Received: 27 April 2020 | Accepted: 22 June 2022

DOI: 10.1111/rssb.12548

ORIGINAL ARTICLE

Series B Statistical Methodology

Causal inference with spatio-temporal data: Estimating the effects of airstrikes on insurgent violence in Iraq

Georgia Papadogeorgou[1] | **Kosuke Imai**[2] | **Jason Lyall**[3] | **Fan Li**[4]

[1]Department of Statistics, University of Florida, Gainesville, Florida, USA

[2]Department of Government and Department of Statistics, Institute for Quantitative Social Science, Harvard University, Cambridge, Massachusetts, USA

[3]Department of Government, Dartmouth College, Hanover, New Hampshire, USA

[4]Department of Statistical Science, Duke University, Durham, North Carolina, USA

Correspondence
Georgia Papadogeorgou, Department of Statistics, University of Florida, Gainesville, FL 32611, USA.
Email: gpapadogeorgou@ufl.edu

Funding information
Air Force Office of Scientific Research, Grant/Award Number: #FA9550-14-1-0072; National Science Foundation, Grant/Award Numbers: 2124124, 2124323, 2124463

Abstract

Many causal processes have spatial and temporal dimensions. Yet the classic causal inference framework is not directly applicable when the treatment and outcome variables are generated by spatio-temporal point processes. We extend the potential outcomes framework to these settings by formulating the treatment point process as a stochastic intervention. Our causal estimands include the expected number of outcome events in a specified area under a particular stochastic treatment assignment strategy. Our methodology allows for arbitrary patterns of spatial spillover and temporal carryover effects. Using martingale theory, we show that the proposed estimator is consistent and asymptotically normal as the number of time periods increases. We propose a sensitivity analysis for the possible existence of unmeasured confounders, and extend it to the Hájek estimator. Simulation studies are conducted to examine the estimators' finite sample performance. Finally, we illustrate the proposed methods by estimating the effects of American airstrikes on insurgent violence in Iraq from February 2007 to July 2008. Our analysis suggests that increasing the average number of daily airstrikes for up to 1 month may result in more insurgent attacks. We also find some evidence that airstrikes can displace attacks from Baghdad to new locations up to 400 km away.

KEYWORDS

carryover effects, inverse probability of treatment weighting, point process, sensitivity analysis, spillover effects, stochastic intervention, unstructured interference

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 flows, and the various consequences of disease outbreaks. These applications also illustrate key methodological challenges. First, when the treatment and outcome variables are generated by spatio-temporal processes, there exists an infinite number of possible treatment and event locations at each point in time. In addition, spatial spillover and temporal carryover effects are likely to be complex and may not be well understood.

Unfortunately, the classical causal inference framework that dates back to Neyman (1923) and Fisher (1935) is not directly applicable to such settings. Indeed, standard causal inference approaches assume that the number of units that can receive the treatment is finite (e.g., Robins, 1997; Rubin, 1974). Although a small number of studies develop a continuous time causal inference framework, they do not incorporate a spatial dimension (e.g., Gill & Robins, 2001; Zhang et al., 2011). In addition, causal inference methods have been used for analysing functional magnetic resonance imaging data, which have both spatial and temporal dimensions. For example, Luo et al. (2012) apply randomisation-based inference, while Sobel and Lindquist (2014) employ structural modelling. We instead focus on data generated by different underlying processes, leading to new estimands and estimation strategies.

Specifically, we consider settings in which the treatment and outcome events are assumed to be generated by spatio-temporal point processes (Section 3). The proposed method 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 researchers have considered unstructured interference, they assume non-spatial and cross-sectional settings (see e.g., Basse & Airoldi, 2018; Sävje et al., 2019, and references therein). In addition, Aronow et al. (2019) study spatial randomised experiments in a cross-sectional setting, and under the assumption that the number of potential intervention locations is finite and their spatial coordinates are known and fixed. By contrast, our proposed spatio-temporal causal inference framework allows for *temporally and spatially unstructured interference* over an infinite number of locations.

Second, instead of separately estimating the causal effects of treatment received at each location, we consider the impacts of different *stochastic treatment assignment strategies*, defined formally as the intervention distributions over treatment point patterns. Stochastic interventions have been used to estimate effects of realistic treatment assignment strategies (Díaz Muñoz & van der Laan, 2012; Papadogeorgou et al., 2019; Young et al., 2014) and to address challenging causal inference problems including violation of the positivity assumption (Kennedy, 2019), interference (Hudgens & Halloran, 2008; Imai et al., 2021), mediation analysis (Díaz & Hejazi, 2019; Lok, 2016) and multiple treatments (Imai & Jiang, 2019). We show that this approach is also useful for causal inference with spatio-temporal treatments and outcomes.

Finally, our methodology allows for arbitrary patterns of spatial and temporal interference. As such, our estimation method neither requires the separation of units into minimally interacting sets (e.g., Tchetgen Tchetgen et al., 2017) nor relies on an outcome modelling approach that entails specifying a functional form of spillover effects based on, for example, geographic distance. Instead, we view our data as a single time series of maps, which record the locations of treatment and outcome realisations as well as the geographic coordinates of other relevant events.

Our estimation builds on the time-series causal inference approach pioneered by Bojinov and Shephard (2019).

We propose a spatially smoothed inverse probability weighting estimator that is consistent and asymptotically normal under a set of reasonable assumptions, regardless of whether the propensity scores are known, or estimated from a correctly specified model (Section 4). To do so, we establish a new central limit theorem for martingales that can be widely used for causal inference in observational, time series settings. We also show that the proposed estimator based on the estimated propensity score has a lower asymptotic variance than when the true propensity score is known. This generalises the existing theoretical result under the independently and identically distributed setting (Hirano et al., 2003) to the spatially and temporally dependent setting. Finally, to assess the potential impact of unobserved confounding, we develop a sensitivity analysis method by generalising the sensitivity analysis of Rosenbaum (2002) to our spatio-temporal context and to the Hájek estimator with standardised weights (Section 5). We conduct simulation studies to assess the finite sample performance of the proposed estimators (Section 6).

Our motivating illustration 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 during each day anywhere in Iraq as a *treatment pattern*. Instead of focusing on the causal effects of each airstrike, we estimate the effects of different *airstrike strategies*, defined formally as the distributions of airstrikes throughout Iraq (Section 7). The proposed methodology enables us to capture spatio-temporal variations in treatment effects, shedding new light on how airstrikes affect the location, distribution and intensity of insurgent violence.

Specifically, under a set of assumptions, our analysis suggests that a higher number of airstrikes, without modifying their spatial distribution, may increase 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 can shift insurgent attacks from Baghdad to Mosul and its environs. Under our assumptions, these findings suggest that airstrikes can increase insurgent attacks *and* disperse them over considerable distances. Furthermore, our analysis shows that increasing the number of airstrikes may initially reduce attacks but ultimately increase them over the long run. Our sensitivity analysis indicates, however, that these findings are somewhat sensitive to the potential existence of unmeasured confounders. Thus, further analyses are necessary in order for us to reach more definitive conclusions about the impacts of airstrikes.

The proposed methodology has a wide range of applications beyond the specific example analysed in this paper. For example, the causal effects of pandemics and crime on a host of economic and social outcomes could be evaluated using our methodology. With the advent of massive and granular datasets, we expect the need to conduct causal analysis of spatio-temporal data will only continue to grow.

2 | MOTIVATING APPLICATION: AIRSTRIKES AND INSURGENT ACTIVITIES IN IRAQ

Airstrikes have emerged as a principal tool for fighting against insurgent and terrorist organisations 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 (Lyall, 2019b), we have few rigorous studies that evaluate the impact of airstrikes on subsequent

insurgent violence. Even these 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., Dell & Querubin, 2018; Kocher et al., 2011; Lyall, 2019a; Mir & Moore, 2019).

Moreover, all existing studies have two interrelated methodological shortcomings: they carve continuous geographic space into discrete, often arbitrary, units, and they make simplifying assumptions about patterns of spatial and temporal interference. Mir and Moore (2019b), for example, argue that drone strikes in Pakistan have reduced terrorist violence. But they use a coarse estimation strategy that bins average effects of drone strikes into broad half-year increments over entire districts that cannot capture local spatial and temporal dynamics. Similarly, Rigterink (2021) draws on 443 drone strikes to estimate airstrike effects on 13 terrorist groups in Pakistan, concluding that they have mixed effects. Yet their group-month estimation strategy cannot detect spillover effects nor accurately capture the timing of insurgent responses. In short, we need a flexible methodological approach that avoids the pitfalls of binning treatment and outcome measures into too-aggregate, possibly misleading, temporal and spatial units.

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 organisations in a bid to turn the war's tide.

Aircraft were assigned to bomb targets via two channels. First, airstrikes were authorised in response to American forces coming under insurgent attack. These close air support missions represented the vast majority of airstrikes in 2007–2008. 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 drive the location of airstrikes. Aid is often provided in tandem with airstrikes to drive out insurgents, while these same insurgents often attack aid locations 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. We emphasise that we may not observe all factors used for decisions on airstrikes. We will address this limitation by developing and applying a sensitivity analysis.

Figure 1 summarises 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 2246 airstrikes during this period. Figure 1b plots the spatial locations of these airstrikes across Iraq, with spatial clustering observed around Baghdad and the neighbouring 'Sunni Triangle', 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 68,573 insurgent attacks were recorded by the US Army's CIDNE database during this time period. Finally, Figure 1d plots the

