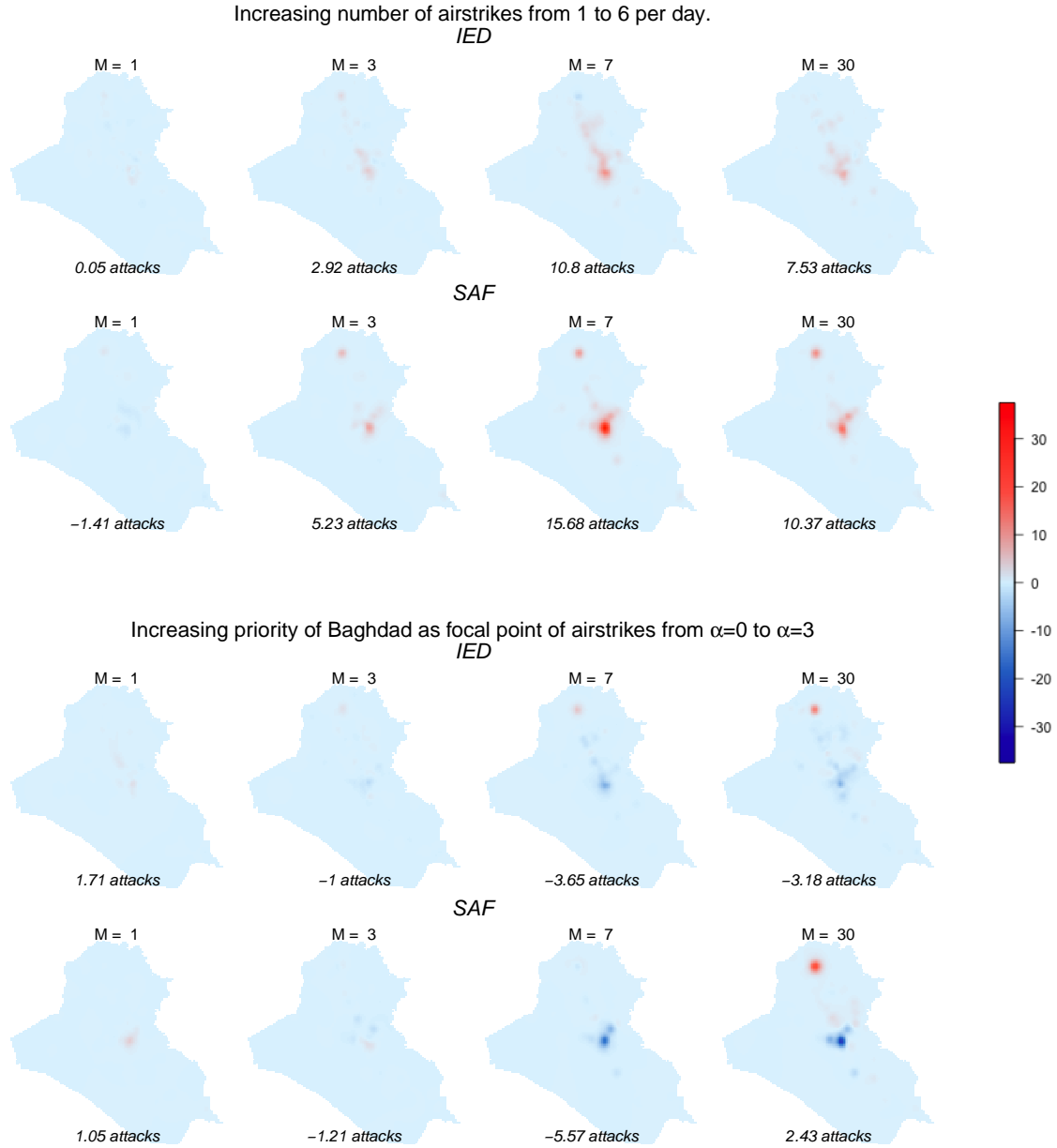


Figure 7: Changes in Estimated Intensity of Insurgent Attacks when Increasing the Intensity of Airstrikes (the first two rows) and when Shifting the Focal Point of Airstrikes to Baghdad (the bottom two rows). Insurgent attacks are measured using Improvised Explosive Devices (IEDs; the first and third rows) and small arms fire (SAFs; second and fourth rows) with the varying number of intervention duration, $M = 1, 3, 7, 30$ days (columns). The number shown below each map represents the estimated change in the total number of attacks per day over the entire country, whereas the legend represents the difference in estimated intensities.

6.4 Findings

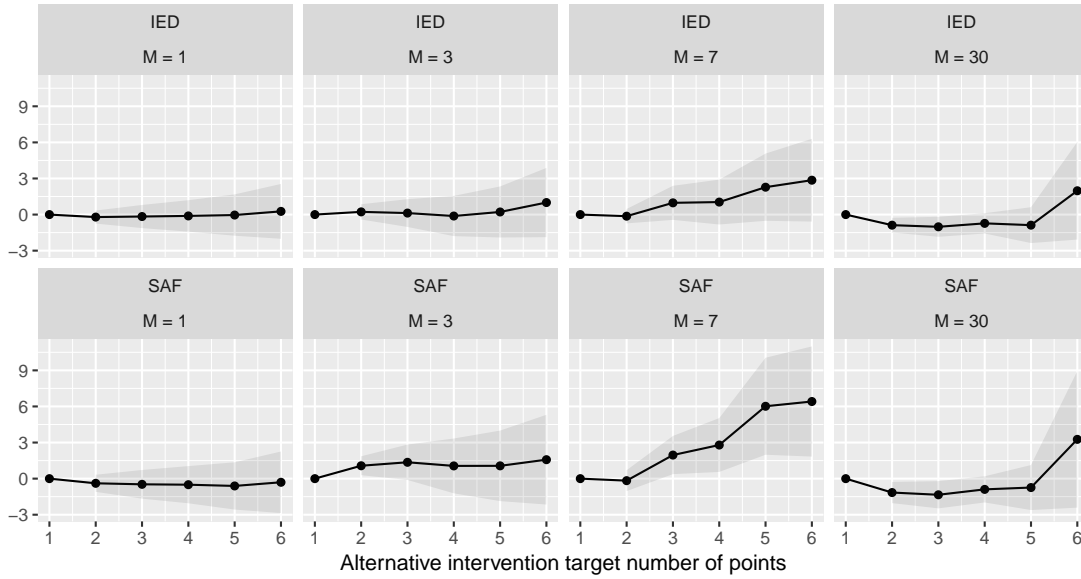
Figure 7 illustrates changes in the estimated intensity surfaces for insurgent attacks (measured using IEDs and SAFs) when increasing the intensity of airstrikes (the first two rows)

and when shifting the focal point of airstrikes to Baghdad (the bottom two rows), with the varying duration of interventions, $M = 1, 3, 7, 30$ days (columns). These surfaces can be used to estimate the causal effect of a change in the intervention over any region. Dark blue areas represent areas where the change in the military strategy would reduce insurgent attacks, whereas red areas correspond to those with an increase in insurgent attacks.

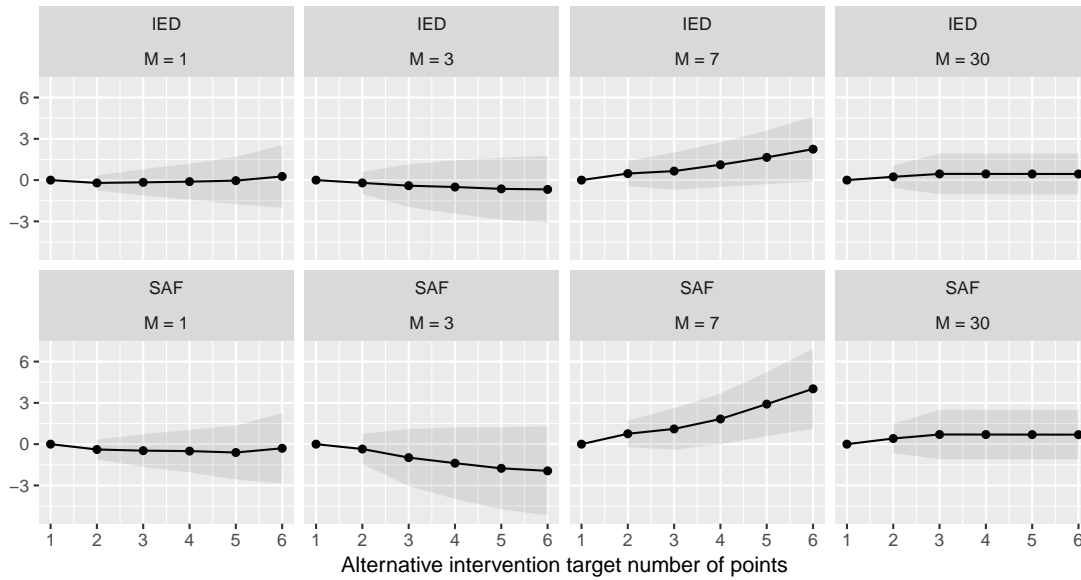
The figure reveals a number of interesting findings. First, there is no substantial change in insurgent attacks if these interventions last only for one day. When increasing airstrikes for a longer duration, however, a greater number of insurgent attacks are expected to occur, with the largest number of attacks per day occurring with the 7 day change in the intervention. These changes are concentrated in the Baghdad area and the roads that connect Baghdad and the northern city of Mosul. These estimated effects, measured in terms of attacks per day, appear to decline when the intervention change lasts for 30 days. These patterns apply to both IEDs and SAFs with slightly greater effects estimated for SAFs.

In contrast, when shifting the focal point of airstrikes to Baghdad for 7 days, the total number of insurgent attacks over the entire country decreases. However, when the intervention lasts for 30 days, we find that insurgents shift their attacks to the areas around Mosul while reducing the number of attacks in Baghdad. This displacement pattern is particularly pronounced for SAFs. For IEDs, insurgents appear to move their attacks to the Mosul area even with the intervention of 7 days and yet the effect size is smaller. Statistical significance of these results is shown in Appendix [E.2](#).

Figure [8a](#) shows the changes in the estimated average number of insurgent attacks *in Baghdad* as the airstrike intensity increases from 1 to 2, 3, . . . , 6 airstrikes per day in the entire country (horizontal axis). We also vary the duration of intervention from $M = 1$ day to $M = 30$ days (columns). Both the point estimate (solid lines) and 95% CIs (grey bands) are shown. Consistent with Figure [7](#), we find that increasing the number of airstrikes leads to a greater number of attacks when the duration of intervention is 7 days. These effects



(a) Estimated effect of increasing airstrike intensity for M days



(b) Estimated effect of increasing airstrike intensity M days ago

Figure 8: Changes in the Estimated Number of Insurgency Attacks in Baghdad when Increasing the Intensity of Airstrikes (a) for M days, and (b) M days ago. In each plot, the horizontal axis represents the expected number of airstrike per day under the alternative intervention. The vertical axis represents a change in the estimated average number of insurgency attacks in Baghdad for IEDs (first & third row) and SAFs (second & fourth row) when number of airstrikes per day increases from 1 to the value on the horizontal axis. Each column shows different (a) duration or (b) lag length of intervention, $M = 1, 3, 7, 30$ days. 95% confidence intervals are shown as grey bands.

appear to decrease when the intervention is much shorter or lasts for a month. The patterns are similar for both IEDs and SAFs.

Figure 8b shows the change in the estimated number of IEDs and SAFs attacks in Bagh-

dad when increasing the number of airstrikes M days before, while the expected number of airstrikes during the following $M - 1$ days equals one per day. We find that for $M = 3$ all estimated lagged effects are negative, whereas the lagged effects for $M = 7$ are estimated to be positive. This suggests that increasing the number of airstrikes may reduce insurgent violence in a short term while leading to an increase in a longer term. Appendix E.2 presents the effect estimates and 95% CIs for various interventions and outcomes.

