CESifo Economic Studies, 2018, 127–149
doi: 10.1093/cesifo/ify004
Advance Access Publication Date: 16 March 2018
Original article

Geographic Natural Experiments with Interference: The Effect of All-Mail Voting on Turnout in Colorado

Luke Keele* and Rocío Titiunik†

*McCourt School of Public Policy and Government Department, Georgetown University, 304 Old North, 37th & O St, Washington, DC, 20057, USA. e-mail: lk681@georgetown.edu and †Department of Political Science, University of Michigan, 5700 Haven Hall, 505 South State St, Ann Arbor, MI 48109, USA. e-mail: titiunik@umich.edu

Abstract

We analyze a geographic natural experiment during the 2010 Colorado primary election in the USA, when counties in the state of Colorado had the option to have an all-mail election or retain traditional in-person voting on Election Day. The town of Basalt, in the southwestern part of the state, is split in half by two counties that chose different modes of voting. Our research design compares these two counties to understand whether turnout levels were altered by all-mail elections. Our analysis considers the possibility that social interactions may lead to spillover effects—a situation in which one unit's outcome may be affected by the treatment received by other units. In our application, treated and control voters lived in very close proximity and spillovers are possible. Using the potential outcomes framework, we consider different estimands under the assumption that interference occurs only when treated individuals are in close geographic proximity to a sufficiently high number of control individuals. Under our assumptions, our empirical analysis suggests that all-mail voting decreased turnout in Colorado, and shows no evidence of spatial interference between voters. (JEL codes: C18, C99)

Key words: econometric and statistical methods, spatial models

1. Introduction

In recent decades, the USA has witnessed an increase in methods of voting that differ from traditional in-person voting on Election Day. These reforms, commonly referred to as 'convenience voting', include in-person early voting (where voters may cast a vote in person before election day), no-excuse absentee voting (where voters may apply for an absentee ballot without providing a reason for doing so), and all-mail voting (where voting by mail is mandatory)—see Gronke et al. (2008) for a review. These policies are often implemented with the goal of reducing the costs of voting, which is in turn expected to increase voter participation.

Among the different convenience voting policies that have been adopted, all-mail voting is the most drastic, since, under this policy, in-person precinct voting is eliminated and there are no polling places; instead, citizens receive a ballot in the mail several weeks in advance of Election Day and then return it by mail to the election administration office. Since all-mail voting is the only convenience method that eliminates precinct-place voting, its effects on turnout could be different from the effects of other types of convenience voting. While all-mail voting may be more convenient, the move to all-mail elections reduces the social aspect of voting, which can be a key motivator for political participation (Gerber et al. 2008). All-mail voting also removes the possibility of using Election Day as a focal point for mobilization efforts by political parties. Combined, these factors might result in fewer voters casting a vote.

Given the far-reaching nature of all-mail voting reforms, scholars have been interested in studying whether they affect voter turnout. As various states have implemented vote-by-mail systems either on a statewide or more local basis, a number of studies have attempted to estimate whether this mode of voting increases or decreases turnout. Much of the focus has been on the state of Oregon, where polling-place voting was gradually eliminated during the 1990s, and since 1998 all statewide primary and general elections are conducted by mail only. Several studies have concluded that Oregon's all-mail voting reform increased turnout, though the estimated magnitude of the change varies considerably from study to study (Southwell and Burchett 2000; Karp and Banducci 2000; Berinsky et al. 2001; Richey 2008; Gronke and Miller 2012). Some additional evidence on this question comes from other states. Kousser and Mullin (2007) and Bergman and Yates (2011) study California, where county election officials can assign voters to all-mail voting precincts in low-population areas, and find that turnout appears to be lower under an all-mail voting system. Gerber et al. (2013) study the large-scale move from polling-place to all-mail elections in the state of Washington, and they find that all-mail voting increases turnout by 2–4 percentage points.

We examine this question using a geographic natural experiment in Colorado, where in 2010 counties were given the choice to require that votes be cast by mail during the primary election. Counties that adopted all-mail elections removed other alternative methods of voting, while counties that did not adopt all-mail voting still offered traditional polling-place voting on Election Day (and also allowed by-mail no-excuse absentee voting). In general, given that voter administration is conducted by county governments, counties may chose their mode of voting to try to accomplish their specific voter turnout goals. This type of strategic decision-making may complicate naive statistical inferences that simply compare all-mail counties to in-person counties. In an attempt to minimize these complications, our study focuses on voters in Basalt, a town that is split by the border between Eagle county, which adopted all-mail election voting, and Pitkin county, which retained in-person voting. Our research design focuses on a narrow area around the boundary between both counties, and makes the assumption that, within this small area, voters in the town of Basalt are split in a haphazard fashion between Eagle and Pitkin counties, after conditioning on covariates. Based on this research strategy, we draw inferences about the effects of all-mail elections on voter turnout.