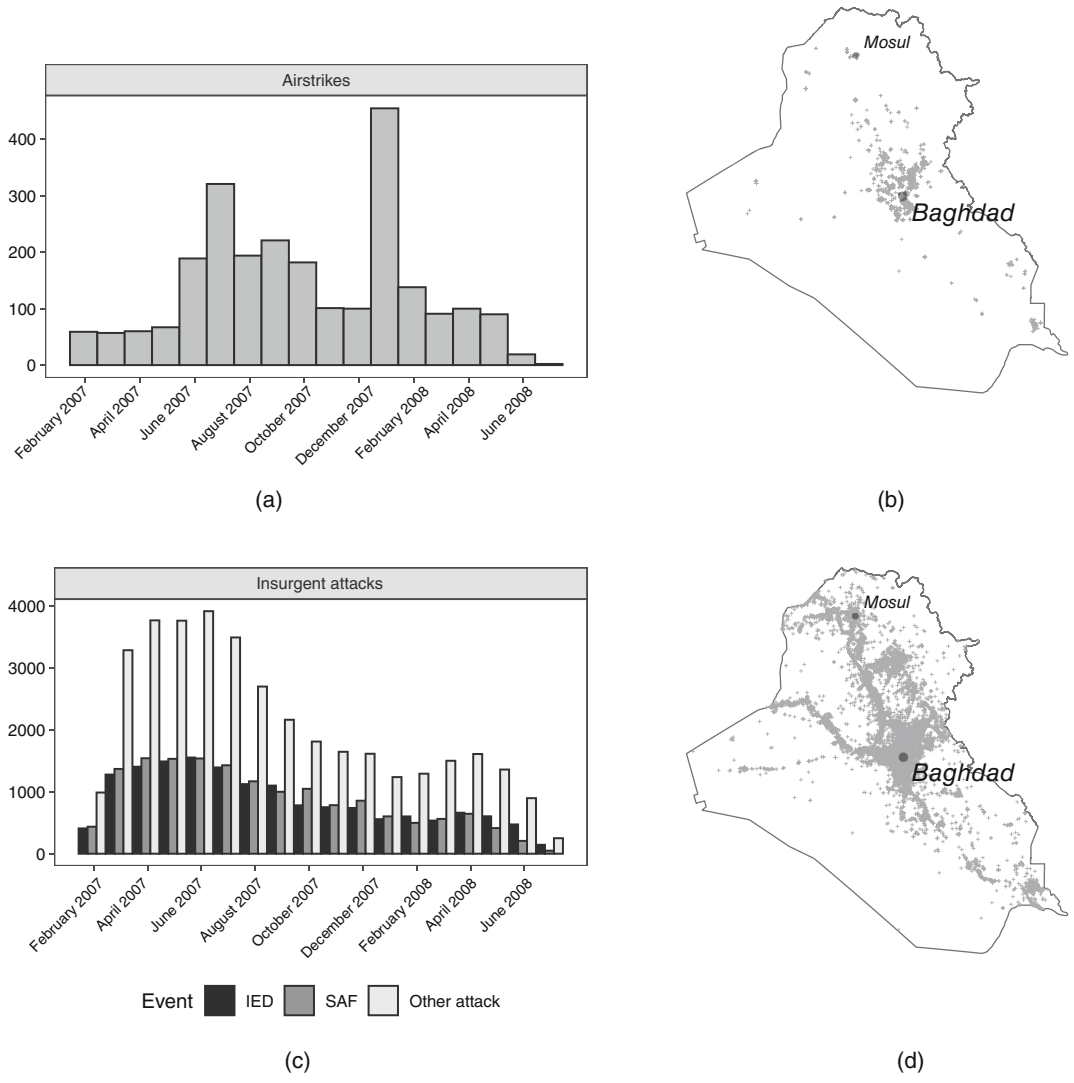


FIGURE 1 Distribution of the treatment and outcome point processes over time and space. Plots (b) and (d) show the locations of airstrikes and insurgent attacks, respectively, during the time period 23 February 2007 to 5 July 2008. Insurgent attacks are sorted into one of three categories: Improvised Explosive Devices (IEDs), Small Arms Fire (SAF), and other attacks. (a) Airstrikes over time; (b) Airstrikes over space; (c) Insurgent violence over time; (d) Insurgent violence over space

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 based on 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 location 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 in a set Ω. The set Ω is *not* assumed to be a finite grid, but it is allowed to include an infinite number of locations 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, that is, $W_t \in \mathcal{W}$ for each t. The set of *treatment-active locations*, that is, the locations 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. Finally, $\overline{\boldsymbol{W}}_t = (W_1, W_2, \dots, W_t)$ denotes the collection of treatments over the time periods $1, 2, \dots, t$.

We use w_t to represent a realisation of W_t and $\overline{\boldsymbol{w}}_t = (w_1, w_2, \dots, w_t)$ to denote the history of treatment point pattern realisations from time 1 through time t. Let $Y_t(\overline{\boldsymbol{w}}_t)$ represent the potential outcome at time $t \in \mathcal{T}$ for any given treatment sequence $\overline{\boldsymbol{w}}_t \in \mathcal{W}^t = \mathcal{W} \times \cdots \times \mathcal{W}$, depending on *all* previous treatments. Similar to the treatment, $Y_t(\overline{\boldsymbol{w}}_t)$ represents a point pattern with locations $S_{Y_t(\overline{\boldsymbol{w}}_t)}$, which are referred to as the *outcome-active locations*. In our study, $S_{Y_t(\overline{\boldsymbol{w}}_t)}$ represents the locations of insurgent attacks if the patterns of airstrikes had been $\overline{\boldsymbol{w}}_t$. Let $\overline{\mathcal{Y}}_T = \{Y_t(\overline{\boldsymbol{w}}_t) : \overline{\boldsymbol{w}}_t \in \mathcal{W}^t, t \in \mathcal{T}\}$ denote the collection of potential outcomes for all time periods and for all treatment sequences.

Among all of these potential outcomes for time t, we only observe the one corresponding to the observed treatment sequence, denoted by $Y_t = Y_t(\overline{\boldsymbol{W}}_t)$. We use $\overline{\boldsymbol{Y}}_t = \{Y_1, Y_2, \dots, Y_t\}$ to represent the collection of observed outcomes up to and including time period t. In addition, let $\boldsymbol{X}_t$ be the set of possibly time-varying confounders that are realised prior to W_t but after W_{t-1}. No assumption is necessary about the temporal ordering of any variables in $\boldsymbol{X}_t$ and Y_{t-1}. Let $\overline{\boldsymbol{\mathcal{X}}}_T = \{\boldsymbol{X}_t(\overline{\boldsymbol{w}}_{t-1}) : \overline{\boldsymbol{w}}_{t-1} \in \mathcal{W}^{t-1}, t \in \mathcal{T}\}$ be the set of potential values of $\boldsymbol{X}$ under any possible treatment history and for all time periods. We also assume that the observed covariates correspond to the covariates under the observed treatment path, $\boldsymbol{X}_t = \boldsymbol{X}_t(\overline{\boldsymbol{W}}_{t-1})$, and use $\overline{\boldsymbol{X}}_t = (\boldsymbol{X}_1, \boldsymbol{X}_2, \dots, \boldsymbol{X}_t)$ to denote the collection of observed covariates over the time periods $1, 2, \dots, t$. Finally, we use $\overline{H}_t = \{\overline{\boldsymbol{W}}_t, \overline{\boldsymbol{Y}}_t, \overline{\boldsymbol{X}}_{t+1}\}$ to denote all observed history preceding the treatment at time $t + 1$.

Since our statistical inference is based on a single time series, we consider all potential outcomes and potential values of the time-varying confounders as *fixed*, pre-treatment quantities. Then, the randomness we quantify is with respect to the assignment of treatment W_t given the complete history including all counterfactual values $\overline{H}^*_{t-1}$ where $\overline{H}^*_t = \{\overline{\boldsymbol{W}}_t, \overline{\mathcal{Y}}_T, \overline{\boldsymbol{\mathcal{X}}}_T\}$ and $\overline{H}_t \subset \overline{H}^*_t$.

3.2 | Causal estimands under stochastic interventions

A notion central to our proposed causal inference framework is *stochastic intervention*. Instead of setting a treatment variable to a fixed value, a stochastic intervention specifies the probability distribution that generates the treatment under a potentially counterfactual scenario. Although our framework accommodates a large class of intervention distributions, for concreteness, we consider intervention distributions based on Poisson point processes, which are fully

characterised by an intensity function $h : \Omega \to [0, \infty)$. 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 Poisson($h|\Omega|$) distribution, with locations distributed independently and uniformly over Ω. In general, the specification of stochastic intervention should be motivated by policy or scientific objectives. Such examples in the context of our study are given in Section 7.1.

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

$$N_{Bt}(F_h) = \int_{\mathcal{W}} N_B\left(Y_t\left(\overline{\boldsymbol{W}}_{t-1}, w_t\right)\right)\, \mathrm{d}F_h(w_t) = \int_{\mathcal{W}} \left|S_{Y_t\left(\overline{\boldsymbol{W}}_{t-1}, w_t\right)} \cap B\right|\, \mathrm{d}F_h(w_t). \tag{1}$$

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 $t-1$. The region B does not need to be defined as a connected subset of Ω, and it can be the union of potentially non-bordering sets (e.g. the suburbs of two cities).

We can extend the above estimand to an intervention taking place over M consecutive time periods. Consider an intervention, denoted by $F_{\mathbf{h}} = F_{h_1} \times \cdots \times F_{h_M}$, 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 2a. Then, we define a general estimand as

$$\begin{aligned} N_{Bt}(F_{\mathbf{h}}) &= \int_{\mathcal{W}^M} N_B\left(Y_t\left(\overline{\boldsymbol{W}}_{t-M}, w_{t-M+1}, \ldots, w_t\right)\right)\, \mathrm{d}F_{h_1}(w_t) \cdots \mathrm{d}F_{h_M}(w_{t-M+1}) \\ &= \int_{\mathcal{W}^M} \left|S_{Y_t\left(\overline{\boldsymbol{W}}_{t-M}, w_{t-M+1}, \ldots, w_t\right)} \cap B\right|\, \mathrm{d}F_{h_1}(w_t) \cdots \mathrm{d}F_{h_M}(w_{t-M+1}). \end{aligned} \tag{2}$$

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 was to follow the stochastic

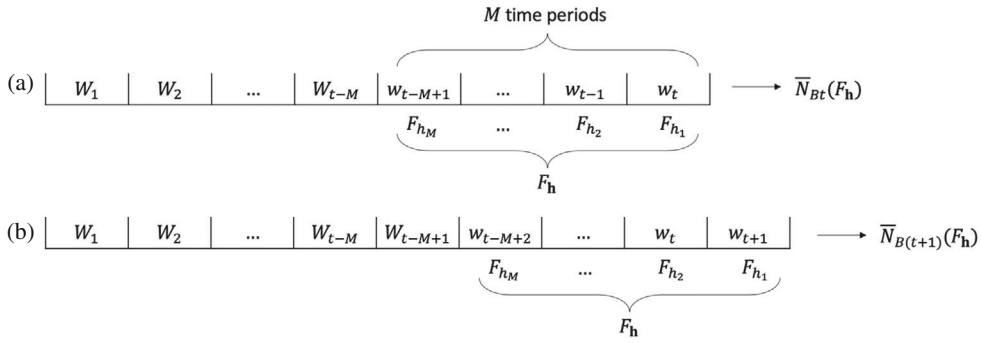


FIGURE 2 Graphical illustration of stochastic intervention over multiple time periods for time period t and $t+1$. Under intervention $F_{\mathbf{h}}$, treatments during time periods $t - M + 1, \ldots, t-1, t$ are assigned according to distributions $F_{h_M}, \ldots, F_{h_2}, F_{h_1}$.

intervention with distribution $F_{\mathbf{h}}$. Treatments during the initial $t-M$ time periods were the same as observed. A special case of $F_{\mathbf{h}}$ assumes that treatments during the M time periods are independent and identically distributed draws from the same distribution 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

$$\tau_{Bt}(F_{\mathbf{h}'}, F_{\mathbf{h}''}) = N_{Bt}(F_{\mathbf{h}''}) - N_{Bt}(F_{\mathbf{h}'}), \tag{3}$$

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 consider 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 $F_{\mathbf{h}}$. This quantity is defined as

$$N_B(F_{\mathbf{h}}) = \frac{1}{T-M+1} \sum_{t=M}^{T} N_{Bt}(F_{\mathbf{h}}). \tag{4}$$

Figure 2 shows two of the terms averaged in Equation (4), that is, $N_{Bt}(F_{\mathbf{h}})$ and $N_{B(t+1)}(F_{\mathbf{h}})$. For $N_{Bt}(F_{\mathbf{h}})$, 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 $N_{B(t+1)}(F_{\mathbf{h}})$, but intervention time periods are shifted by 1: treatments up to $t-M+1$ are set to their observed values, while treatments during time periods $t-M+2, \dots, t+1$ are drawn from $F_{\mathbf{h}}$. In Equation (4), the summation starts at $t = M$ since the quantity $N_{Bt}(F_{\mathbf{h}})$ assumes that there exist M prior time periods during which treatments are intervened on. We suppress the dependence of $N_B(F_{\mathbf{h}})$ on T for notational simplicity.