7 Concluding Remarks

In this paper, we provide a framework for causal inference with spatio-temporal point process treatments and outcomes. We demonstrate the flexibility of this proposed methodology by applying it to the estimation of airstrike effects on insurgent violence in Iraq. Our central idea is to use a stochastic intervention, one that represents a distribution of treatments, instead of the standard causal inference approach that estimates the average potential outcomes under some fixed treatment values. A key advantage of our approach is its flexibility: it permits unstructured patterns of both spatial spillover and temporal carryover effects. This flexibility is crucial since for many spatio-temporal causal inference problems, including our own application, we know little about how the treatments in one region affect the outcomes in other regions across different time periods.

The proposed methodology combines spatio-temporal point process modeling with the inverse probability of treatment assignment weighting approach, thereby avoiding the direct modeling of the outcome. Instead, we use kernel smoothing to estimate the spatial surface of causal quantities, enabling effective visualization of estimated causal effects using maps. Our approach enables the specification of various military strategies across space and time. We illustrate this advantage by varying the intensity and focal locations of airstrikes and their effects over multiple temporal windows. In addition, we also evaluate how long it takes for these effects to manifest, offering new insights into how quickly insurgents respond to airstrikes and how far their violence is displaced.

The proposed framework can also be applied to other high-dimensional, and possibly unstructured, treatments. The standard approach to causal inference, which bases causal effects on fixed treatment values, cannot perform well in such settings. Indeed, the sparsity of observed treatment patterns alone makes it difficult to satisfy the required overlap assumption (Imai and Jiang, 2019). In short, the stochastic intervention approach proposed here offers an effective solution to a broad class of causal inference problems. Future research should further develop the methodology for stochastic intervention. In particular, it is important to develop a sensitivity analysis for the potential violation of the unconfoundedness assumption, an improved weighting method for balancing covariates, and an inferential approach with better finite sample performance.

References

- Abramson, I. S. (1982). On bandwidth variation in kernel estimates—a square root law. *The annals of Statistics* 1217–1223.
- Aronow, P. M., Samii, C., and Wang, Y. (2019). Design-based inference for spatial experiments with interference. *Annual Summer Meeting of the Society for Political Methodology* <https://polmeth.mit.edu/sites/default/files/documents/AronowSamiiWang.pdf>.
- Baddeley, A., Rubak, E., and Turner, R. (2015). *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC Press, London.
- Basse, G. and Airolidi, E. M. (2018). Limitations of design-based causal inference and a/b testing under arbitrary and network interference. *Sociological Methodology* **48**, 1, 136–151.
- Bojinov, I. and Shephard, N. (2019). Time series experiments and causal estimands: Exact randomization tests and trading. *Journal of the American Statistical Association* Forthcoming.

- Csörgö, M. (1968). On the Strong Law of Large Numbers and the Central Limit Theorem for Martingales. *Transactions of the American Mathematical Society* **131**, 1, 259–275.
- Dell, M. and Querubin, P. (2018). Nation building through foreign intervention: Evidence from discontinuities in military strategies. *Quarterly Journal of Economics* **133**, 2, 701–764.
- Díaz Muñoz, I. and van der Laan, M. (2012). Population Intervention Causal Effects Based on Stochastic Interventions. *Biometrics* **68**, 541–549.
- Fisher, R. A. (1935). *The Design of Experiments*. Oliver and Boyd, London.
- Gill, R. D. and Robins, J. M. (2001). Causal inference for longitudinal data: The continuous case. *Annals of Statistics* **29**, 6, 1785–1811.
- Hirano, K. K., Imbens, G. W., and Ridder, G. (2003). Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica* **71**, 4, 1161–1189.
- Imai, K. and Jiang, Z. (2019). Comment on “The Blessings of Multiple Causes” by Wang and Blei. *Journal of the American Statistical Association* **114**, 528, 1605–1610.
- Kennedy, E. H. (2019). Nonparametric causal effects based on incremental propensity score interventions. *Journal of the American Statistical Association* **114**, 526, 645–656.
- Kocher, M., Pepinsky, T., and Kalyvas, S. (2011). Aerial bombing and counterinsurgency in the vietnam war. *American Journal of Political Science* **55**, 2, 201–218.
- Lok, J. J. (2016). Defining and estimating causal direct and indirect effects when setting the mediator to specific values is not feasible. *Statistics in Medicine* **35**, 2, 4008–4020.
- Luo, X., Small, D. S., Li, C.-S. R., and Rosenbaum, P. R. (2012). Inference with interference between units in an fmri experiment of motor inhibition. *Journal of the American Statistical Association* **107**, 498, 530–541.

- Lyall, J. (2019a). Bombing to lose? airpower, civilian casualties, and the dynamics of violence in counterinsurgency wars. *Unpublished Paper* .
- Lyall, J. (2019b). Civilian casualties, humanitarian aid, and insurgent violence in civil wars. *International Organization* **73**, 4, 901–926.
- Mir, A. and Moore, D. (2019). Drones, surveillance, and violence: Theory and evidence from a us drone program. *International Studies Quarterly* **63**, 4, 846–862.
- Neyman, J. (1923). On the application of probability theory to agricultural experiments: Essay on principles, section 9. (translated in 1990). *Statistical Science* **5**, 465–480.
- Robins, J. M. (1997). *Latent Variable Modeling and Applications to Causality*, vol. 120 of *Lecture Notes in Statistics*, chap. Causal Inference from Complex Longitudinal Data, 69–117. Springer Verlag, New York.
- Robins, J. M. (1999). Association, causation, and marginal structural models. *Synthese* 151–179.
- Robins, J. M., Hernán, M. A., and Brumback, B. (2000). Marginal structural models and causal inference in epidemiology. *Epidemiology* **11**, 5, 550–560.
- 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 non-randomized studies. *Journal of Educational Psychology* **66**, 688–701.
- Sävje, F., Aronow, P. M., and Hudgens, M. G. (2019). Average treatment effects in the presence of unknown interference. arXiv:1711.06399.
- Scott, D. W. (1992). *Multivariate Density Estimation*. New York : Wiley.

- Sobel, M. E. and Lindquist, M. A. (2014). Causal inference for fmri time series data with systematic errors of measurement in a balanced on/off study of social evaluative threat. *Journal of the American Statistical Association* **109**, 507, 967–976.
- Tchetgen Tchetgen, E. J., Fulcher, I., and Shpitser, I. (2017). Auto-g-computation of causal effects on a network. *arXiv preprint arXiv:1709.01577* .
- van der Vaart, A. W. (2010). TIME SERIES. *VU University Amsterdam, lecture notes* .
- Zhang, M., Joffe, M. M., and Small, D. S. (2011). Causal inference for continuous-time processes when covariates are observed only at discrete times. *Annals of Statistics* **39**, 1, 131–173.

SUPPLEMENTARY APPENDIX FOR “CAUSAL INFERENCE WITH SPATIO-TEMPORAL DATA”

A Notation

Table A.1: Notation.

Paths	\overline{w}_t	Realization of the treatment assignments for time periods $1, \dots, t$
	$\overline{\mathcal{Y}}_t$	Collection of all potential outcomes for time periods $1, \dots, t$
	\overline{Y}_t^{obs}	Observed outcomes for time periods $1, \dots, t$
Intervention	M	The number of time periods over which we intervene
	h	The Poisson point process intensity defining the stochastic intervention
Estimands	$\overline{N}_t, \overline{N}$	Expected number of outcome-active locations for time period t for an intervention over M time periods, and their average over time
	τ_t^M, τ^M	Expected change in the number of outcome-active locations comparing two interventions for time period t and their average over time
Estimators	\hat{Y}_t^M	Estimated continuous surface the integral of which is used for calculating \hat{N}_t
	\hat{N}_t, \hat{N}	Estimated expected number of points during time period t for an intervention taking place over the preceding M time periods, and their average over time
	$\hat{\tau}_t^M, \hat{\tau}^M$	Estimated expected change in the number of outcome-active locations for time period t comparing two interventions, and their average over time
Arguments	B	The set over which the number of outcome-active locations are counted

B Theoretical Proofs

Proof of Theorem 1. Note that the collection of variables temporally precedent to treatment at time period t is the expanded history \overline{H}_{t-1}^* , defined in Assumption 3. The expanded history \overline{H}_{t-1}^* is a filtration generated by the collection of potential confounders $\overline{\mathcal{X}}_T$, the col-

lection of potential outcomes $\overline{\mathcal{Y}}_T$, and the previous treatments, and satisfies $\overline{H}_{t-1}^* \subset \overline{H}_t^*$.

Let $e_t = \widehat{N}_{Bt}(F_h) - \overline{N}_{Bt}(F_h)$ be the estimation error for time period t and lag M . We will decompose e_t in two components, one corresponding to error due to the treatment assignment (A_{1t}), and one corresponding to error due to spatial smoothing (A_{2t}). Since the bandwidth parameter of the kernel depends on T , we write K_{b_T} instead of K_b . Specifically,

$$\begin{aligned} e_t &= \left[\prod_{j=t-M+1}^t \frac{f_h(W_j)}{e_j(W_j)} \right] \int_B \sum_{s \in S_{Y_t}} K_{b_T}(\omega, s) d\omega - \overline{N}_{Bt}(F_h^M) \\ &= \underbrace{\left[\prod_{j=t-M+1}^t \frac{f_h(W_j)}{e_j(W_j)} \right] N_B(Y_t(\overline{\mathbf{w}}_{t-M}, W_{t-M+1}, \dots, W_t)) - \overline{N}_{Bt}(F_h^M)}_{A_{1t}} + \\ &\quad \underbrace{\left[\prod_{j=t-M+1}^t \frac{f_h(W_j)}{e_j(W_j)} \right] \left[\int_B \sum_{s \in S_{Y_t}} K_{b_T}(\omega, s) d\omega - N_B(Y_t(\overline{\mathbf{w}}_{t-M}, W_{t-M+1}, \dots, W_t)) \right]}_{A_{2t}} \end{aligned}$$

We show that

1. $\sqrt{T} \left(\frac{1}{T-M+1} \sum_{t=M}^T A_{1t} \right)$ is asymptotically normal, and
2. $\sqrt{T} \left(\frac{1}{T-M+1} \sum_{t=M}^T A_{2t} \right)$ converges to zero in probability.

Showing asymptotic normality of the first error.

We use the central limit theorem for martingale difference series (Theorem 4.16 of [van der Vaart \(2010\)](#)) to show asymptotic normality of $(T - M + 1)^{-1} \sum_{t=M}^T A_{1t}$.

Claim. A_{1t} is a martingale difference series with respect to the filtration $\mathcal{F}_t = \overline{H}_{t-M+1}^*$.

To prove this, we show that $E(|A_{1t}|) < \infty$ and $E(A_{1t} \mid \mathcal{F}_{t-1}) = E(A_{1t} \mid \overline{H}_{t-M}^*) = 0$.