Similarly, based on $N_B(F_{\mathbf{h}})$, we define the causal effect of intervention $F_{\mathbf{h}'}$ versus $F_{\mathbf{h}''}$ as

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

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 $\overline{\boldsymbol{W}}_T$ was followed until $t-M$ with subsequent treatments $W_{t-M+1}, \dots, W_t$ arising according to $F_{\mathbf{h}'}$ versus $F_{\mathbf{h}''}$.

The effect size of a point pattern treatment would depend on M, and a greater value of M allows one to study slow-responding outcome processes. Moreover, specifying $F_{\mathbf{h}'}$ and $F_{\mathbf{h}''}$ such that they are identical except for the assignment at M time periods prior, h'_M, h''_M, yields the lagged effect of a treatment change, which resembles the lagged effects defined by Bojinov and Shephard (2019) for binary treatments and non-stochastic interventions.

The above estimands are defined while conditioning on the treatments of all previous time periods. This is important because we do not want to restrict the range of temporal carryover effects. Although the proposed estimand is generally data-dependent, the quantity becomes fixed under some settings. For example, if the potential outcomes at time t are restricted to depend at most on the latest L treatment point patterns, then the estimands for stochastic interventions that take place over $M \geq L$ time periods will no longer depend on the observed treatment path.

4 | ESTIMATION AND INFERENCE

In this section, we introduce a set of causal assumptions and the proposed estimator that combines inverse probability of treatment weighting (IPW) with kernel smoothing. We then derive its asymptotic properties. All proofs are given in Appendix B.

4.1 | The assumptions

Similar to the standard causal inference settings, variants of the unconfoundedness and overlap assumptions based on stochastic interventions are required for the proposed methodology. For simplicity, we focus on stochastic interventions with identical and independent distribution over M periods, $F_{\mathbf{h}} = F_h^M$, 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 does not depend on any, past or future, potential outcomes and potential confounders conditional on the observed history of treatments, confounders and outcomes up to time $t-1$:

$$f(W_t|\overline{\boldsymbol{W}}_{t-1}, \overline{\mathcal{Y}}_T, \overline{\boldsymbol{\mathcal{X}}}_T) = f(W_t|\overline{H}_{t-1}).$$

Assumption 1 resembles the sequential ignorability assumption in the standard longitudinal settings (Robins, 1999; Robins et al., 2000), but it is more restrictive. The assumption requires that the treatment assignment does not depend on both past and future potential values of the time-varying confounders as well as those of the outcome variable, conditional on their past observed values. In contrast, the standard sequential ignorability assumption only involves future potential outcomes.

Unfortunately, sequential ignorability would not suffice in the current setting. The reason is that we utilise data from a *single* unit measured repeatedly over many time periods to draw causal conclusions. This contrasts with the typical longitudinal settings where data are available on a large number of independent units over a short time period. Our assumption is similar to the non-anticipating treatment assumption of Bojinov and Shephard (2019) for binary non-stochastic treatments, while explicitly showing the dependence on the time-varying confounders. By requiring the treatment to be conditionally independent of the time-varying confounders, we assume that all "back-door paths" from treatment to either the outcome or the time-varying confounders are blocked (Pearl, 2000).

Next, we consider the overlap assumption, also known as positivity, in the current setting. We define the probability *density* of treatment realisation w at time t given the history, $e_t(w) = f(W_t = w|\overline{H}_{t-1})$, as the propensity score at time period t. Also, let f_h denote the probability density function of the stochastic intervention F_h. The assumption requires the ratio of propensity score over the density for the stochastic intervention, rather than the propensity score itself, is bounded away from zero.

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. This assumption enforces that the support of the intervention distribution has to be included in the support of the propensity score, and does not allow for interventions that assign positive mass to fixed treatments w.

4.2 | The propensity score for point process treatments

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 & 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. That is, $f(W_t = w|e_t(w), \overline{H}_{t-1}) = f(W_t = w|e_t(w))$ holds for all t.*

In practice, Proposition 1 allows us to empirically assess the propensity score model specification by checking the predictive power of covariates in $\overline{H}_{t-1}$ for the treatment W_t conditional on the propensity score. For example, if a covariate significantly improves prediction in a point process model of W_t after adjusting for the estimated propensity score, then the covariate is not balanced and the propensity score model is likely to be misspecified.

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

$$f(W_t|\overline{\boldsymbol{W}}_{t-1}, \overline{\mathcal{Y}}_T, \overline{\boldsymbol{\mathcal{X}}}_T) = f(W_t|e_t(W_t)).$$

Proposition 2 shows that the potentially high-dimensional sets, $\overline{H}^*_{t-1}$ and $\overline{H}_{t-1}$, can be reduced to the one dimensional propensity score $e_t(w)$ as a conditioning set sufficient for estimating the causal effects of W_t.

4.3 | The estimators

To estimate the causal estimands defined in Section 3, we propose propensity-score-based estimators that combine IPW with kernel smoothing of spatial point patterns. The estimation 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.

An alternative approach would be the direct modelling of the outcome. For example, one would model the outcome point process as a function of the past history following the g-computation in the standard longitudinal settings (Robins, 1986). However, such an approach would require an accurate specification of spatial spillover and temporal carryover effects. This is a difficult task in many applications. Instead, we focus on modelling the treatment assignment mechanism.

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

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

where $\|\cdot\|$ denotes the Euclidean norm. The summation represents the spatially smoothed version of the outcome point pattern at time period t. The product of ratios represents a weight similar to those in the marginal structural models (Robins et al., 2000), but in accordance with the stochastic intervention F_h^M: each of the M terms represents the likelihood ratio of treatment W_j 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 given in Equation (6) defines a continuous surface over Ω. The continuity of $\widehat{Y}\left(F_h^M\right)$ allows us to use it as an intensity function when estimating causal quantities. This leads to the following estimator for the expected number of outcome-active locations in any region B at time t, defined in Equation (2),

$$\widehat{N}_{Bt}\left(F_h^M\right) = \int_B \widehat{Y}_t\left(F_h^M;\omega\right)\,\mathrm{d}\omega. \tag{7}$$

We can now construct the following estimator for the temporally-expected average potential outcome defined in Equation (4),

$$\widehat{N}_B\left(F_h^M\right) = \frac{1}{T-M+1}\sum_{t=M}^{T}\widehat{N}_{Bt}\left(F_h^M\right). \tag{8}$$

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

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

An alternative estimator of $N_{Bt}\left(F_h^M\right)$ could be obtained by replacing the kernel-smoothed version of the outcome in Equation (6) with the number of observed outcome active locations in B at time t. Even though this estimator has the same asymptotic properties discussed below, the kernel-smoothing of the outcome ensures that, for a specific intervention F_h^M, once the surface in Equation (6) is calculated, it can then be used to estimate the temporally expected effects defined in Section 3 for any $B \subset \Omega$. In addition, it allows for the visualisation of the outcome surface under an intervention, making it easier to identify the areas of increased or decreased activity as illustrated in Section 7.

In the next section we establish the asymptotic properties of the proposed IPW estimators. In our simulations (Section 6) and empirical study in (Section 7), we also use the Hájek estimator, which standardises the IPW weights and replaces Equation (8) with

$$\widehat{N}_B\left(F_h^M\right) = \sum_{t=M}^{T}\widehat{N}_{Bt}\left(F_h^M\right) \Big/ \sum_{t=M}^{T}\left\{\prod_{j=t-M+1}^{t}\frac{f_h(W_j)}{e_j(W_j)}\right\}. \tag{10}$$

We find that this Hájek estimator outperforms the corresponding IPW estimators in finite samples, mirroring the existing results under other settings (e.g., Cole et al., 2021; Liu et al., 2016).

4.4 | The asymptotic properties of the proposed IPW estimators

Below, we establish the asymptotic properties of the proposed IPW estimators. Our results differ from the existing asymptotic normality results in the causal inference literature in several

ways. First, our inference is based on a single time series of point patterns that are both spatially and temporally dependent. Second, we employ a kernel-smoothed version of the outcome. Third, using martingale theory, we derive a new central limit theorem in time-dependent, observational settings. We now present the main theoretical results. All proofs are given in Appendix B.

Theorem 1 (Asymptotic normality). *Suppose that Assumptions 1 and 2 as well as the regularity conditions (Assumption A.1) hold. Then, if the bandwidth $b_T \to 0$ and as $T \to \infty$, we have that*

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

where v represents the probability limit of $(T-M+1)^{-1}\sum_{t=M}^{T} v_t$ as $T \to \infty$ with

$$v_t = \mathrm{Var}\left[\prod_{j=t-M+1}^{t} \frac{f_h(W_j)}{e_j(W_j)} N_B(Y_t) \mid \overline{H}_{t-M}^{*}\right] \quad \text{for } t \geq M.$$

The key idea of our proof is to separate the estimation error arising due to the treatment assignment W_t given the complete history $\overline{H}_{t-1}^{*}$, from the error due to spatial smoothing. Using martingale theory, we show that the former is $\sqrt{T}$-asymptotically normal, where the temporal dependence is controlled based on Assumption 1. The latter is shown to converge to zero at a rate faster than $1/\sqrt{T}$.

According to Theorem 1, the knowledge of v would enable asymptotic inference about the temporally expected potential outcome. The variance v is the converging point of $(T-M+1)^{-1}\sum_{t=M}^{T} v_t$ where v_t represents a time period-specific variance. Unfortunately, since we only observe one treatment path for each time period t, we cannot directly estimate the time-specific variances, 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 $v_t^* = \mathrm{E}\left\{\left[\widehat{N}_{Bt}\left(F_h^M\right)\right]^2 | \overline{H}_{t-M}^{*}\right\}$. For v^* such that $(T-M+1)^{-1}\sum_{t=M}^{T} v_t^* \xrightarrow{p} v^*$, we have $v \leq v^*$. Then, an α-level confidence interval for $N_B\left(F_h^M\right)$ based on the asymptotic variance bound v^*/T will achieve the nominal asymptotic coverage. Although v^* cannot be directly calculated either, there exists a consistent estimator of its value, as stated in the following lemma:

Lemma 1 (Consistent estimation of variance upper bound). *Suppose that Assumptions 1 and 2 and the regularity conditions (Assumption A.1) hold. Then, as $b_T \to 0$ and $T \to \infty$, we have*

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

In Appendix B.3 we extend the above results to the estimator $\widehat{\tau}_B\left(F_{h_1}^M, F_{h_2}^M\right)$.

So far, all of the theoretical results presented above have been established with the true propensity score $e_t(w)$. However, in practice, the propensity score is unknown and must be estimated. The next theorem shows that, when the propensity score is estimated under the correct model specification, the proposed estimator maintains its consistency and asymptotic normality. To prove this result, we extended classic M-estimation theory to multivariate martingale difference series, established a new central limit theorem for time series data, and derived the properties of the

propensity score models under the spatio-temporal settings. To our knowledge, these results are new even though related results exist under the continuous time setting (Crimaldi & Pratelli, 2005; Küchler & Sørensen, 1999). We believe that our results may be useful when studying the asymptotic properties of causal estimators in other dependent, observational settings (see Appendix B.4 more details).

Theorem 2 (Asymptotic normality using the estimated propensity score). *Suppose that Assumptions 1 and 2 as well as the regularity conditions (Assumptions A.1, A.2, A.3) hold. If the bandwidth $b_T \to 0$, then as $T \to \infty$, we have*

$$\sqrt{T}\left(\widehat{N}_B\left(F_h^M\right) - N_B\left(F_h^M\right)\right) \xrightarrow{d} \mathcal{N}(0, v^e).$$

Next, we show that using the estimated propensity scores from a correctly specified model yields more efficient estimates than using the true propensity scores. This generalises the well-known analogous result proved for the independent and identically distributed setting (e.g., Hirano et al., 2003) to the spatially and temporally dependent case (see Zeng et al., 2021, for a similar result in a different dependent setting). Thus, even with the estimated propensity score, we can make asymptotically conservative inference based on the variance upper bound derived above.

Theorem 3 (Asymptotic efficiency under the estimated propensity score). *The estimator in Equation (8) based on the estimated propensity score from a correctly specified parametric model has asymptotic variance that is no larger than the asymptotic variance of the same estimator using the known propensity score. That is, for v in Theorem 1 and v^e in Theorem 2, we have $v^e \leq v$.*

The asymptotic results presented here require the area of interest, Ω, to be fixed while the number of time periods T increases. We note that point pattern treatments and outcomes might also arise in situations where the number of time periods T is fixed, but the area under study Ω grows to include more regions. In Appendix B.5, we provide an alternative causal inference framework for point pattern treatments under this new design by extending our causal estimands, estimation and asymptotic results to the spatio-temporal setting with an area consisting of a growing number of independent regions.

5 | SENSITIVITY ANALYSIS

The validity of our estimators critically relies upon the assumption of no unmeasured confounding (Assumption 1). We develop a sensitivity analysis to address the potential violation of this key identification assumption. Specifically, we extend the sensitivity analysis pioneered by Rosenbaum (2002) to the spatio-temporal context and to the Hájek estimator with standardised weights, which we consider in our simulation and empirical studies.

Suppose there exists an unmeasured, potentially time-varying confounder U_t. We assume that the unconfoundedness assumption holds only after conditioning on the realised history of this unobserved confounder as well as $\overline{H}_{t-1}$, i.e.,

$$f(W_t|\overline{\boldsymbol{W}}_{t-1}, \overline{\boldsymbol{y}}_T, \overline{\boldsymbol{x}}_T, \overline{\boldsymbol{U}}_T) = f(W_t|\overline{H}_{t-1}, \overline{U}_t),$$

where $\overline{\mathcal{U}}_T$ represents the collection of all potential values of U_t across all time points $t = 1, 2, \ldots, T$ whereas $\overline{U}_t$ represents the history of realised but unmeasured confounder U up to time t. Note that U_t can be correlated with the observed confounders.

The existence of an unmeasured confounder invalidates the inference based on the propensity score with observed covariates alone $e_t(w)$ because the true propensity score, denoted by $e_t^*(w) = f_{W_t}(w|\overline{H}_{t-1}, \overline{U}_t)$, conditions on the history of the unmeasured confounder $\overline{U}_t$. To develop a sensitivity analysis, we assume the ratio of estimated versus true propensity scores for the realised treatment W_t is bounded by a value $\Gamma(\geq 1)$,

$$\frac{1}{\Gamma} \leq \rho_t = \frac{e_t(W_t)}{e_t^*(W_t)} \leq \Gamma.$$

A larger value of Γ allows a greater degree of violation of the unconfoundedness assumption.

In our application, we use the Hájek-version of the proposed estimator, which we find to be more stable than the IPW estimator (see Section 6). Thus, to develop a sensitivity analysis, we derive an algorithm for bounding the Hájek estimator for stochastic interventions for each fixed value of Γ (see Appendix D for the sensitivity analysis of the IPW estimator). Specifically, for all values of $\boldsymbol{\rho} = (\rho_1, \rho_2, \ldots, \rho_T) \in [\Gamma^{-1}, \Gamma]^T$, we wish to bound the following two quantities:

$$\widehat{N}_{\boldsymbol{\rho}}(F_h) = \frac{\sum_{t=1}^T \rho_t \, w_t(F_h) \, \tilde{N}_B(Y_t)}{\sum_{t=1}^T \rho_t \, w_t(F_h)}, \quad \text{and}$$

$$\widehat{\tau}_{\boldsymbol{\rho}}(F_{h_1}, F_{h_2}) = \frac{\sum_{t=1}^T \rho_t \, w_t(F_{h_2}) \, \tilde{N}_B(Y_t)}{\sum_{t=1}^T \rho_t \, w_t(F_{h_2})} - \frac{\sum_{t=1}^T \rho_t \, w_t(F_{h_1}) \, \tilde{N}_B(Y_t)}{\sum_{t=1}^T \rho_t \, w_t(F_{h_1})},$$

where

$$w_t(F_h) = \frac{f_h(W_t)}{e_t(W_t)} \quad \text{and} \quad \tilde{N}_B(Y_t) = \int_B \sum_{s \in S_{Y_t}} K_b(\|\omega - s\|) \mathrm{d}\omega.$$

Below we show how to formulate the bounding problem for $\widehat{N}_{\boldsymbol{\rho}}(F_h)$ as a linear programme, and how to use the bounds for $\widehat{N}_{\boldsymbol{\rho}}(F_h)$ to also bound the effect estimator $\hat{\tau}_{\boldsymbol{\rho}}(F_{h_1}, F_{h_2})$.

Theorem 4 (Bounding the causal quantities).

1. *The problem of maximising $\widehat{N}_{\boldsymbol{\rho}}(F_h)$ over $\boldsymbol{\rho} \in [\Gamma^{-1}, \Gamma]^T$ is equivalent to the following linear program,*

$$\begin{aligned} &\textit{maximise}_{\rho^*} \sum_{t=1}^T \rho_t^* \, w_t(F_h) \, \tilde{N}_B(Y_t) \\ &\qquad \textit{subject to} \quad \frac{\kappa}{\Gamma} \leq \rho_t^* \leq \Gamma\kappa, \ \sum \rho_t^* \, w_t(F_h) = 1, \ \textit{and} \ \kappa \geq 0, \end{aligned}$$

where $\boldsymbol{\rho} = \boldsymbol{\rho}^*/\kappa$.

2. *Suppose that* $\boldsymbol{\rho}_j^{\max}$ *and* $\boldsymbol{\rho}_j^{\min}$ *represent the values of* $\boldsymbol{\rho}$ *that maximise and minimise* $\widehat{N}_{\boldsymbol{\rho}}(F_{h_j})$, *respectively, for* $j = 1, 2$. *Then, the bounds for the causal effect are obtained as,*

$$\widehat{N}_{\boldsymbol{\rho}_2^{min}}(F_{h_2}) - \widehat{N}_{\boldsymbol{\rho}_1^{max}}(F_{h_1}) \leq \widehat{\tau}_{\boldsymbol{\rho}}(F_{h_1}, F_{h_2}) \leq \widehat{N}_{\boldsymbol{\rho}_2^{max}}(F_{h_2}) - \widehat{N}_{\boldsymbol{\rho}_1^{min}}(F_{h_1}). \tag{11}$$

The proof for bounding $\widehat{N}_{\boldsymbol{\rho}}(F_h)$ is based on the Charnes–Cooper transformation of linear fractionals (Charnes & Cooper, 1962), and the proof for bounding $\hat{\tau}_{\rho}(F_{h_1}, F_{h_2})$ is given in Appendix D. For bounding $\widehat{N}_{\boldsymbol{\rho}}(F_h)$, this proposition allows us to use a standard linear algorithm to obtain the optimal solution for $(\boldsymbol{\rho}^*, \kappa)$ and transform it back to the optimal solution $\boldsymbol{\rho}$. Then, we can use these bounds to also acquire bounds on the effect estimator. Since all bounds are wider for a greater value of Γ, the estimated effects are robust to propensity score misspecification up to the smallest value of Γ for which the interval of bounds in Equation (11) includes 0. Due to the standardisation of weights in the Hájek estimator, the bound in Equation (11) is conservative, in the sense that, if the causal estimate is shown to be robust up to some value Γ, then it is robust up to an even greater degree of propensity score model misspecification $\Gamma^* \geq \Gamma$. Similar bounds can be derived for the stochastic interventions that take place over multiple time periods (see Appendix Data for details).

The propensity score modelling in our spatio-temporal setting is much more complex with an infinite number of potential treatment locations than in the conventional cross-sectional setting. As a result, the modelling uncertainty for the propensity score is much greater. This makes it difficult to compare the scale of Γ between the spatio-temporal and conventional cross-section settings. In particular, the value of Γ is expected to be much closer to the null value of one in the spatio-temporal context.

6 | SIMULATION STUDIES

We conduct simulation studies to empirically investigate several key theoretical properties of the proposed methodology: (a) the performance of our estimator under different stochastic interventions and 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, (e) the relative efficiency of the estimator when using the true and estimated propensity scores, and (f) 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 fit Poisson process models to the simulated data.

6.1 | The study design

To construct a realistic simulation design, we base our data generating process on the observed data from our application. We consider a time series of point patterns of length $T \in \{200, 400, 500\}$. For each time series length T, 200 datasets are generated. The scenario with $T = 500$ closely resembles our observed data, which have $T = 469$.