For the first part, Assumptions 2 and Assumption 3(a) imply that A_{1t} is bounded and hence $E[|A_{1t}|] < \infty$:

$$\begin{aligned} |A_{1t}| &\leq \left| \prod_{j=t-M+1}^t \frac{f_h(W_j)}{e_j(W_j)} N_B(Y_t(\overline{\mathbf{w}}_{t-M}, W_{t-M+1}, \dots, W_t)) \right| + \left| \overline{N}_{Bt}(F_h^M) \right| \\ &\leq \delta_W^M \delta_Y + \delta_Y \end{aligned} \tag{A.1}$$

For the second part, it suffices to show that

$$E \left\{ \left[\prod_{j=t-M+1}^t \frac{f_h(W_j)}{e_j(W_j)} \right] N_B(Y_t(\bar{\mathbf{w}}_{t-M}, W_{t-M+1}, \dots, W_t)) \mid \bar{H}_{t-M}^* \right\} = \bar{N}_{Bt}(F_h^M),$$

where the expectation is taken with respect to the assignment of treatments $W_{(t-M+1):t}$.

$$\begin{aligned} & E \left\{ \left[\prod_{j=t-M+1}^t \frac{f_h(W_j)}{e_j(W_j)} \right] N_B(Y_t(\bar{\mathbf{w}}_{t-M}, W_{t-M+1}, \dots, W_t)) \mid \bar{H}_{t-M}^* \right\} \\ &= \int \left[\prod_{j=t-M+1}^t \frac{f_h(w_j)}{e_j(w_j)} \right] N_B(Y_t(\bar{\mathbf{w}}_{t-M}, \underbrace{w_{t-M+1}, \dots, w_t}_{\mathbf{w}_{(t-M+1):t}})) \times \\ & \quad f(w_{t-M+1} \mid \bar{H}_{t-M}^*) f(w_{t-M+2} \mid \bar{H}_{t-M}^*, W_{t-M+1}) \cdots \times \\ & \quad f(w_t \mid \bar{H}_{t-M}^*, W_{(t-M+1):(t-1)}) d\mathbf{w}_{(t-M+1):t} \\ &= \int \left[\prod_{j=t-M+1}^t \frac{f_h(w_j)}{e_j(w_j)} \right] N_B(Y_t(\bar{\mathbf{w}}_{t-M}, w_{t-M+1}, \dots, w_t)) \times \\ & \quad f(w_{t-M+1} \mid \bar{H}_{t-M}^*) f(w_{t-M+2} \mid \bar{H}_{t-M+1}^*) \cdots f(w_t \mid \bar{H}_{t-1}^*) d\mathbf{w}_{(t-M+1):t} \\ & \quad \quad \quad \text{(because } \bar{H}_{t'+1}^* = \bar{H}_{t'}^* \cup \{W_{t'+1}\}) \\ &= \int N_B(Y_t(\bar{\mathbf{w}}_{t-M}, w_{t-M+1}, \dots, w_t)) \left[\prod_{j=t-M+1}^t f_h(w_j) \right] d\mathbf{w}_{(t-M+1):t} \\ & \quad \quad \quad \text{(By Assumption 1)} \\ &= \bar{N}_{Bt}(F_h^M). \end{aligned}$$

This establishes the claim that A_{1t} is a martingale difference series with respect to filtration \mathcal{F}_{t-1} .

Claim. $(T - M + 1)^{-1} \sum_{t=M}^T E\{A_{1t}^2 I(|A_{1t}| > \epsilon \sqrt{T - M + 1}) \mid \mathcal{F}_{t-1}\} \xrightarrow{p} 0$ for every $\epsilon > 0$.

Let $\epsilon > 0$. Note that A_{1t} is bounded by $\delta_Y(\delta_W^M + 1)$ (see (A.1)). Choose T_0 as

$$T_0 = \operatorname{argmin}_{t \in \mathbb{N}^+} \{\epsilon \sqrt{t - M + 1} > \delta_Y(\delta_W^M + 1)\}$$

$$\begin{aligned}
&= \operatorname{argmin}_{t \in \mathbb{N}^+} \left\{ t > M - 1 + \left\lceil \frac{\delta_Y(\delta_W^M + 1)}{\epsilon} \right\rceil^2 \right\} \\
&= \left\lceil M - 1 + \left\lceil \frac{\delta_Y(\delta_W^M + 1)}{\epsilon} \right\rceil^2 \right\rceil
\end{aligned}$$

Then, for $T > T_0$, $\epsilon\sqrt{T - M + 1} > \epsilon\sqrt{T_0 - M + 1} > \delta_Y(\delta_W^M + 1)$ which leads to $I(|A_{1t}| > \epsilon\sqrt{T - M + 1}) = 0$ and $E(A_{1t}^2 I(|A_{1t}| > \epsilon\sqrt{T - M + 1}) \mid \mathcal{F}_{t-1}) = 0$. This proves the claim.

Combining the previous results to show asymptotic normality of the first error: Since A_{1t} has mean zero, $E(A_{1t}^2 \mid \mathcal{F}_{t-1}) = \operatorname{Var}(A_{1t} \mid \mathcal{F}_{t-1})$, and since $\bar{N}_{Bt}(F_h)$ is fixed, $\operatorname{Var}(A_{1t} \mid \mathcal{F}_{t-1}) = \operatorname{Var}(\hat{N}_{Bt}(F_h) \mid \mathcal{F}_{t-1}) = \operatorname{Var}(\hat{N}_{Bt}(F_h) \mid \bar{H}_{t-M}^*)$. This gives us that

$$\frac{1}{T - M + 1} \sum_{t=M}^T E(A_{1t}^2 \mid \mathcal{F}_{t-1}) = \frac{1}{T - M + 1} \sum_{t=M}^T \operatorname{Var}(\hat{N}_{Bt}(F_h^M) \mid \bar{H}_{t-M}^*) \xrightarrow{p} v,$$

from Assumption 3(b). Combining these results, using that $\sqrt{T}/\sqrt{T - M + 1} \rightarrow 1$ and Theorem 4.16 of [van der Vaart \(2010\)](#),

$$\sqrt{T} \left(\frac{1}{T - M + 1} \sum_{t=M}^T A_{1t} \right) \xrightarrow{d} N(0, v).$$

Showing convergence to zero of the second error.

The second error compares the integral of the kernel-smoothed outcome surface over the region of interest B with the actual number of points within the set B . We show that as T goes to infinity, and since the bandwidth of the kernel converges to 0, the error due to kernel smoothing also goes to zero. Specifically, we will show that

$$\sqrt{T} \left(\frac{1}{T - M + 1} \sum_{t=M}^T A_{2t} \right) \xrightarrow{p} 0.$$

We start by introducing some notation and making an assumption about the set B . For

$\epsilon > 0$, we use $\mathcal{N}_\epsilon(A)$ to denote the ϵ -neighborhood of a set A : $\mathcal{N}_\epsilon(A) = \{\omega \in \Omega : \text{there exists } a \in A \text{ with } \text{dist}(\omega, a) < \epsilon\}$. Also, we use ∂B to denote the boundary of B (its closure excluding the interior points), $\partial B = \overline{B} \setminus B^\circ$.

Assumption A.1 (Number of time periods with outcome active locations within a neighborhood of B 's boundary). *There exists $\delta_B > 0$ and $Q^* \in (1/2, 1)$ such that*

$$P \left(\sum_{t=M}^T I \left(\exists s \in S_{Y_t} \cap \mathcal{N}_{\delta_B}(\partial B) \right) > T^{1-Q^*} \right) \rightarrow 0, \text{ as } T \rightarrow \infty.$$

Assumption A.1 states that the probability that we observe more than T^{1-Q^*} time periods with outcome-active locations within a δ_B -neighborhood of B 's boundary goes to zero as the number of observed time periods increases.

Let $c_t = \prod_{j=t-M+1}^t f_h(W_j)/e_j(W_j)$, and write

$$\left| \frac{1}{T-M+1} \sum_{t=M}^T A_{2t} \right| = \left| \frac{1}{T-M+1} \sum_{t=M}^T c_t \left[\int_B \sum_{s \in S_{Y_t}} K_{b_T}(\omega; s) d\omega - N_B(Y_t) \right] \right|.$$

Then:

$$\begin{aligned} & \int_B \sum_{s \in S_{Y_t}} K_{b_T}(\omega; s) d\omega - N_B(Y_t) \\ &= \sum_{s \in S_{Y_t} \cap B} \int_B K_{b_T}(\omega; s) d\omega + \sum_{s \in S_{Y_t} \cap B^c} \int_B K_{b_T}(\omega; s) d\omega - N_B(Y_t) \\ &= \sum_{s \in S_{Y_t} \cap B} \left[1 - \int_{B^c} K_{b_T}(\omega; s) d\omega \right] + \sum_{s \in S_{Y_t} \cap B^c} \int_B K_{b_T}(\omega; s) d\omega - N_B(Y_t) \\ &= \sum_{s \in S_{Y_t} \cap B^c} \int_B K_{b_T}(\omega; s) d\omega - \sum_{s \in S_{Y_t} \cap B} \int_{B^c} K_{b_T}(\omega; s) d\omega. \end{aligned}$$

This shows that the error from smoothing the outcome surface at time t comes from (1) the kernel weight from points outside of B that falls within B , and (2) the kernel weight from

points inside B that falls outside B . Using this, we write:

$$\left| \frac{1}{T-M+1} \sum_{t=M}^T A_{2t} \right| = \left| \frac{1}{T-M+1} \sum_{t=M}^T c_t \left[\sum_{s \in S_{Y_t} \cap B^c} \int_B K_{b_T}(\omega; s) d\omega - \sum_{s \in S_{Y_t} \cap B} \int_{B^c} K_{b_T}(\omega; s) d\omega \right] \right|.$$

Take $\epsilon > 0$, and $Q \in (1/2, Q^*)$ where Q^* is the one in Assumption A.1. Then, we will show that $P(T^Q \{ \left| \frac{1}{T-M+1} \sum_{t=M}^T A_{2t} \right| > \epsilon \}) \rightarrow 0$ as $T \rightarrow \infty$, which implies that the second error converges to zero faster than \sqrt{T} (since $Q > 1/2$).

$$\begin{aligned} & P\left(T^Q \left\{ \left| \frac{1}{T-M+1} \sum_{t=M}^T A_{2t} \right| > \epsilon \right\}\right) \\ &= P\left(\left| \frac{1}{T-M+1} \sum_{t=M}^T c_t \left[\sum_{s \in S_{Y_t} \cap B^c} \int_B K_{b_T}(\omega; s) d\omega - \sum_{s \in S_{Y_t} \cap B} \int_{B^c} K_{b_T}(\omega; s) d\omega \right] \right| > \frac{\epsilon}{T^Q}\right) \\ &\leq P\left(\frac{1}{T-M+1} \sum_{t=M}^T c_t \sum_{s \in S_{Y_t} \cap B^c} \int_B K_{b_T}(\omega; s) d\omega > \frac{\epsilon}{2T^Q}\right) + \\ &\quad P\left(\frac{1}{T-M+1} \sum_{t=M}^T c_t \sum_{s \in S_{Y_t} \cap B} \int_{B^c} K_{b_T}(\omega; s) d\omega > \frac{\epsilon}{2T^Q}\right), \end{aligned}$$

where the last equation holds because $|A - B| > \epsilon$ implies that at least one of $|A|, |B| > \epsilon/2$.