(a) (b) (c)

FIGURE 3 Simulated data. Panel (a) shows one of the two time-invariant confounders representing the exponential decay of distance to the road network. Panel (b) shows one realisation for one of the time-varying confounders. After generating points from a non-homogeneous Poisson process, which depends on the observed airstrike density, we define the time-varying confounder as the exponential decay of distance to these points. Panel (c) shows the estimated log-density of treatment patterns, which is used as the density ϕ in the definition of stochastic interventions. (a) Time-invariant confounder $X^1(\omega)$; (b) Realisation of time-varying confounder $X_t^3(\omega)$; (c) Distribution of treatment point patterns

6.1.1 | Time-varying and time-invariant confounders

Our simulation study includes two time-invariant and two time-varying confounders. We base the first time-invariant confounder on the distance from Iraq's road network and its borders, by defining its value at location $\omega \in \Omega$ as $X^1(\omega) = \exp\{-3D_1(\omega)\} + \log(D_2(\omega))$, where $D_1(\omega)$ is the distance from ω to the closest road, and $D_2(\omega)$ is the distance to the country's border. This covariate is shown in Figure 3a. The second covariate is defined similarly as $X^2(\omega) = \exp\{-D_3(\omega)\}$ where $D_3(\omega)$ is the distance from ω to Baghdad.

We generate the time-varying confounders, $X_t^3(\omega)$ and $X_t^4(\omega)$, using the kernel-smoothed density of the observed airstrike and attack patterns. Specifically, we pool all airstrike locations across time and estimate the density of airstrike patterns, denoted by $\hat{f}(\omega)$ at location ω (shown in the right plot of Figure 6). Based on this density, we draw a point pattern from a non-homogeneous Poisson point process with intensity function $\lambda^{X^3}(\omega) = \exp\left\{\rho_0 + \rho_1 \hat{f}(\omega)\right\}$, for $\rho_0 \approx -2.7$ and $\rho_1 = 8$, and define $X^3(\omega)$ as $\exp\{-D_4(\omega)\}$, where $D_4(\omega)$ is the distance from location ω to the closest point. We generate $X^4(\omega)$ similarly based on the estimated density for insurgent attacks, and for corresponding values $\rho_0 \approx -3.2$ and $\rho_1 = 7$. Figure 3c shows one realisation of $X_t^3(\omega)$.

6.1.2 | Spatio-temporal point processes for treatment and outcome variables

For each time period $t \in \mathcal{T}$, we generate W_t from a non-homogeneous Poisson process that depends on all confounders $\boldsymbol{X}_t(\omega) = (X^1(\omega), X^2(\omega), X_t^3(\omega), X_t^4(\omega))^\top$, as well as the previous treatment and outcome realisations, W_{t-1} and Y_{t-1}. The intensity of this process is given by

$$\lambda_t^W(\omega) = \exp\left\{\alpha_0 + \boldsymbol{\alpha}_{\boldsymbol{X}}^\top \boldsymbol{X}_t(\omega) + \alpha_W W_{t-1}^*(\omega) + \alpha_Y Y_{t-1}^*(\omega)\right\}, \tag{12}$$

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 from 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

$$\lambda_t^Y(\omega) = \exp\left\{\gamma_0 + \boldsymbol{\gamma}_{\boldsymbol{X}}^\top \boldsymbol{X}_t(\omega) + \gamma_2 X_{t-1}^2(\omega) + \gamma_W W^*_{(t-3):t}(\omega) + \gamma_Y Y^*_{t-1}(\omega)\right\}, \tag{13}$$

where $W^*_{(t-3):t}(\omega) = \exp\{-2D^*_W(\omega)\}$ with $D^*_W(\omega)$ being the distance from ω to the closest points in $\bigcup_{j=t-3}^{t} S_{W_j}$. This specification imposes a lag-three dependence of the outcome on the treatment process. The model leads to an average of 5.5 treatment-active locations and 31 outcome-active locations within each time period, resembling the frequency of events in our observed data. The spatial distribution of generated treatment point patterns also resembles the observed one (compare Figure 3c to the right plot of Figure 6). The simulated and observed outcome point patterns also have similar distributions.

6.1.3 | Stochastic interventions

We consider stochastic interventions of the form F_h^M for a non-homogeneous Poisson process with intensity h, which is defined as $h(\omega) = c\phi(\omega)$ for c ranging from 3 to 8, and surface ϕ set to the density shown (in logarithm) in Figure 3c. This definition of stochastic intervention based on the treatment density aligns with the specification in our study in Section 7. We consider varying the intervention duration by setting $M \in \{1, 3, 7, 30\}$. We also examine lagged interventions over three time periods, that is, $F_{\mathbf{h}} = F_{h_3} \times F_{h_2} \times F_{h_1}$. The intervention for the first time period F_{h_3} is a Poisson process with intensity $h_3(\omega) = c\phi(\omega)$ for c ranging from 3 to 7, whereas $F_{h_2} = F_{h_1}$ is a non-homogeneous Poisson process with intensity $5\phi(\omega)$. For each stochastic intervention, we consider three regions of interest, B, of different sizes, representing the whole country, the Baghdad administrative unit, and a small area in northern Iraq which includes the town of Mosul.

6.1.4 | Approximating the true estimands

Equation (13) shows that the potential outcomes depend on the realised treatments during the last four time periods as well as the realised 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 each dataset in the following manner. For each time period t, and each $r = 1, 2, \dots, R$ repetition, we generate realisations $w_{t-M+1}^{(r)}, \dots, w_{t-1}^{(r)}, w_t^{(r)}$ from the intervention distribution $F_{\mathbf{h}}$. Based on the treatment path $(\overline{\boldsymbol{W}}_{t-M}, w_{t-M+1}^{(r)}, \dots, w_t^{(r)})$, we generate outcomes $y_{t-M+1}^{(r)}, \dots, y_t^{(r)}$ using Equation (13). This yields $S_{y_t^{(r)}}$, which contains the outcome-active locations based on one realisation 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 $N_{Bt}(F_{\mathbf{h}})$, and further averaging these over time gives an approximation of $N_B(F_{\mathbf{h}})$.

6.1.5 | Estimation

We estimate the expected number of points $N_B(F_{\mathbf{h}})$ and the effect of a change in the intervention on this quantity $\tau_B(F_{\mathbf{h}'}, F_{\mathbf{h}''})$ using the following estimators: (a) the proposed estimators defined

in Equations (8) and (9) with the true propensity scores; (b) the same proposed estimators with the estimated propensity scores based on the correctly specified model; (c) the above two estimators with the Hájek-type standardisation in (10); and (d) the unadjusted estimator based on the propensity score model using a homogeneous Poisson process with no predictor.

All estimators utilise the smoothed outcome point pattern. Spatial smoothing is performed using Gaussian kernels with SD equal to $10T^{-2/3}\delta$, which is decreasing in T, and for δ scaling the bandwidth according to the size of the geometry under study. We choose this bandwidth such that for $T = 500$ (the longest time series in our simulation scenario) the bandwidth is approximately 0.5, slightly smaller than the size of the smallest region of interest B (square with edge equal to 0.75). We discuss the choice of the bandwidth in Section 7.3.

6.1.6 | Theoretical variance and its upper bound

Theorem 1 provides the expression for the asymptotic variance of the proposed IPW estimator. We compute Monte Carlo approximations to this variance and its upper bound. 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 Equations (12) and (13), (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 Equations (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.

Their averages over time give the desired Monte Carlo approximations. We use a similar procedure to approximate the theoretical variance and variance bound of $\hat{\tau}_B(F_{\mathbf{h}'}, F_{\mathbf{h}''})$.

6.1.7 | 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 estimator is discussed in Appendix C.

6.1.8 | 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 Equation (12). 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.

6.1.9 | Relative efficiency of estimators based on the true and estimated propensity score

According to Theorem 3, the asymptotic variance of the estimator based on the true propensity score is at least as large as that of the estimator based on the estimated propensity score. We

investigate the relative magnitude of the Monte Carlo approximations of the corresponding two variances.

6.2 | Simulation results

Figure 4 presents the results for all the stochastic interventions that were considered. The top four rows show how the (true and estimated) average potential outcomes in the whole region ($B = \Omega$) change as the intensity varies under interventions F_h^M for $M \in \{1, 3, 7, 30\}$, respectively. The last row shows how the true and estimated average potential outcomes in the same region change under the three time period *lagged* 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 column) to 500 (right column).

The unadjusted estimator returned values that are too far from the truth and are not shown here. We find that the accuracy of the proposed estimator improves as the number of time periods increases. Notice that the convergence is slower for larger values of M. 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 most simulation scenarios even when T is relatively small and M is large. The IPW estimator (investigated more thoroughly in Appendix F) tends to suffer from extreme weights because the weights are multiplied over the intervention time periods as shown in Equation (6). These results indicate a deteriorating performance of the IPW estimator as the value of M increases, whereas the standardisation of weights used in the Hájek estimator appears to partially alleviate this issue. Results were comparable for the two other sets B.

Next, we compare the true theoretical variance, v/T, with the variance bound v^*/T and its consistent estimator (see Lemma 1). We assess the conservativeness of the theoretical variance bound by focusing on the proposed estimators with the true propensity score. Figure 5 shows the results for an intervention F_h^M for $M \in \{1, 3\}$, and for region $B = \Omega$. The results for the other regions are similar and hence omitted.

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, and the theoretical variance bound is at least as large as the variance. In the setting with $M = 3$, the theoretical variance bound follows the variance closely, evident by the fact that the two lines are essentially indistinguishable. We have found this to be the case in all scenarios with higher uncertainty, indicating that the theoretical variance bound is not overly conservative. Indeed, the variance bound is visibly larger than the true variance only in the low-variance scenarios of interventions over a single time period, as shown in the top row of Figure 5 (and in Appendix E.1).

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 more 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 E.1).