Also, since all quantities are positive, we can drop the absolute value. Then, since $c_t \leq \delta_W^M$ from Assumption 2,

$$\begin{aligned} & P\left(T^Q \left\{ \left| \frac{1}{T-M+1} \sum_{t=M}^T A_{2t} \right| > \epsilon \right\}\right) \\ &\leq P\left(\frac{1}{T-M+1} \sum_{t=M}^T \sum_{s \in S_{Y_t} \cap B^c} \int_B K_{b_T}(\omega; s) d\omega > \frac{\epsilon}{2T^Q \delta_W^M}\right) + \\ &\quad P\left(\frac{1}{T-M+1} \sum_{t=M}^T \sum_{s \in S_{Y_t} \cap B} \int_{B^c} K_{b_T}(\omega; s) d\omega > \frac{\epsilon}{2T^Q \delta_W^M}\right). \end{aligned}$$

Use $s_{Y_t}^{out}$ to denote the point in S_{Y_t} that lies outside B and is the closest to B : $s_{Y_t}^{out} = \{s \in S_{Y_t} \cap B^c : \text{dist}(s, B) = \min_{s' \in S_{Y_t} \cap B^c} \text{dist}(s', B)\}$. Similarly, $s_{Y_t}^{in}$ is the point in $S_{Y_t} \cap B$

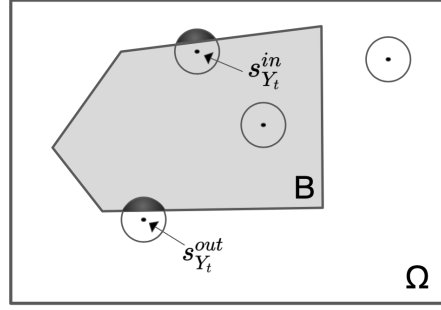


Figure A.1: Kernel-smoothed outcome surface, and points $s_{Y_t}^{in}, s_{Y_t}^{out}$ as the points closest to the boundary of B that lie within and outside B respectively. The amount of kernel weight falling within B from points outside of B is necessarily less or equal to the kernel weight from $s_{Y_t}^{out}$ (shaded), and similarly for $s_{Y_t}^{in}$.

that is closest to B^c . These points are shown graphically in Figure A.1. Because there are at most δ_Y outcome-active locations, from the definition of $s_{Y_t}^{in}, s_{Y_t}^{out}$, and because kernels are defined to be decreasing in distance, we have that

$$\begin{aligned}
& P\left(T^Q \left\{ \left| \frac{1}{T-M+1} \sum_{t=M}^T A_{2t} \right| \right\} > \epsilon \right) \\
& \leq P\left(\frac{1}{T-M+1} \sum_{t=M}^T \int_B K_{b_T}(\omega; s_{Y_t}^{out}) d\omega > \frac{\epsilon}{2T^Q \delta_W^M \delta_Y} \right) \\
& \quad + P\left(\frac{1}{T-M+1} \sum_{t=M}^T \int_{B^c} K_{b_T}(\omega; s_{Y_t}^{in}) d\omega > \frac{\epsilon}{2T^Q \delta_W^M \delta_Y} \right) \\
& = \underbrace{P\left(\sum_{t=M}^T \int_B K_{b_T}(\omega; s_{Y_t}^{out}) d\omega > \frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} \right)}_{B_1} \\
& \quad + \underbrace{P\left(\sum_{t=M}^T \int_{B^c} K_{b_T}(\omega; s_{Y_t}^{in}) d\omega > \frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} \right)}_{B_2}.
\end{aligned}$$

We show that B_1, B_2 converge to zero separately. Take B_1 :

$$\begin{aligned}
B_1 &= P\left(\sum_{t=M}^T \int_B K_{b_T}(\omega; s_{Y_t}^{out}) d\omega > \frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} \left| \sum_{t=M}^T I(s_{Y_t}^{out} \in \mathcal{N}_{\delta_B}(\partial B)) > T^{1-Q^*} \right) \right. \\
& \quad \times P\left(\sum_{t=M}^T I(s_{Y_t}^{out} \in \mathcal{N}_{\delta_B}(\partial B)) > T^{1-Q^*} \right) \\
& \quad + P\left(\sum_{t=M}^T \int_B K_{b_T}(\omega; s_{Y_t}^{out}) d\omega > \frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} \left| \sum_{t=M}^T I(s_{Y_t}^{out} \in \mathcal{N}_{\delta_B}(\partial B)) \leq T^{1-Q^*} \right) \right)
\end{aligned}$$

$$\times P\left(\sum_{t=M}^T I(s_{Y_t}^{out} \in \mathcal{N}_{\delta_B}(\partial B)) \leq T^{1-Q^*}\right)$$

From Assumption A.1 we have that

$$\begin{aligned} & P\left(\sum_{t=M}^T I(s_{Y_t}^{out} \in \mathcal{N}_{\delta_B}(\partial B)) > T^{1-Q^*}\right) \\ & \leq P\left(\sum_{t=M}^T I\left(\exists s \in S_{Y_t} \cap \mathcal{N}_{\delta_B}(\partial B)\right) > T^{1-Q^*}\right) \rightarrow 0, \end{aligned}$$

and $\lim_{T \rightarrow \infty} B_1$ is equal to

$$\lim_{T \rightarrow \infty} P\left(\sum_{t=M}^T \int_B K_{b_T}(\omega; s_{Y_t}^{out}) d\omega > \frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} \left| \sum_{t=M}^T I(s_{Y_t}^{out} \in \mathcal{N}_{\delta_B}(\partial B)) \leq T^{1-Q^*} \right.\right).$$

Studying the latter quantity, we have that

$$\begin{aligned} & P\left(\sum_{t=M}^T \int_B K_{b_T}(\omega; s_{Y_t}^{out}) d\omega > \frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} \left| \sum_{t=M}^T I(s_{Y_t}^{out} \in \mathcal{N}_{\delta_B}(\partial B)) \leq T^{1-Q^*} \right.\right) \\ & \leq P\left(\sum_{\substack{t=M \\ s_{Y_t}^{out} \notin \mathcal{N}_{\delta_B}(\partial B)}}^T \int_B K_{b_T}(\omega; s_{Y_t}^{out}) d\omega > \frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} - T^{1-Q^*}\right) \\ & \leq P\left(\sum_{\substack{t=M \\ s_{Y_t}^{out} \notin \mathcal{N}_{\delta_B}(\partial B)}}^T \int_{\omega: \|\omega\| > \delta_B} K_{b_T}(\omega; \mathbf{0}) d\omega > \frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} - T^{1-Q^*}\right) \\ & \leq P\left((T-M+1) \int_{\omega: \|\omega\| > \delta_B} K_{b_T}(\omega; \mathbf{0}) d\omega > \frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} - T^{1-Q^*}\right) \\ & = I\left((T-M+1) \int_{\omega: \|\omega\| > \delta_B} K_{b_T}(\omega; \mathbf{0}) d\omega > \frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} - T^{1-Q^*}\right) \quad (\text{A.2}) \end{aligned}$$

where the first inequality follows from the fact that at most T^{1-Q^*} time periods had $s_{Y_t}^{out}$ within δ_B of set's B boundary, and $\int_B K_{b_T}(\omega; s_{Y_t}^{out}) \leq 1$ for those time periods. The second inequality follows from the fact that during the remaining time periods $s_{Y_t}^{out}$ was further than δ_B from B and $\int_B K_{b_T}(\omega; s_{Y_t}^{out}) \leq \int_{\omega: \|\omega - s_{Y_t}^{out}\| > \delta_B} K_{b_T}(\omega; s_{Y_t}^{out}) = \int_{\omega: \|\omega\| > \delta_B} K_{b_T}(\omega; \mathbf{0})$. The third inequality follows from not excluding the time periods with $s_{Y_t}^{out} \in \mathcal{N}_{\delta_B}(\partial B)$. Finally,

the last equality holds because there is no uncertainty in the statement so the probability turns to an indicator.

Since the bandwidth $b_T \rightarrow 0$ as $T \rightarrow \infty$, there exists $T_1 \in \mathbb{N}$ such that $b_T < \delta_B$ and $\int_{\omega: \|\omega\| > \delta_B} K_{b_T}(\omega; \mathbf{0}) d\omega = 0$ for all $T \geq T_1$. Also, since $\frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} - T^{1-Q^*} \rightarrow \infty$, there exists $T_2 \in \mathbb{N}$ such that $\frac{\epsilon(T-M+1)}{2T^Q \delta_W^M \delta_Y} - T^{1-Q^*} > 1$ for all $T \geq T_2$. Then, for all $T \geq T_0 = \max\{T_1, T_2\}$ we have that the quantity in (A.2) is equal to 0, showing that $\lim_{T \rightarrow \infty} B_1 = 0$. Similarly, we can show that $\lim_{T \rightarrow \infty} B_2 = 0$.

Combining all of these results we have that

$$P\left(T^Q \left\{ \left| \frac{1}{T-M+1} \sum_{t=M}^T A_{2t} \right| \right\} > \epsilon\right) \rightarrow 0,$$

as $T \rightarrow \infty$, establishing that the second error converges to zero faster than $1/\sqrt{T}$. \square

Proof of Lemma 1. Define $\Psi_t = [\hat{N}_{Bt}(F_h^M)]^2 - v_t^*$. Then, Ψ_t is a martingale difference series with respect to $\mathcal{F}_t = \bar{H}_{t-M+1}$: (1) $E(|\Psi_t|) < \infty$ since Ψ_t is bounded, and (2) $E(\Psi_t | \mathcal{F}_{t-1}) = E\left\{ [\hat{N}_{Bt}(F_h^M)]^2 | \bar{H}_{t-M}^* \right\} - v_t^* = 0$. Also, since $\hat{N}_{Bt}(F_h^M)$ is bounded we have that $\sum_{t=M}^\infty t^{-2} E(\Psi_t^2) < \infty$. From Theorem 1 in Csörgő (1968) we have that

$$\frac{1}{T-M+1} \sum_{t=M}^T \Psi_t = \frac{1}{T-M+1} \sum_{t=M}^T [\hat{N}_{Bt}(F_h^M)]^2 - \frac{1}{T-M+1} \sum_{t=M}^T v_t^* \xrightarrow{p} 0.$$

\square

Proof of Theorem 2. In order to prove asymptotic normality of $\hat{\tau}_B(F_{h_1}^M, F_{h_2}^M)$ we will rely on results in the proof of Theorem 1 above. Take

$$\begin{aligned} \hat{\tau}_{Bt}(F_{h_1}^M, F_{h_2}^M) - \tau_{Bt}(F_{h_1}^M, F_{h_2}^M) = \\ \left\{ \prod_{j=t-M+1}^t \frac{f_{h_2}(W_j)}{e_j(W_j)} - \prod_{j=t-M+1}^t \frac{f_{h_1}(W_j)}{e_j(W_j)} \right\} \int_B \sum_{s \in S_{Y_t}} K_{b_T}(\omega, s) d\omega - \tau_{Bt}(F_{h_1}^M, F_{h_2}^M) = \end{aligned}$$

$$\begin{aligned}
& \underbrace{\left\{ \prod_{j=t-M+1}^t \frac{f_{h_2}(W_j)}{e_j(W_j)} - \prod_{j=t-M+1}^t \frac{f_{h_1}(W_j)}{e_j(W_j)} \right\} N_B(Y_t(\bar{\mathbf{w}}_{t-M}, W_{t-M+1}, \dots, W_t)) - \tau_{Bt}(F_{h_1}^M, F_{h_2}^M) +}_{C_{1t}} \\
& \underbrace{\left[\prod_{j=t-M+1}^t \frac{f_{h_2}(W_j)}{e_j(W_j)} \right] \left[\int_B \sum_{s \in S_{Y_t}} K_{b_T}(\omega, s) d\omega - N_B(Y_t(\bar{\mathbf{w}}_{t-M}, W_{t-M+1}, \dots, W_t)) \right]}_{C_{2t}^2} + \\
& \underbrace{\left[\prod_{j=t-M+1}^t \frac{f_{h_1}(W_j)}{e_j(W_j)} \right] \left[\int_B \sum_{s \in S_{Y_t}} K_{b_T}(\omega, s) d\omega - N_B(Y_t(\bar{\mathbf{w}}_{t-M}, W_{t-M+1}, \dots, W_t)) \right]}_{C_{2t}^1}
\end{aligned}$$

Following steps identical to showing $\sqrt{T} \left[(T - M + 1)^{-1} \sum_{t=M}^T A_{2t} \right] \xrightarrow{p} 0$ in the proof of Theorem 1, we can equivalently show that $\sqrt{T} \left[(T - M + 1)^{-1} \sum_{t=M}^T C_{2t}^1 \right] \xrightarrow{p} 0$ and $\sqrt{T} \left[(T - M + 1)^{-1} \sum_{t=M}^T C_{2t}^2 \right] \xrightarrow{p} 0$.

Therefore, all we need to show is that $\sqrt{T} \left[(T - M + 1)^{-1} \sum_{t=M}^T C_{1t} \right] \xrightarrow{d} N(0, \eta)$. We will do so by showing again that C_{1t} is a martingale difference series with respect to the filtration \mathcal{F}_{t-1} :

1. Since $E(|A_{1t}|) < \infty$, from the triangular inequality we straightforwardly have that $E(|C_{1t}|) < \infty$.
2. Since $E(A_{1t} | \mathcal{F}_{t-1}) = 0$, we also have that $E(C_{1t} | \mathcal{F}_{t-1}) = 0$, from linearity of expectation.

Then, using the triangular inequality and (A.1), we have that C_{1t} is bounded by $2\delta_Y(\delta_W^M + 1)$. Then, for $\epsilon > 0$, choosing $T_0 = \underset{t \in \mathbb{N}^+}{\operatorname{argmin}} \{ \epsilon \sqrt{t - M + 1} > 2\delta_Y(\delta_W^M + 1) \}$ satisfies that, for $T > T_0$, $E(C_{1t}^2 I(|C_{1t}| > \epsilon \sqrt{T - M + 1}) | \mathcal{F}_{t-1}) = 0$. Combining these results, we have that $\sqrt{T} [\hat{\tau}_{Bt}(F_{h_1}^M, F_{h_2}^M) - \tau_{Bt}(F_{h_1}^M, F_{h_2}^M)] \rightarrow N(0, \eta)$.

To show $(T - M + 1)^{-1} \sum_{t=M}^T \left\{ [\hat{\tau}_{Bt}(F_{h_1}^M, F_{h_2}^M)]^2 - E\{[\tau_{Bt}(F_{h_1}^M, F_{h_2}^M)]^2 | \bar{H}_{t-M}^*\} \right\} \xrightarrow{p} 0$, the proof follows exactly the same way as the proof of Lemma 1 and is omitted here.

□

Proof of Proposition 1. Note that $f(W_t = w \mid e_t(w), \bar{H}_{t-1}) = f(W_t = w \mid \bar{H}_{t-1}) =$

$e_t(w)$ since $e_t(w)$ is a function of \bar{H}_{t-1} . Therefore, it suffices to show that $f(W_t = w \mid e_t(w)) = e_t(w)$:

$$f(W_t = w \mid e_t(w)) = E[f(W_t = w \mid \bar{H}_{t-1}) \mid e_t(w)] = E[e_t(w) \mid e_t(w)] = e_t(w). \quad (\text{A.3})$$

□

Proof of Proposition 2. We show that $f(W_t = w \mid e_t(w), \bar{\mathcal{Y}}_T, \bar{\mathcal{X}}_T) = f(W_t = w \mid e_t(w))$.

$$\begin{aligned} & f(W_t = w \mid e_t(w), \bar{\mathcal{Y}}_T, \bar{\mathcal{X}}_T) \\ &= E \left[f(W_t = w \mid \bar{H}_{t-1}, e_t(w), \bar{\mathcal{Y}}_T, \bar{\mathcal{X}}_T) \mid e_t(w), \bar{\mathcal{Y}}_T, \bar{\mathcal{X}}_T \right] \\ &= E \left[f(W_t = w \mid \bar{H}_{t-1}, \bar{\mathcal{Y}}_T, \bar{\mathcal{X}}_T) \mid e_t(w), \bar{\mathcal{Y}}_T, \bar{\mathcal{X}}_T \right] \\ & \quad \text{(Because } e_t(w) \text{ is a function of } \bar{H}_{t-1}) \\ &= E \left[f(W_t = w \mid \bar{H}_{t-1}) \mid e_t(w), \bar{\mathcal{Y}}_T, \bar{\mathcal{X}}_T \right] \quad \text{(From Assumption 1)} \\ &= E \left[e_t(w) \mid e_t(w), \bar{\mathcal{Y}}_T, \bar{\mathcal{X}}_T \right] \\ &= e_t(w) \\ &= f(W_t = w \mid e_t(w)) \quad \text{(From (A.3))} \end{aligned}$$

□

C The Hájek Estimator

The standardization of weights used in the Hájek estimator is known to be effective in the settings where the weights are extreme. Its sample boundedness property guarantees that the resulting estimate is always within the range of the observed outcome. In our case, the Hájek estimator replaces the division by $T - M + 1$ with that by $\sum_{t=M}^T v_t$ where v_t is the

product of fractions in (6). For example,

$$\widehat{N}_B(F_h^M)_{\text{Hájek}} = \frac{1}{\sum_{t=M}^T v_t} \sum_{t=M}^T \widehat{N}_{Bt}(F_h^M)$$

Unfortunately, our asymptotic results (i.e., Theorems 1 and 2) do not easily extend to the Hájek estimator. The reason is that the estimator includes $\sum_{t=M}^T v_t$, a quantity that depends on all time periods, making it difficult to apply the martingale theory used in our proofs. Therefore, we use a heuristic approach to estimating the variance bound of the Hájek estimator. Since the Hájek estimator simply rescales the corresponding IPW estimator by $(T - M + 1) / \sum_{t=M}^T v_t$, we scale the variance bound derived for the estimator by $[(T - M + 1) / (\sum_{t=M}^T v_t)]^2$.

D Additional Simulation Results

D.1 Point Estimates for the Interventions over Many Time Periods

Figure A.2 shows the performance of the estimators for the interventions over many time periods. The plots show the estimated change in the number of outcome-active locations over the sub-region $B = [0.75, 1]$ for a change in the stochastic intervention from 3 per time period to the value on the horizontal axis. The rows correspond to the interventions over $M = 3, 7$, and 30 time periods, respectively, whereas the columns represent the different lengths of time series, i.e., $T = 200, 400$ and 500. The results are shown for the IPW estimators based on the true propensity score (purple lines with open triangles) and the estimated propensity score (blue lines with x's) as well as the Hájek estimator based on the estimated propensity score (orange lines with open rhombuses). Only the Hájek estimates are shown for $M = 30$ as the extremely small weights arising from a large number of time periods make the estimates from the other estimators essentially equal to zero. The lines and points in the plot show the median estimate and the rectangles show the interquartile range of estimates across 200 simulated data sets.

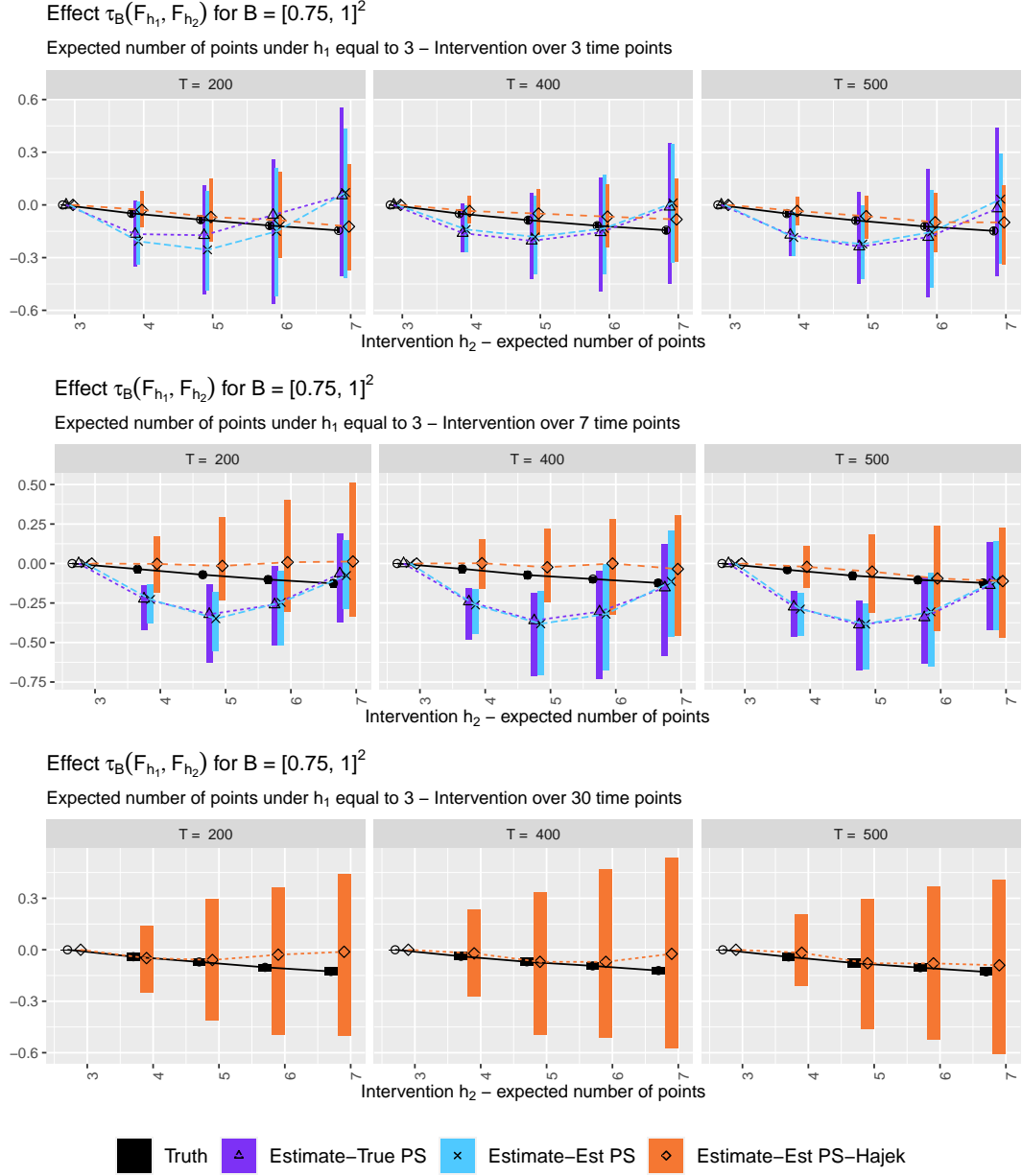


Figure A.2: Simulation Results for the Interventions of Increasing Time Lengths. Rows correspond to the interventions taking place over $M = 3, 7$, and 30 time periods. Columns correspond to the increasing length of the time series from 200 (left plots) to 500 (right plots). The vertical axis shows the change in the expected number of the outcome active locations over $[0.75, 1]^2$ for a change in the intervention intensity from 3 under h_1 to the value shown in the horizontal axis under h_2 , for M time periods. The points in the plot show the median estimate over 200 data sets, and the rectangles show the interquartile range of estimates. Only the Hájek estimates are shown for $M = 30$ as the extremely small weights arising from a large number of time periods make the estimates from the other estimators close to zero.

The results indicate that the IPW estimator’s performance deteriorates as the length of the intervention period M increases. This is expected since the weight at each time point is the product of M ratios as shown in (6). The data provide little information about what would happen under a hypothetical intervention taking place over a long time period unless the specified intervention resembles the actual data generating mechanism for the observed data. However, Figure A.2 shows that the standardization of weights performed in the Hájek estimator can alleviate some of these issues, yielding effect estimates close to the true values.