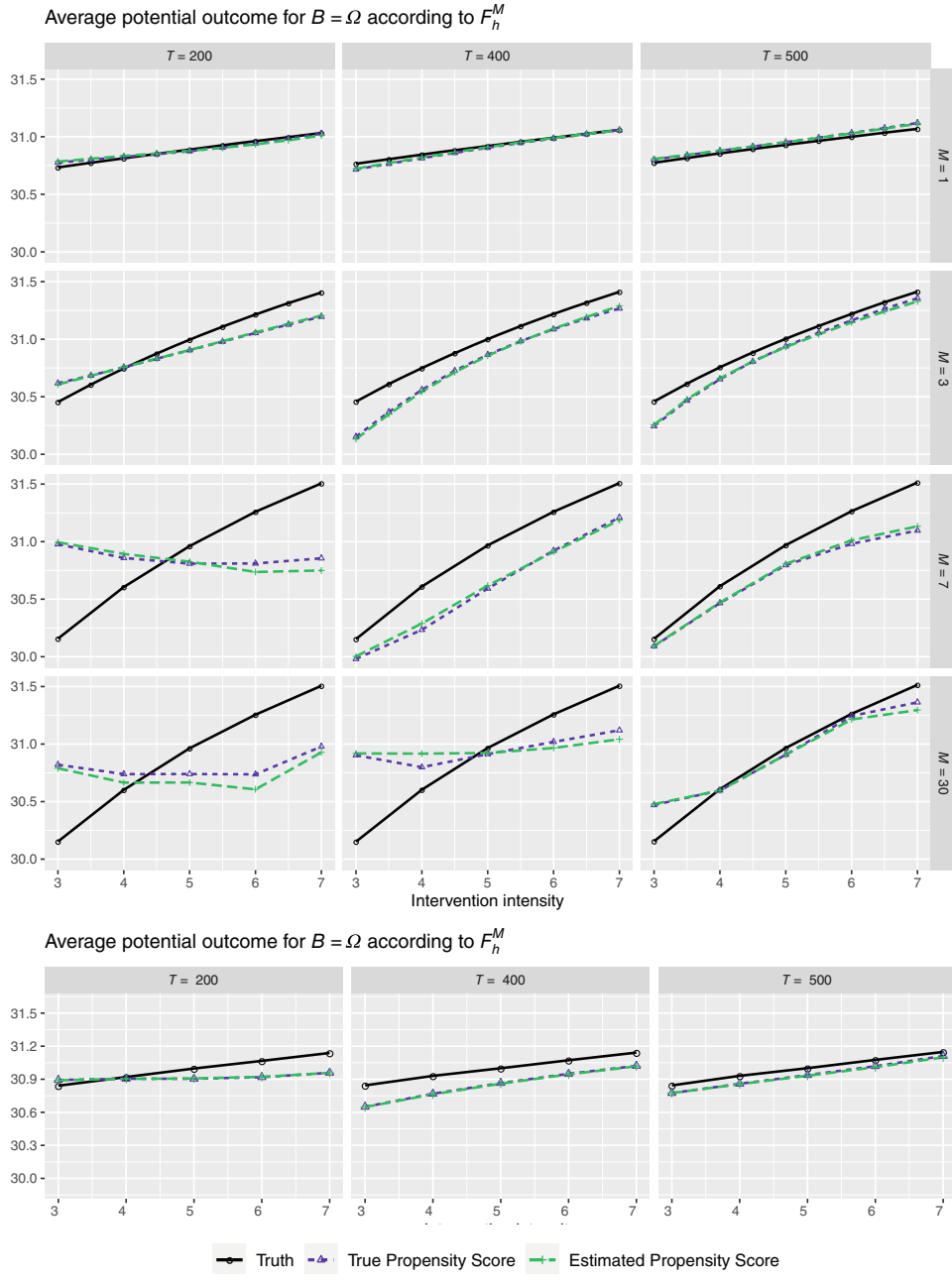


FIGURE 4 Simulation results for the average potential outcomes. In the top four rows, we present the true and estimated average potential outcomes in $B = \Omega$ under interventions F_h^M with the varying intensity (horizontal axis) and $M \in \{1, 3, 7, 30\}$ (rows), respectively. In the bottom row, we consider the average potential outcome for the lagged intervention over three time periods $F_{\mathbf{h}}$, with the varying intensity of F_{h_3} shown on the horizontal axis. The black lines with solid circles represent the truths, the Hájek estimator based on the true propensity score is shown in purple, and the Hájek estimator based on the estimated propensity score is in green.

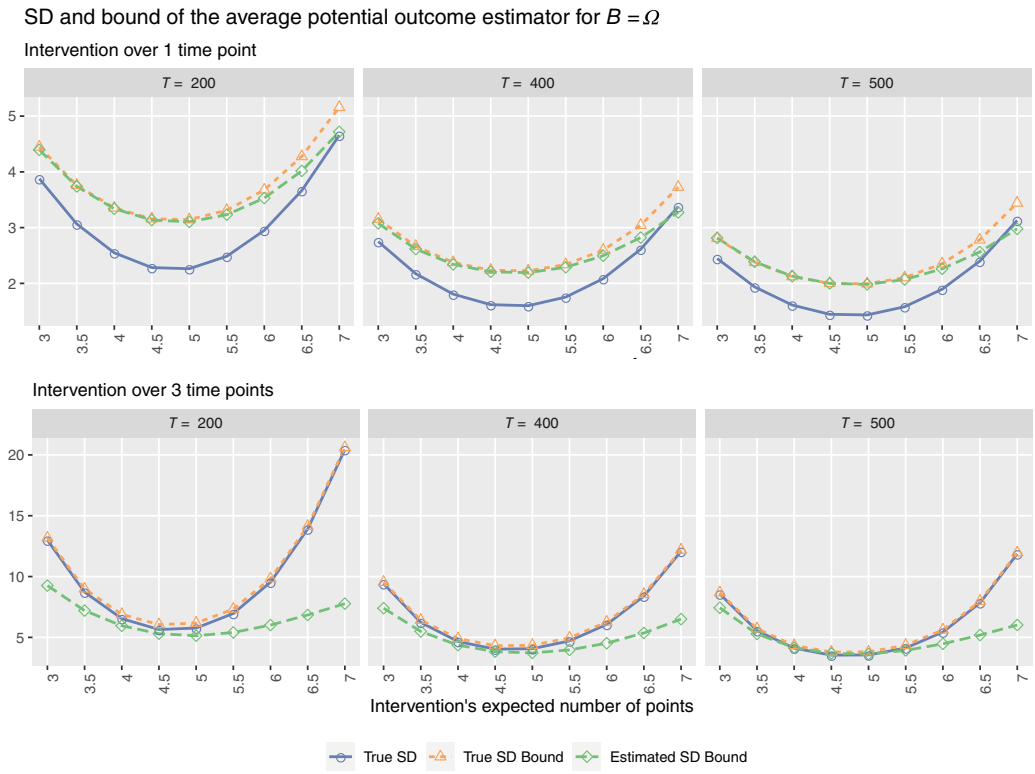


FIGURE 5 Comparison of the asymptotic SD with the true and estimated asymptotic SD bound for the average potential outcome estimator. The comparison is based on the varying number of expected points (horizontal axis) under the stochastic intervention F_h^M taking place over $M = 1$ (top row) and $M = 3$ (bottom row) time periods. The columns correspond to a simulation setting with a different time series length.

We also compare the variance of the estimator based on the true propensity score with that of the estimator based on the estimated propensity score. Table 1 shows the ratio of the Monte Carlo variances which, according to Theorem 3, should be larger than 1, asymptotically. Consistent with the above simulation results, we find that the ratio is above 1 for interventions over one and three time periods. In addition, the ratio is largest in the low uncertainty scenarios where either the number of intervention periods, M, is small, or the expected number of points is near the observed value under the intervention ($c \approx 5$). In contrast, in the high uncertainty situations with longer intervention periods, for example $M \in \{7, 30\}$, the ratio remains below 1, implying that the asymptotic approximation may not be sufficiently accurate for the sample sizes considered.

In Appendices E.2 and E.3, we also investigate the performance of the inferential procedure based on the true variance, true variance bound, and estimated variance bound, for both the IPW and Hájek estimators. The confidence interval for the IPW estimator tends to yield coverage close to its nominal level only for the interventions over a small number of time periods. In contrast, the confidence interval for the Hájek estimator has good coverage probability even for the interventions over a larger number of time periods. Partly based on these findings, we use the Hájek estimator and its associated confidence interval in our empirical application (see Section 7).

TABLE 1 Variance ratio of the proposed estimator based on the true propensity score over the proposed estimator based on the estimated propensity score

	Expected number of treatment active locations under the intervention				
	$c = 3$	$c = 4$	$c = 5$	$c = 6$	$c = 7$
$M = 1$	1.24	1.38	1.41	1.32	1.08
$M = 3$	1.09	1.18	1.24	1.14	1.06
$M = 7$	1.07	0.85	1.08	0.61	0.54
$M = 30$	0.60	0.75	0.87	0.58	0.75
Lagged	1.08	1.21	1.24	1.20	1.11

Notes: The results are based on Monte Carlo approximation with $T = 500$. The estimated propensity score is obtained from the correctly specified model. If the ratio is greater than 1, the estimated propensity score yields a more efficient estimator than the true propensity score. We consider interventions that are constant over all intervention time periods, F_h^M for $M \in \{1, 3, 7, 30\}$ (top four rows), and the lagged intervention over three time periods $F_{\mathbf{h}} = F_{h_3} \times F_{h_2} \times F_{h_1}$ (bottom row).

Finally, we evaluate the performance of the propensity score as a balancing score (Proposition 1). In Appendix E.4, we show that the p-values of the previous outcome-active locations variable (Y^*_{t-1} in Equation 12) are substantially greater in the weighted propensity score model than in the unweighted model, where the weights are equal to the inverse of the estimated propensity score. These findings are consistent with the balancing property of the propensity score.

In Appendix F we present an alternative simulation study, though all qualitative conclusions remain unchanged.

7 | EMPIRICAL ANALYSES

In this section, we present our empirical analyses of the data introduced in Section 2. We first describe the airstrike strategies of interest and then discuss the causal effect estimates obtained under those strategies.

7.1 | Airstrike strategies and causal effects of interest

We consider hypothesised stochastic interventions that generate airstrike locations based on a simple non-homogeneous Poisson point process with finite and non-atomic intensity $h : \Omega \to [0, \infty)$. 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 1 January–24 September 2006 to define 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 with an anisotropic Gaussian kernel and bandwidth specified according to Scott's rule of thumb (Scott, 1992).

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 prioritisation of certain locations for airstrikes change the spatial pattern of insurgent attacks? (3) How long does it take for

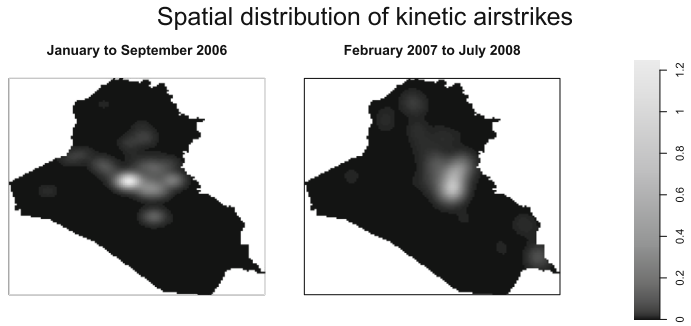


FIGURE 6 Spatial density estimate of airstrike locations during 1 January–24 September 2006 (left) and the entire study period February 2007–July 2008 (right)

the effects of change in these airstrike strategies to be realised? 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 $\int_\Omega h(\omega)\mathrm{d}\omega$ 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 monotonically increases as a function of c. In our analysis, we consider $\{1, 2, \ldots, 6\}$ as the range of c which corresponds to the expected number of airstrikes per day, in agreement with the observed data.

For the second question, we fix the expected number of airstrikes but vary their focal locations. To do this, we specify a distribution over Ω with power-density $d_\alpha(\omega) = d(\omega)^\alpha / \left(\int_\Omega d(\omega)^\alpha\right)$ 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 $\int_\Omega h_\alpha(\omega)\mathrm{d}\omega = c$, so that the overall expected number of airstrikes remains constant. Locations in s_f are increasingly prioritised under h_α for increasing α. For our analysis, we choose the centre of Baghdad to be the focal point s_f and d_α to be the normal distribution centred 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 visualisation of the spatial distributions in h_α for the different values of α is shown in Figure A15.