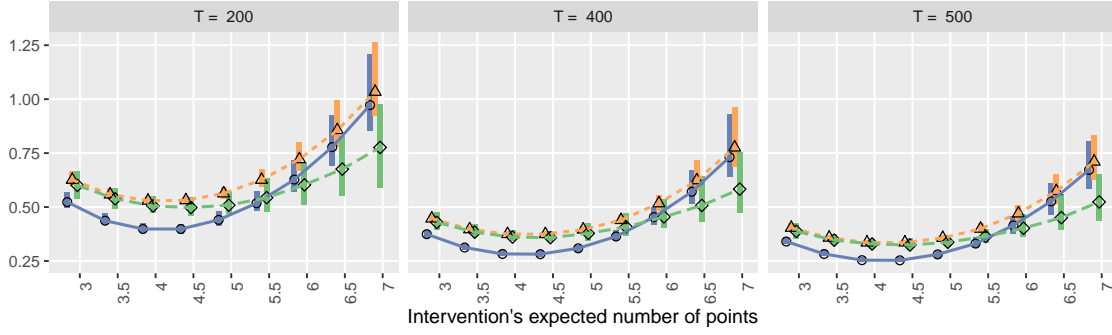
D.2 Asymptotic Variance and Bound, and Estimated Variance Bound

Figure 5 shows the average (over 200 simulated data sets) of the true asymptotic standard deviation and true bound as well as the estimated standard deviation bound of the IPW estimator for the average potential outcome using the true propensity score, for interventions taking place over $M = 3$ time periods. Figure A.3 is a similar plot for the interventions taking place over $M = 1, 3$, and 7 (rows) time periods, and observed time series of length $T = 200, 400, 500$ (columns). These plots show the median and interquartile range of the asymptotic standard deviation, true bound, and estimated bound over 200 simulated data sets.

We begin by focusing on low uncertainty scenarios, corresponding to the interventions taking place over $M = 1$ or 3 time periods with the distribution resembling the actual data generating mechanism. We think that the intervention distribution resembles the data generating mechanism in scenarios where the intervention intensity is close to 5, which is the average number of treatment-active locations for the data generating process. In these scenarios, the asymptotic variance bound is distinctly higher than the true asymptotic variance, indicating that the inference based on the true asymptotic bound would be conservative. We find that in these low uncertainty scenarios, the estimated bound is close to the true bound. For that reason, we would expect the confidence intervals for the IPW estimator based on

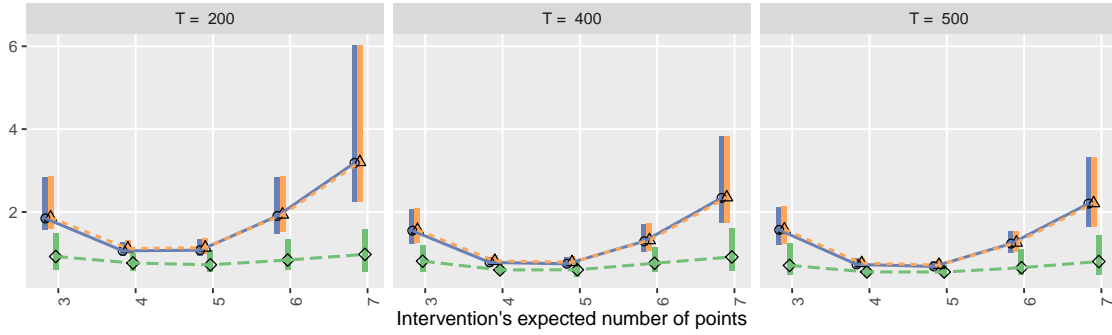
Standard deviation and bound of the average potential outcome estimator for $B = [0, 0.5]^2$

Intervention over 1 time point



Standard deviation and bound of the average potential outcome estimator for $B = [0, 0.5]^2$

Intervention over 3 time points



Standard deviation and bound of the average potential outcome estimator for $B = [0, 0.5]^2$

Intervention over 7 time points

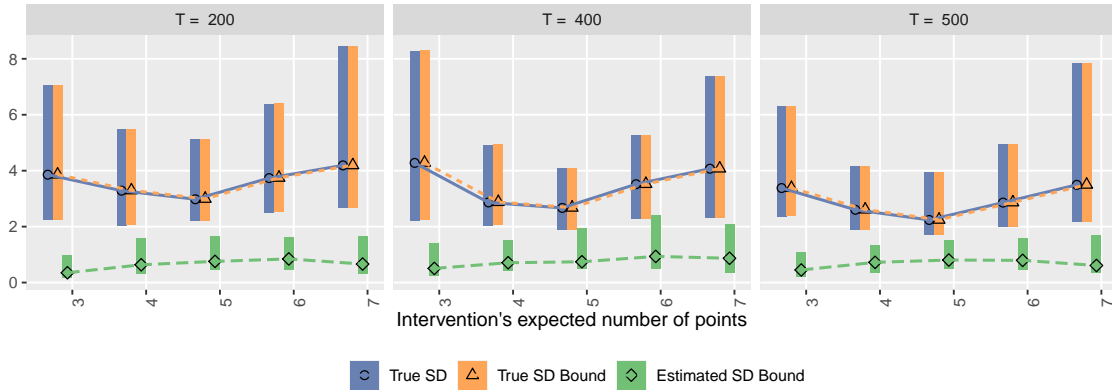


Figure A.3: Asymptotic Standard Deviation and Bound, and Estimated Bound. This figure shows the true asymptotic standard deviation (blue circles), the true asymptotic bound (orange triangles), and the estimated bound (green rhombuses) of the IPW estimator for the average potential outcome using the true propensity score, under interventions that take place over $M = 1, 3$ and 7 time periods (rows), and for increasing length of the time series (columns). The horizontal axis shows the intensity of the intervention at each time period. The points show the median value, and the rectangles show the interquartile range over 200 simulated data sets.

the estimated bound to have a higher coverage probability than its nominal coverage (see Appendix D.3 for the coverage results).

In contrast, under high uncertainty scenarios such as the interventions over longer time periods, e.g., $M = 7$, the asymptotic standard deviation and theoretical bound are essentially indistinguishable. However, under these scenarios, the estimate of the theoretical bound tends to be biased downwards, suggesting that the confidence intervals for the IPW estimator based on the estimated bound would be anti-conservative. Furthermore, we expect it to take a longer time series in order for the estimated bound to converge to its theoretical value when the intervention takes place over a longer time period.

D.3 Coverage of the Confidence Intervals for the IPW and Hájek Estimators

IPW estimator. The results in Figure A.3 indicate that the coverage of confidence intervals based on the asymptotic variance bound should be similar to those based on the true variance under high uncertainty scenarios, while they should be slightly higher under low uncertainty scenarios. Furthermore, confidence intervals based on the estimated variance bound should yield coverage probability close to (lower than) the nominal coverage under low (high) uncertainty scenarios.

These expectations are indeed reflected in the coverage results shown in Figure A.4. Furthermore, for $M \in \{1, 3, 7\}$ and $T = 500$, the confidence intervals using the true variance achieve a coverage that is close to the nominal level. In contrast, for $M = 30$, the confidence intervals based on the true asymptotic variance have a coverage of 50% or less, indicating that for interventions taking place over longer time periods, more data are needed to make use of the asymptotic approximation.

Hájek estimator. Motivated by the good performance of the Hájek estimator shown in Figure A.2, we also investigate the coverage probability of the 95% confidence interval as described in Appendix C. The rows of Figure A.5 show the coverage results for increasingly small regions, i.e., $B_1 = [0, 1]^2$, $B_2 = [0, 0.5]^2$, and $B_3 = [0.75, 1]^2$, whereas the columns

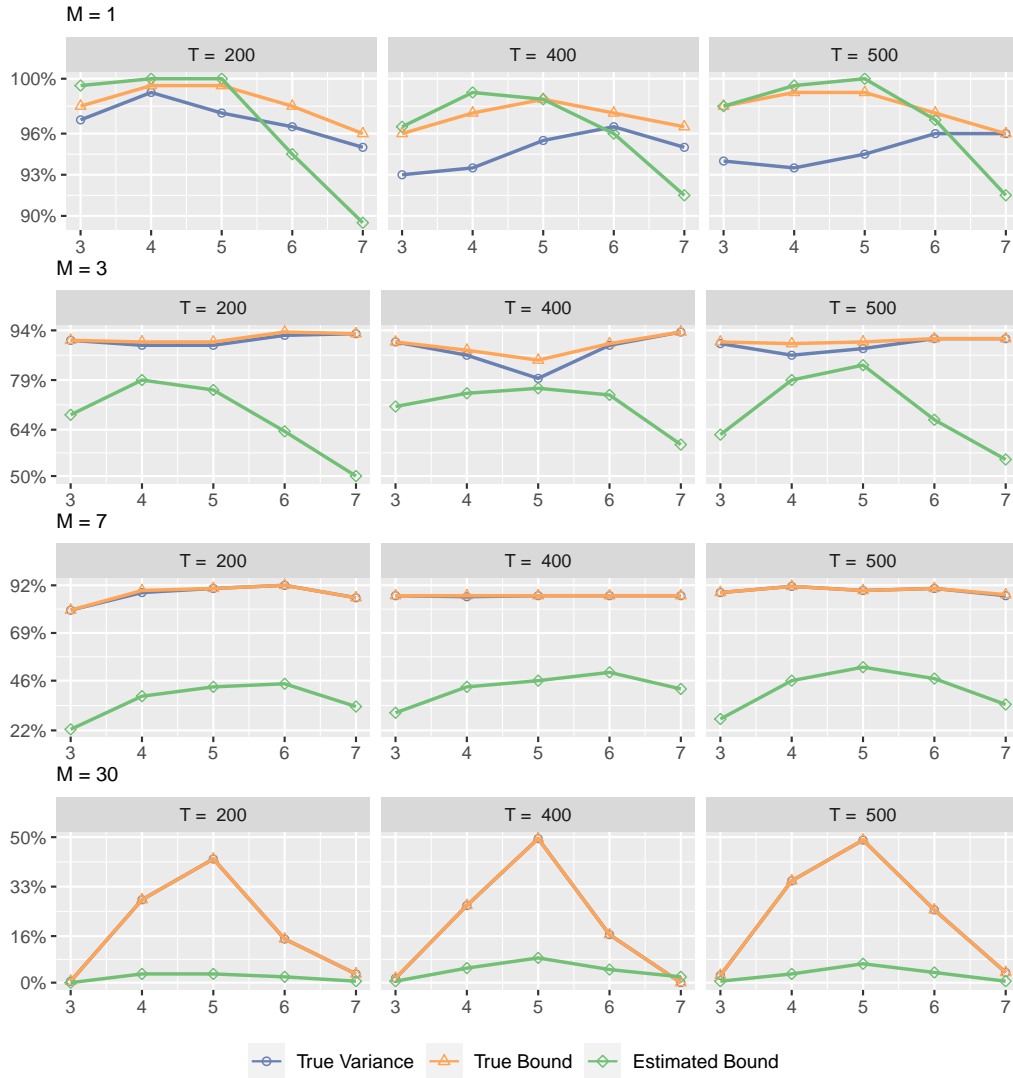


Figure A.4: Coverage of the IPW Estimator 95% Confidence Intervals. This figure shows the coverage of 95% confidence intervals for the average potential outcome over $B = \Omega$ based on the IPW estimator using the true variance (blue lines open circles), the true bound (orange lines with triangles), and the estimated bound (green lines with rhombuses), for interventions taking place over $M \in \{1, 3, 7, 30\}$ time periods (rows) and increasing length of the observed time series (columns).

show the results for increasingly long observed time series ($T = 200, 400, 500$). Different colors correspond to the coverage results under interventions taking place over $M = 1$ (black), 3 (green), 7 (red), and 30 (blue) time periods. We find that the coverage is above 85% for all cases, even when an intervention takes place over 30 time periods. As expected, the coverage is higher for smaller values of M , since these correspond to lower-uncertainty situations. We also find that the coverage is lower for smaller regions, a phenomenon we do not observe for the IPW estimator.

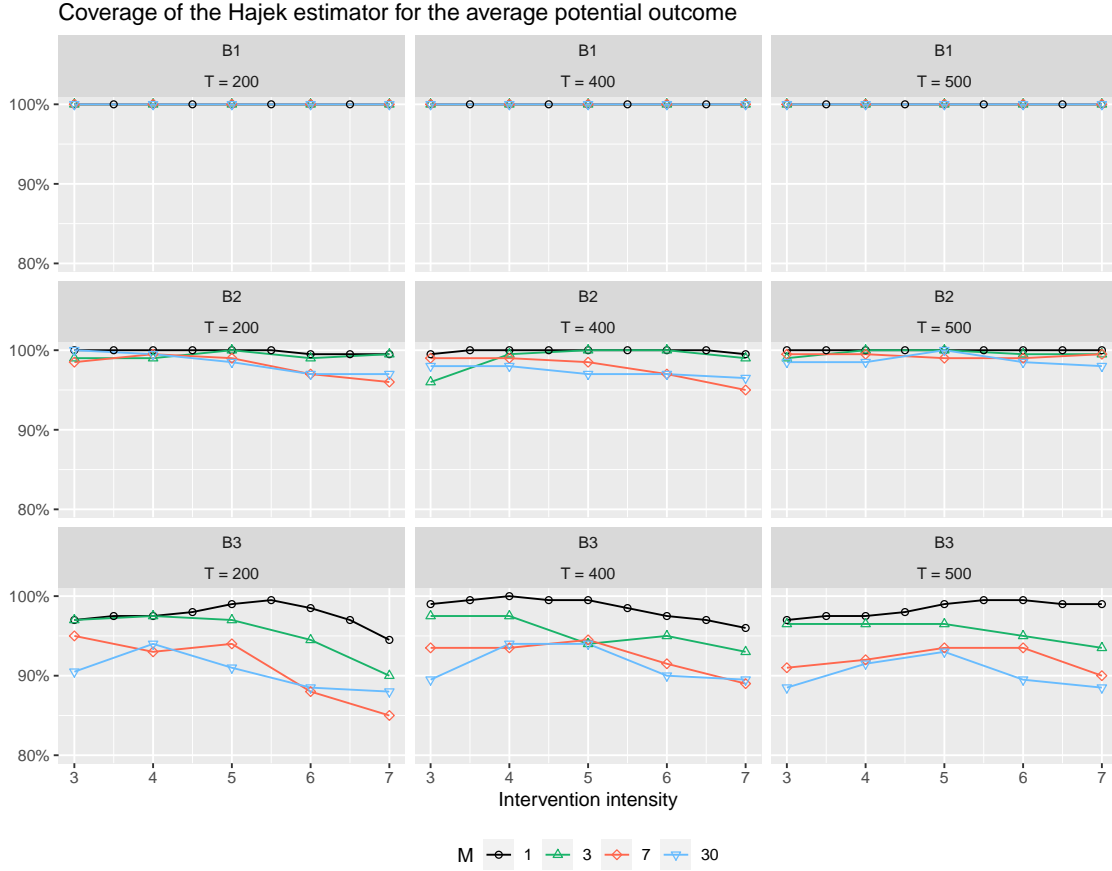


Figure A.5: Coverage of the Hájek Estimator’s 95% Confidence Intervals for the Average Potential Outcomes under Various Interventions. We vary the intervention intensity h (horizontal axis), and the length of intervention $M = 1, 3, 7, 30$ (different lines). Each row represents the coverage for different regions of interest, i.e., $B_1 = [0, 1]^2$, $B_2 = [0, 0.5]^2$ and $B_3 = [0.75, 1]^2$, whereas each column represents the length of time series, i.e., $T = 200, 400$ and 500 .

D.4 Uncertainty Estimates

We also compute the standard deviation of the estimated average potential outcome across simulated data sets and compare it with the mean of the standard deviations, each of which is used to create the confidence intervals. The similarity of these two quantities implies the accuracy of our uncertainty estimates. Figure A.6 presents the results as the ratio of these two quantities. A value below (above) 1 indicates that the true variability in our point estimates is smaller (greater) than our uncertainty estimate.

While the ratios are always below 1 for the Hájek estimator (bottom panel), they are almost always above 1 for the IPW estimator (top panel). This is consistent with the above

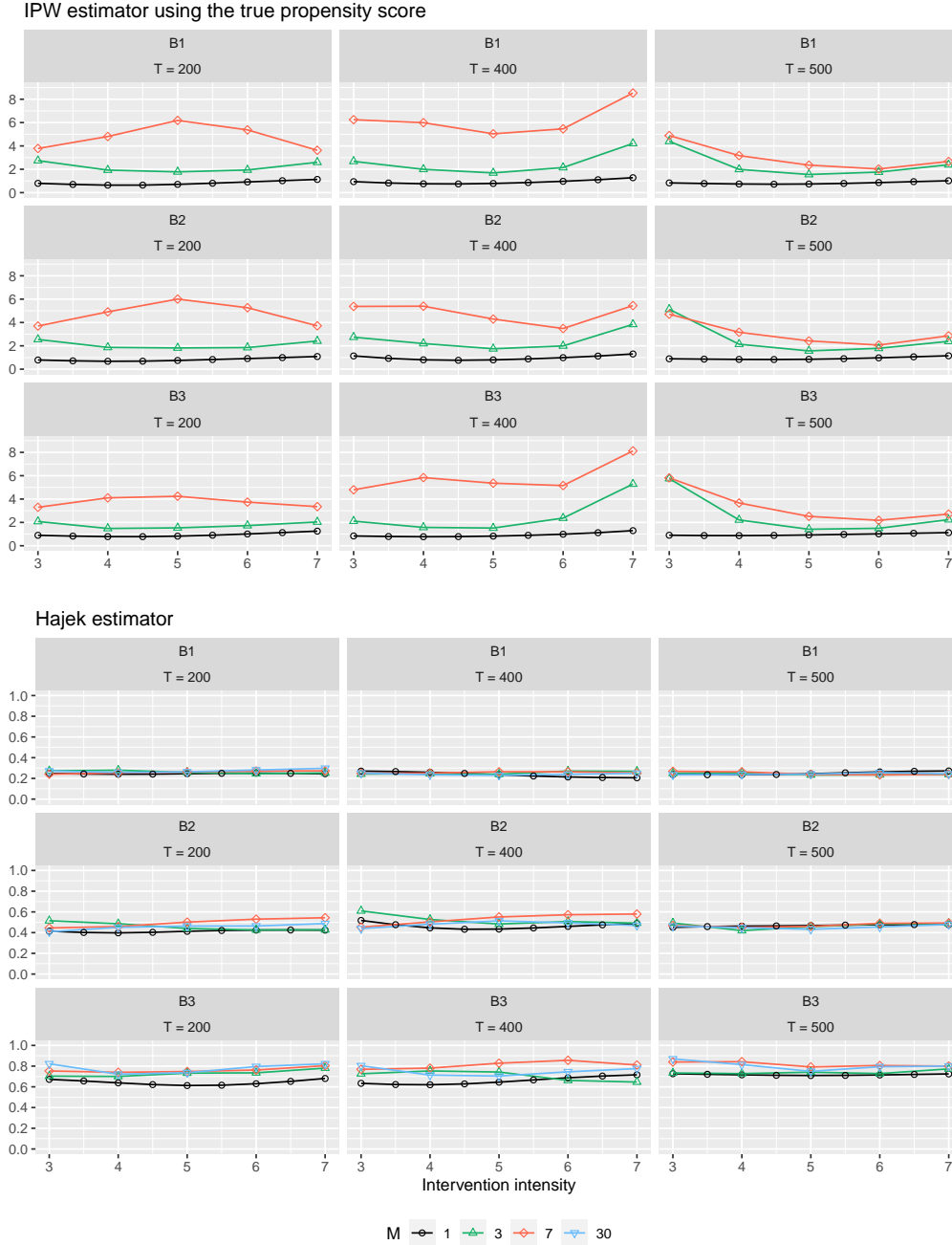


Figure A.6: Comparison of the Estimated and True Uncertainty for the Inverse Probability of Treatment and Hájek Estimators. Each plot presents the ratios between the standard deviation of each estimator and the mean estimated standard deviation across simulated data sets. A value smaller (greater) than 1 implies underestimation (overestimation) of uncertainty. The top (bottom) panel presents the results for the IPW (Hájek) estimator with the varying intensity under the intervention (horizontal axis) and for the whole region B_1 (first and forth row) and two sub-regions, $B_2 = [0, 0.5]^2$ (second and fifth row) and $B_3 = [0.75, 1]^2$ (third and sixth row). We also vary the length of intervention, $M = 1, 3, 7$ and 30 time periods (black, green, red, and blue lines, respectively). The columns correspond to different lengths of the time series $T = 200, 400$ and 500.

results, showing that we tend to overestimate (underestimate) the uncertainty for the Hájek (IPW) estimator. We find that the confidence interval for the Hájek estimator tends to be most conservative when M is small and the region of interest is large. For the IPW estimator, the degree of underestimation decreases as the length of time series T increases but increases as the length of intervention M increases. In fact, when $M = 30$, some of the ratios are as large as 20 (hence they are not included in the figure). The results suggest that in practice the Hájek estimator should be preferred over the IPW estimator especially for stochastic interventions over a long time period.

D.5 Covariate Balance

We evaluate the balance of covariates based on the estimated propensity score by comparing their p-values in the propensity score model, and in a model with functional form as in the propensity score model but weighted by the inverse of the estimated propensity score. The left plot of Figure A.7 shows the p-value for the previous outcome-active locations, which are one of the time-varying confounders, across 200 simulated data sets. Evidently, the p-values in the unweighted model are close to 0, indicating that previous outcome-active locations form an important predictor of the treatment assignment. However, in the weighted model, the p-values of the same confounder are more evenly distributed across the $(0, 1)$ range, indicating that this confounder is better balanced in the weighted time series. The right plot of Figure A.7 shows the distribution (over 200 simulated data sets) of the estimated coefficient for this confounder in the weighted model. This distribution is centered around zero, indicating that balance is achieved on average across data sets.

E Additional Empirical Results

E.1 Visualization

As discussed in Section 6.1, we consider a stochastic intervention whose focal point is the center of Baghdad. The degree of concentration is controlled by the precision parameter α

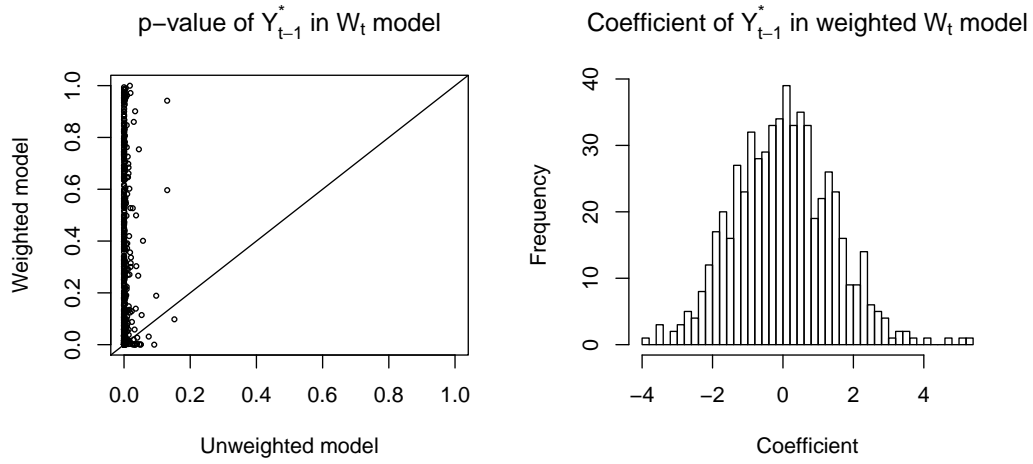


Figure A.7: Balance of the Previous Outcome-Active Locations in Treatment Model. In the left plot, each point shows the relative magnitude of the p-value for the previous outcome-active locations in the unweighted propensity score model (horizontal axis) over that of the model weighted by the inverse of the estimated propensity score (vertical axis). The right plot shows the distribution of the estimated coefficient of the previous outcome-active locations in the weighted propensity score model.

whose greater value, implying that more airstrikes are occurring near the focal point. We vary the value of α from 0 to 3, while keeping the expected number of airstrikes constant at 3 per day. Figure A.8 illustrates intensities for the different values of α . The first plot in the figure does not focus on Baghdad at all, representing the baseline spatial distribution ϕ_0 . As the value of α increases, the spatial distribution of airstrikes becomes concentrated more towards the center of Baghdad.