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 $F_{\mathbf{h}} = F_h^M$. In addition, we may also be interested in the lagged effects of airstrike strategies as in the third question. We specify lagged intervention to be the one which differs only for the M time periods ago, that is, $F_{\mathbf{h}} = F_{h_0}^{M-1} \times F_{h_1}$, 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. 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 M days later. For our analysis, we consider values of M which correspond to 1 day, 3 days, 1 week and 1 month.

Although full investigation is beyond the scope of this paper, in Appendix G.3, we briefly consider an extension to adaptive interventions over a single time period ($M = 1$), and discuss challenges when considering adaptive interventions over multiple time periods ($M > 1$).

7.2 | The specification and diagnostics of the propensity score model

Our propensity score model is a non-homogeneous Poisson point process model with intensity $\lambda_t(\omega) = \exp\{\boldsymbol{\beta}^\top \boldsymbol{X}_t(\omega)\}$ where $\boldsymbol{X}$ includes an intercept, temporal splines, and 32 spatial surfaces including all the covariates. The two main drivers of military decisions over airstrikes are the prior number and locations of observed insurgent attacks and airstrikes, which are expected to approximately satisfy unconfoundedness of Assumption 1. Our model includes the observed airstrikes and insurgent attacks during the last day, week, and month (six spatial surfaces). For example, the airstrike history of time t during the previous week is $\overline{W}^*_{t-1}(\omega) = \sum_{j=1}^{7} \sum_{s \in S_{W_{t-j}}} \exp\{-\text{dist}(s, \omega)\}$, which represents a surface on Ω with locations closer to the airstrikes in the previous week having greater values than more distant locations.

Our propensity score model also includes additional important covariates that might affect both airstrikes and insurgent attacks. We adjust for shows-of-force (i.e. simulated bombing raids designed to deter insurgents) that occurred 1 day, 1 week and 1 month before each airstrike (three spatial surfaces). Patterns of US aid spending might also affect the location and number of insurgent attacks 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 (one spatial surface). Finally, we also incorporate several time-invariant spatial covariates, including the airstrike's distance from major cities, road networks, rivers, and the population (logged, measured in 2003) of the governorate in which the airstrike took place (4 spatial surfaces). Lastly, we include separate predictors for distances from local settlements in each of the Iraqi districts to incorporate any area specific effects (18 spatial surfaces).

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 (see Figure A14 of Appendix E.4 in Data).

7.3 | 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 visualising 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. 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

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.

7.4 | Findings

Figure 7 illustrates changes in the estimated intensity surfaces for insurgent attacks (IEDs and SAFs) when increasing the expected number 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. Statistical significance of these results is shown in Appendix G.2.

The figure reveals a number of findings. First, we find no substantial change in insurgent attacks if these interventions last only for 1 or 3 days. When increasing airstrikes for a longer duration, however, a greater number of insurgent attacks are expected to occur. These changes are concentrated in the Baghdad area and the roads that connect Baghdad and the northern city of Mosul. These patterns apply to both IEDs and SAFs with slightly greater effects estimated for SAFs. Under Assumption 1, these results suggest that, far from suppressing insurgent attacks, airstrikes actually may *increase* them over time. In this setting, airstrikes can be counterproductive, failing to reduce insurgent violence while also victimising civilians. We emphasise that Assumption 1 may be violated and address this issue through our sensitivity analysis.

Under our assumptions, we find that the effect estimates for shifting the focal point of airstrikes to Baghdad for 1, 3, or 7 days are close to null. However, when the intervention change lasts for 30 days, our analysis suggests that insurgents may 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 SAFs, insurgents appear to move their attacks to the Mosul area even with the intervention of 7 days, though the effect size is smaller. In short, the effects of airstrikes may not be localised, but instead can *ripple over long distances* as insurgents respond in different parts of the country. Unlike existing approaches which focus on the effect of an intervention in the nearby area (e.g Schutte & Donnay, 2014), our approach captures this often-considerable displacement of violence.

Figure 8a shows the changes in the estimated average number of insurgent attacks *in Baghdad* as the expected number of airstrikes increases from 1 to $2, 3, \ldots, 6$ 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 or 30 days. These effects appear to be smaller when the intervention is much shorter. The patterns are similar for both IEDs and SAFs.

Figure 8b shows the change in the estimated number of IEDs and SAFs attacks in Baghdad 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 all estimated lagged effects for $M = 3$ are negative, whereas the estimated lagged effects for $M = 7$ are positive. This suggests that increasing the number of airstrikes may reduce insurgent violence in a short term while leading

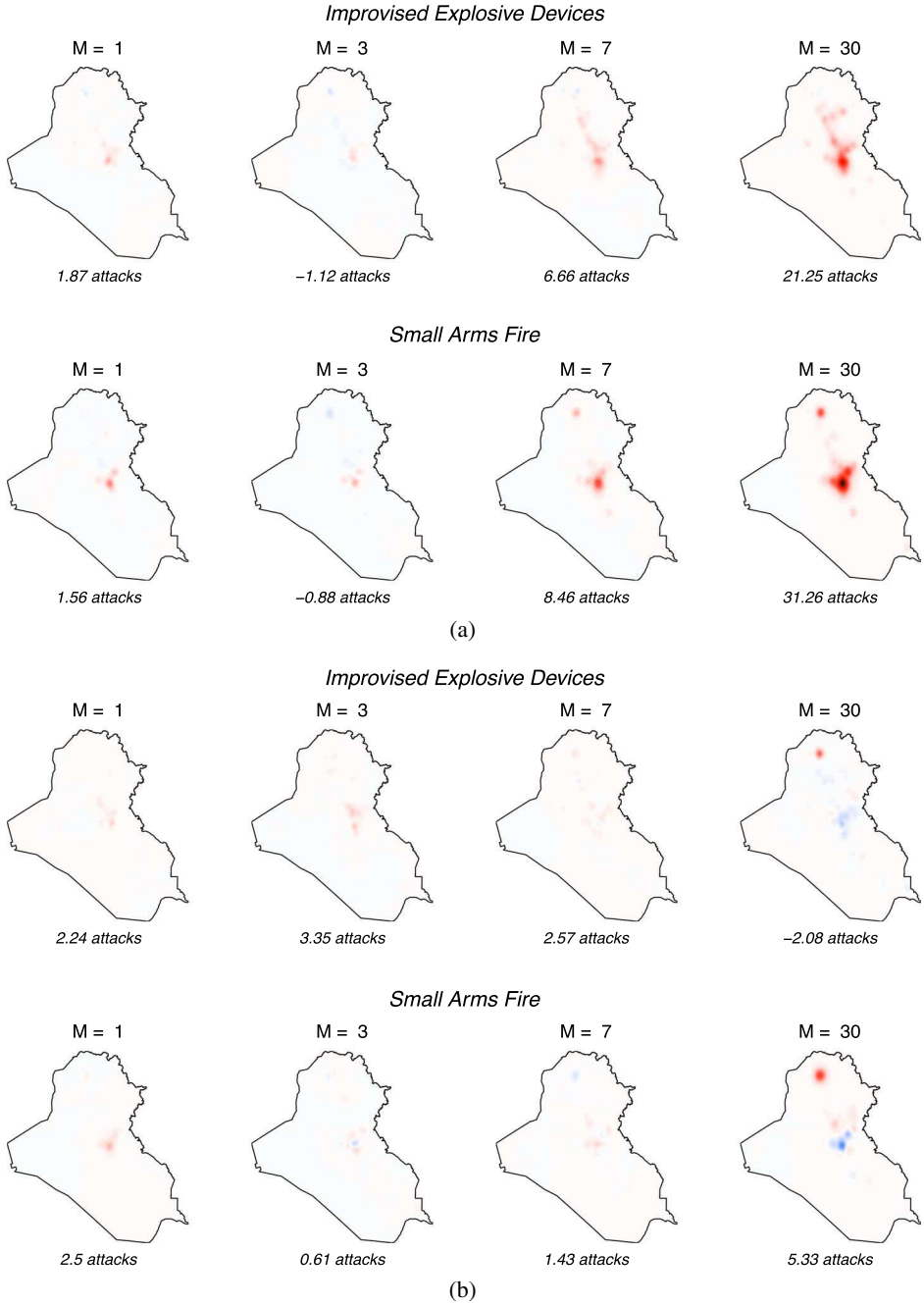
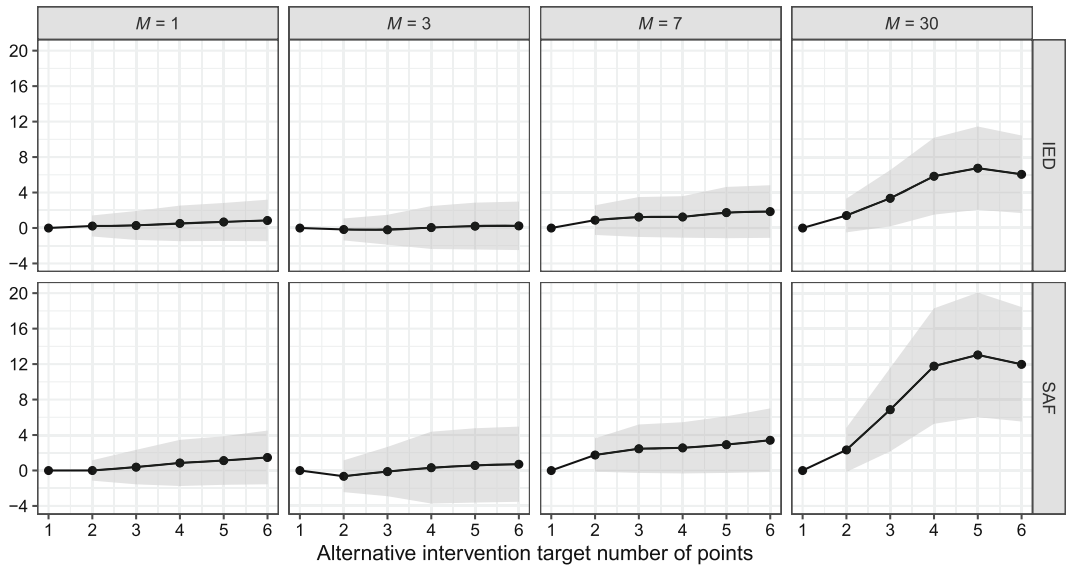
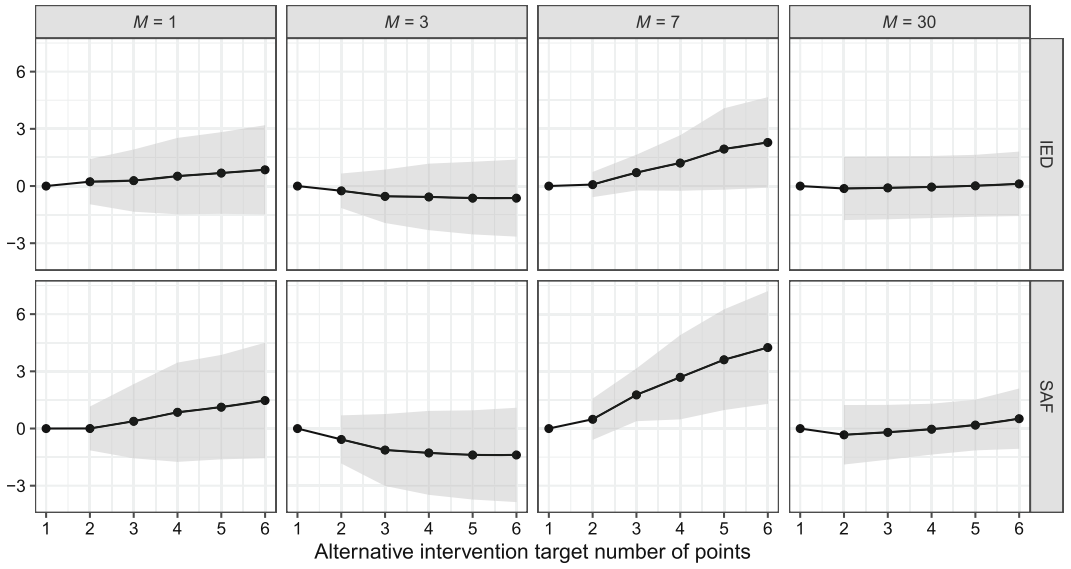