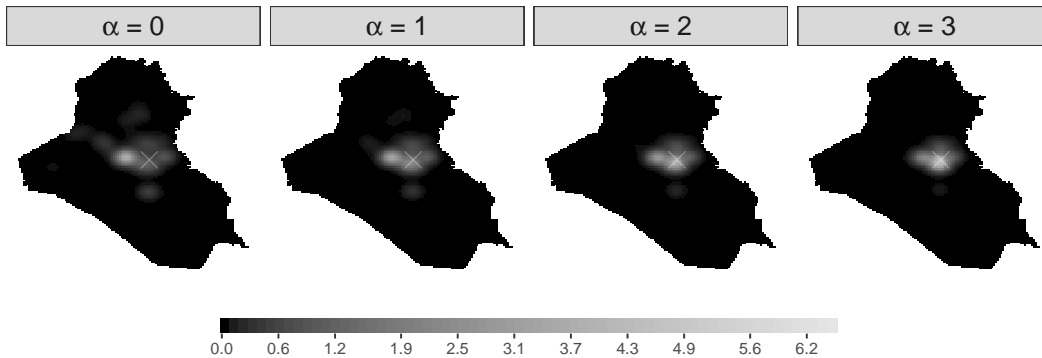


Figure A.8: Visualization of Intensity under Stochastic Interventions whose Focal Point is the Center of Baghdad. Across plots, we vary the degree to which the airstrikes are concentrated around the focal point using the precision parameter, while the expected number of airstrikes is held constant at 3 per day.

E.2 Empirical Results

Table A.2 presents the numerical effect estimates and 95% confidence intervals for various interventions, including those shown in the main text. We also show the effect estimates for the whole Iraq, Baghdad only, and the area outside Baghdad. Statistically significant estimates include the estimated effects of increasing the expected number of airstrikes from 1 to 3 per day for a period of a week on all types of insurgent attacks, with the increase located mostly outside Baghdad. In addition, continuing this intervention for a period of

Type ($F_{h'}$, $F_{h''}$)	M	Outcome	Iraq	Baghdad	Outside Baghdad
Increasing the intensity (1, 3)	3	IED	1.5 (-2.5, 5.5)	0.1 (-1, 1.3)	1.4 (-1.5, 4.3)
		SAF	3.6 (-0.4, 7.7)	1.4 (-0.1, 2.8)	2.3 (-0.4, 4.9)
		Other Attack	5.1 (-3.9, 14)	1.8 (-1, 4.7)	3.2 (-3.1, 9.5)
	7	IED	4.9 (-0.1, 10)	1 (-0.4, 2.4)	4 (0.2, 7.7)
		SAF	7 (1.9, 12.1)	2 (0.4, 3.5)	5 (1.4, 8.6)
		Other Attack	12.3 (0.4, 24.1)	3.5 (0, 7.1)	8.7 (0.3, 17.1)
	30	IED	-2.5 (-5.9, 1)	-1 (-1.8, -0.2)	-1.4 (-4.1, 1.2)
		SAF	-0.6 (-4.7, 3.5)	-1.3 (-2.5, -0.2)	0.8 (-2.3, 3.9)
		Other Attack	-6.6 (-14.8, 1.7)	-4 (-6.3, -1.6)	-2.6 (-8.7, 3.5)
Changing the focal points (0, 3)	3	IED	-1 (-9.5, 7.5)	-0.4 (-2.6, 1.7)	-0.6 (-7, 5.9)
		SAF	-1.2 (-9.9, 7.5)	0.3 (-2.6, 3.3)	-1.5 (-7.5, 4.4)
		Other Attack	-4.9 (-23.9, 14)	-2 (-8, 3.9)	-2.9 (-16.1, 10.3)
	7	IED	-3.6 (-14.3, 7)	-1.9 (-4.6, 0.8)	-1.7 (-9.9, 6.4)
		SAF	-5.6 (-16.6, 5.4)	-3.1 (-6.6, 0.5)	-2.5 (-10.2, 5.2)
		Other Attack	-18.4 (-44, 7.3)	-7.1 (-14.9, 0.7)	-11.3 (-29.4, 6.9)
	30	IED	-3.2 (-18.3, 12)	-1.5 (-5.3, 2.3)	-1.7 (-13.2, 9.8)
		SAF	2.4 (-15.1, 19.9)	-3.9 (-9.5, 1.6)	6.4 (-6.1, 18.8)
		Other Attack	-7.8 (-44.3, 28.7)	-8.1 (-19.6, 3.4)	0.3 (-25.3, 25.9)
Lagged effects (1, 5)	3	IED	-2.3 (-10.8, 6.2)	-0.6 (-2.9, 1.6)	-1.7 (-8, 4.7)
		SAF	-2.8 (-11.7, 6.2)	-1.8 (-4.7, 1.2)	-1 (-7.2, 5.1)
		Other Attack	-7.9 (-28.7, 13)	-3.6 (-10.3, 3)	-4.2 (-18.6, 10.2)
	7	IED	6.7 (-0.6, 14)	1.7 (-0.3, 3.6)	5 (-0.4, 10.5)
		SAF	7 (0.7, 13.4)	2.9 (0.6, 5.2)	4.1 (-0.1, 8.4)
		Other Attack	16.8 (0, 33.5)	6.8 (0.9, 12.8)	10 (-1.1, 21)
	30	IED	1.8 (-3.9, 7.6)	0.4 (-1, 1.9)	1.4 (-2.9, 5.7)
		SAF	1.9 (-3.1, 7)	0.7 (-1.1, 2.5)	1.2 (-2, 4.5)
		Other Attack	5 (-9.1, 19.1)	2.5 (-3.2, 8.1)	2.6 (-5.9, 11)

Table A.2: Causal Effect Estimates and 95% Confidence Intervals for Various Stochastic Interventions. We present the results for three interventions discussed in the main text: increasing the expected number of airstrikes from 1 to 3 per day for M days, changing the focal points of airstrikes from $\alpha = 0$ to $\alpha = 3$ for M days, and the lagged effects of increasing the expected number of airstrikes from 1 to 5 per day M days ago. The range of M we consider is $\{3, 7, 30\}$. The regions of interest are Iraq, Baghdad, and the area outside Baghdad. The results in bold represent statistically significant estimates.

thirty days leads to a decrease in the number of insurgent attacks in Baghdad. For $M = 3$, all estimated lagged effects are negative, whereas all estimated lagged effects for $M = 7$ are positive, indicating that increasing the number of airstrikes might have a short term effect in reducing insurgent violence, but leading to a long-term increase.

E.3 Local Interventions

In this section, we present the results of *local interventions*, which change the airstrike strategy over a region of the country, $L \subset \Omega$. Consider intensities h_1, h_2 which are equal to each other outside this region, i.e., $h_1(\omega) = h_2(\omega) = c_{L^c}\phi_0(\omega)$, for $\omega \in L^c = \Omega \setminus L$. Then, comparing average potential outcomes under this strategies would represent the change in the expected number of insurgent attacks if the airstrike strategy over the area L^c were identical, and the airstrike locations within L were determined according to h_1 versus h_2 .

In our analysis, for the sake of illustration, we set L to the administrative region of Baghdad and consider the airstrike strategy that sets the expected number of airstrikes outside of Baghdad to 3 per day, and that within the Baghdad administrative unit to $\{1, 2, 3, 4\}$. This strategy is shown in Figure A.9.

Table A.3 shows the estimated causal effects for this local intervention. We find that the intervention does not lead to any statistically significant change in insurgent attacks of

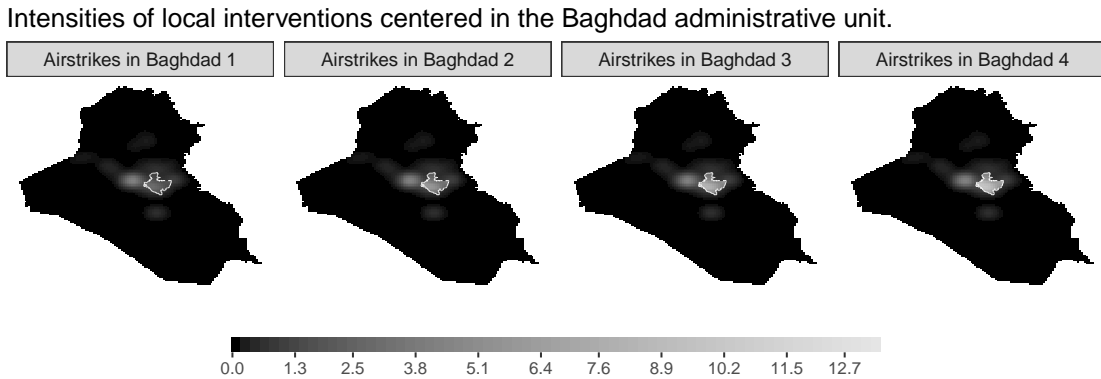


Figure A.9: Visualization of the Local Intervention Strategy. Under this strategy, we set the expected number of airstrikes outside the Baghdad administrative unit (polygon in the map) to 3 with constant spatial distribution, while changing the intensity within Baghdad from 1 per day to 4 per day.

any type.

M	Outcome	Iraq	Baghdad	Outside Baghdad
3	IED	3.9 (-6.5, 14.4)	1.8 (-1.1, 4.6)	2.1 (-5.6, 9.9)
	SAF	-0.1 (-10.8, 10.5)	0.3 (-3.3, 3.9)	-0.4 (-7.7, 6.8)
	Other Attack	8.1 (-17.3, 33.4)	3.7 (-4.8, 12.3)	4.3 (-12.9, 21.5)
7	IED	3.8 (-7.3, 14.9)	1 (-1.9, 3.8)	2.8 (-5.6, 11.3)
	SAF	1.1 (-9.9, 12.1)	0.8 (-2.9, 4.6)	0.3 (-7.1, 7.7)
	Other Attack	8.9 (-18.1, 35.9)	3.7 (-5.1, 12.6)	5.2 (-13.3, 23.7)
30	IED	-3.3 (-16.6, 10)	-1.7 (-5.2, 1.7)	-1.6 (-11.5, 8.4)
	SAF	0.1 (-14.4, 14.6)	-4 (-8.5, 0.5)	4.1 (-6.3, 14.5)
	Other Attack	-10.8 (-43, 21.5)	-7.6 (-18, 2.9)	-3.2 (-25.5, 19.1)

Table A.3: Causal Effect Estimates for Local Intervention Strategy. The duration of the strategy is set to $M \in \{3, 7, 30\}$. The strategy increases the expected number of airstrikes within the Baghdad administrative unit from 1 to 4 per day while keeping the strategy for the rest of the country unchanged at 3 per day. The 95% confidence intervals are in parentheses.