FIGURE 7 Changes in estimated intensity of insurgent attacks when increasing the expected number of airstrikes (the first two rows) and when shifting the focal point of airstrikes to Baghdad (the bottom two rows). Insurgent attacks are measured using improvised explosive devices (IEDs; the first and 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. (a) Increasing the expected number of airstrikes from 1 to 6 per day. (b) Increasing the priority of Baghdad as focal point of airstrikes from $\alpha = 0$ to $\alpha = 3$



(a)



(b)

FIGURE 8 Changes in the Estimated Number of Insurgency Attacks *in Baghdad* when Increasing the Expected Number 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. (a) Estimated effect of increasing the expected number of airstrikes for M days; (b) Estimated effect of increasing the expected number of airstrikes M days ago

to an increase in a longer term. Appendix G.2 presents the effect estimates and 95% CIs for various interventions and outcomes.

We interpret these localised effects around Baghdad as consistent with prior claims (e.g. Hashim, 2011) that Sunni insurgents were sufficiently organised to shift their attacks to new fronts in response to American airstrikes. That is, while heavy bombardment in Baghdad might suppress insurgent attacks locally, we find a net increase in overall violence as insurgent commanders displace their violence to new locations such as Mosul that are experiencing less airstrikes. This displacement effect underscores the danger of adopting too-narrow frameworks for casual estimation that miss spillover and other spatial knock-on effects.

We emphasise that the validity of our results hinges on the reliability of our causal assumptions, of which the unconfoundedness assumption (Assumption 1) is perhaps the strongest. We evaluate the robustness of our results to violations of this assumption using the sensitivity analysis framework developed in Section 5. We investigate the sensitivity of estimated effects for a change in intervention that corresponds to dosage or increased focus in Baghdad, for all values of M, for both SAF and IED outcomes, and for effects in the whole country and in Baghdad only. We find that the estimated effects are robust up to the ratio between the misspecified and the true propensity score (Γ) being bounded by 1.12. The small value of Γ indicates that our causal analysis may be sensitive to violations of the unconfoundedness assumption. As discuss before, however, this sensitivity is partially due to the inherently large uncertainty in estimating the point process intensity functions of the propensity scores from sparse data.

8 | CONCLUDING REMARKS

In this paper, we provide a framework for causal inference with spatio-temporal point process treatments and outcomes. We illustrate 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 that represents a distribution of treatments rather than 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, little is known about how the treatments in one area affect the outcomes in other areas across different time periods.

The estimands and methodology presented in this paper can be applied in a number of settings to estimate the effect of a particular stochastic intervention strategy. There are several considerations that may be useful when defining a stochastic intervention of interest. First, the choice of intervention should be guided by pressing policy questions or important academic debates where undetected spillover might frustrate traditional methods of causal inference. Second, stochastic interventions should satisfy the overlap assumption (Assumption 2). Researchers should not define a stochastic intervention that generates treatment patterns that appear to be far different from those of the observed treatment events. In our application, we achieve this by constructing the stochastic interventions based on the estimated density of point patterns obtained from the past data and the observed number of airstrikes per day.

The proposed framework can also be applied to other high-dimensional, and possibly unstructured, treatments. The standard approach to causal inference, which estimates the causal effects of fixed treatment values, does not perform well in such settings. Indeed, the sparsity of observed treatment patterns alone makes it difficult to satisfy the required overlap assumption (Imai

& Jiang, 2019). We believe that 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 interventions. In particular, it is important to consider an improved weighting method that explicitly targets covariate balance. This might be challenging in the spatiotemporal setting where the notion of covariate balance is not yet well understood. Finally, it is crucial to extend the stochastic intervention framework to adaptive strategies over multiple time periods that might be more reflective of realistic assignments.

FUNDING INFORMATION

This material is based upon work partially supported by the National Science Foundation under Grant No. 2124124, 2124463, and 2124323. 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. In addition, Imai thanks the Sloan Foundation (# 2020–13946) for financial support. The authors would also like to thank Soubhik Barari, Iavor Bojinov, Naoki Egami, Connor Jerzak, Sayar Karmakar and Neil Shephard for their constructive comments.

ORCID

Georgia Papadogeorgou https://orcid.org/0000-0002-1982-2245
Kosuke Imai https://orcid.org/0000-0002-2748-1022

REFERENCES

Abramson, I.S. (1982) On bandwidth variation in kernel estimates-a square root law. *The Annals of Statistics*, 10(4), 1217–1223.

Aronow, P. M., Samii, C. & Wang, Y. (2019) Design-based inference for spatial experiments with interference. Annual summer meeting of the society for political methodology.

Baddeley, A., Rubak, E. & Turner, R. (2015) *Spatial point patterns: methodology and applications with R*. London: Chapman & Hall/CRC Press.

Basse, G. & Airoldi, 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. & Shephard, N. (2019) Time series experiments and causal estimands: exact randomization tests and trading. *Journal of the American Statistical Association*, 114, 1665-1682

Charnes, A. & Cooper, W.W. (1962) Programming with linear fractional functionals. *Naval Research Logistics Quarterly*, 9(3-4), 181–186.

Cole, S.R., Edwards, J.K., Breskin, A. & Hudgens, M.G. (2021) Comparing parametric, nonparametric, and semiparametric estimators: the Weibull trials. *American Journal of Epidemiology*, 190(8), 1643–1651.

Crimaldi, I. & Pratelli, L. (2005) Convergence results for multivariate martingales. *Stochastic Processes and Their Applications*, 115(4), 571–577.

Dell, M. & Querubin, P. (2018) Nation building through foreign intervention: evidence from discontinuities in military strategies. *Quarterly Journal of Economics*, 133(2), 701–764.

Díaz, I. & Hejazi, N. (2019) Causal mediation analysis for stochastic interventions. *arXiv preprint arXiv:1901.02776*.

Díaz Muñoz, I. & 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*. London: Oliver and Boyd.

Gill, R.D. & Robins, J.M. (2001) Causal inference for longitudinal data: the continuous case. *Annals of Statistics*, 29(6), 1785–1811.

Hashim, A. (2011) *Insurgency and counter-insurgency in IRAQ*. Ithaca: Cornell University Press.

Hirano, K., Imbens, G.W. & Ridder, G. (2003) Efficient estimation of average treatment effects using the estimated propensity score. *Econometrica*, 71(4), 1161–1189.

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

Imai, K. & Jiang, Z. (2019) Comment on The blessings of multiple causes by Wang and Blei. *Journal of the American Statistical Association*, 114(528), 1605–1610.

Imai, K., Jiang, Z. & Malai, A. (2021) Causal inference with interference and noncompliance in two-stage randomized experiments. *Journal of the American Statistical Association*, 116(534), 632–644.

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. & Kalyvas, S. (2011) Aerial bombing and counterinsurgency in the Vietnam War. *American Journal of Political Science*, 55(2), 201–218.

Küchler, U. & Sørensen, M. (1999) A note on limit theorems for multivariate martingales. *Bernoulli*, 5(3), 483–493.

Liu, L., Hudgens, M.G. & Becker-Dreps, S. (2016) On inverse probability-weighted estimators in the presence of interference. *Biometrika*, 103(4), 829–842.

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. & 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. & 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.

Papadogeorgou, G., Mealli, F. & Zigler, C.M. (2019) Causal inference with interfering units for cluster and population level treatment allocation programs. *Biometrics*, 75(3), 778–787.

Pearl, J. (2000) *Causality: models, reasoning, and inference*. Cambridge: Cambridge University Press.

Rigterink, A. (2021) The wane of command: evidence on drone strikes and control within terrorist organizations. *American Political Science Review*, 115(1), 31–50.

Robins, J. (1986) A new approach to causal inference in mortality studies with a sustained exposure period-application to control of the healthy worker survivor effect. *Mathematical Modelling*, 7(9-12), 1393–1512.

Robins, J. M. (1997) *Latent variable modeling and applications to causality*, Vol. 120 *Lecture Notes in Statistics*, Causal Inference from Complex Longitudinal Data; 1997:69–117.

Robins, J.M. (1999) Association, causation, and marginal structural models. *Synthese*, 121(1-2):151–179.

Robins, J.M., Hernán, M.A. & Brumback, B. (2000) Marginal structural models and causal inference in epidemiology. *Epidemiology*, 11(5), 550–560.

Rosenbaum, P.R. (2002) *Observational studies*. New York: Springer.

Rosenbaum, P.R. & 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*.

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

Scott, D.W. (1992) *Multivariate density estimation*. New York: Wiley.

Sobel, M.E. & 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*.

Young, J.G., Hernán, M.A. & Robins, J.M. (2014) Identification, estimation and approximation of risk under interventions that depend on the natural value of treatment using observational data. *Epidemiologic Methods*, 3(1), 1–19.

Zeng, S., Li, F., Hu, L. Li, F. (2021) Propensity score weighting analysis for survival outcomes using pseudo observations. *arXiv preprint arXiv:2103.00605*.

Zhang, M., Joffe, M.M. & 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.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Papadogeorgou, G., Imai, K., Lyall, J. & Li, F. (2022) Causal inference with spatio-temporal data: Estimating the effects of airstrikes on insurgent violence in Iraq. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 1–31. Available from: https://doi.org/10.1111/rssb.12548