A key element of our research design is our focus on a small geographic area around the boundary that separates both counties, since citizens who reside close to the county border on either side share important predetermined characteristics that may be related to voter turnout decisions. This focus on treated and control voters who reside near each other, while increasing comparability, may also increase the likelihood that treated voters interact with control voters in a way that affects their outcomes. The presence of interference between voters would undermine the interpretation of our estimates as the average treatment effect of all-mail voting. In general, research designs based on comparisons of units who reside in close geographic proximity to each other are more likely to suffer from spatial forms of interference between units, and will face a trade-off between increasing comparability and reducing interference (Keele et al. 2017). However, the patterns of interference between units may be more general than those induced by residential proximity. For example, workers could be influenced by colleagues whose residence is geographically far from their own, but with whom they interact at the workplace. Even more drastically, interference could arise between units who are never geographically or physically close in any capacity, in particular via social media interactions. The framework we use below could be applied to these more general forms of interference, replacing our notion of residential proximity with a notion of proximity in a social network, assuming that the latter is known—for an example of a study that considers both types of proximity, see Verdery et al. (2012).

One goal in our study is to understand whether interference alters our inferences about the effect of all-mail voting. Our empirical analysis builds on prior work that employs the potential outcomes framework to study interference between units. Sobel (2006) characterizes several estimands of interest under interference, and shows that the usual difference-in-means estimator in a completely randomized experiment is no longer unbiased for the average treatment effect. Hudgens and Halloran (2008) consider a two-stage randomization, in which interference occurs within but not between groups, and define direct and indirect causal effects that consider how the outcomes of one unit change as the treatment assignment of all other units stays constant or changes. Hierarchical models in which interference occurs within but not between groups are also considered, among others, by Tchetgen Tchetgen and VanderWeele (2012), Vanderweele (2008), and VanderWeele et al. (2013). Aronow and Samii (2017) consider the estimation of average causal effects under general forms of known interference. Rosenbaum (2007) and Bowers et al. (2013) consider hypothesis testing under interference in a Fisherian framework. Gerber and Green (2012) consider the problem of spatial interference in randomized experiments, and Sinclair et al. (2012) design a multilevel voter-mobilization experiment to detect spillovers within and between households.

Methodologically, our approach is most similar to the setting in Hong and Raudenbush (2006) and Verbitsky-Savitz and Raudenbush (2012). Hong and Raudenbush (2006) study the effect of retaining low-achieving children in kindergarten versus promoting them to 1st grade, and model interference effects by means of a scalar function of the treatment assignment vector within each school. Verbitsky-Savitz and Raudenbush (2012) apply the ideas in Hong and Raudenbush (2006) to a spatial setting to study the effect of a community policing program on neighborhoods' crime rates in Chicago. They assume that the potential outcome of a given unit depends on the other units' potential outcomes via a scalar function that contains the proportion of contiguous units.

In our analysis of the effects of all-mail voting under spatial interference, we also impose the assumption that interference is a scalar function of the treatment assignment vector. In particular, we assume that interference is a function of geographic proximity to a sufficiently dense area of voters of the opposite treatment status, which vastly reduces the number of potential outcomes for every unit and leads to two estimands of interest.

This approach is similar to the approach in Verbitsky-Savitz and Raudenbush (2012), who model interference as a function of the proportion of contiguous units, although we impose an additional restriction. Our setup treats the geographic locations of the voters in our sample as random, and the boundary between treated and control areas as fixed—an approach that is particularly well suited to geographical natural experiments that focus on a narrow band around a boundary, and differs from other approaches that allow for spatial interference under the assumption that geographic locations are given. Our function of interference can be modified to reflect particular patterns of geographic spillovers.

The remainder of the article is organized as follows. In Section 2 we describe the Colorado application in more detail. In Section 3 we present our notation and describe our causal estimand under the assumption of no interference. In Section 4 we allow for geographic interference, and explore some further issues in Section 5. In Section 6, we use this framework to estimate the treatment effect of Colorado's all-mail voting on turnout, first ignoring interference between units, and then allowing for interference based on residential proximity (Section 6.1). In this section, we also explore sensitivity to differential registration (Section 6.2). We offer concluding remarks in Section 7.

2. All-Mail Voting in 2010 Colorado Primary

In recent years, the state of Colorado has implemented several reforms aimed at making voting more convenient. Starting in 2008, voters could choose to be placed on a permanent vote-by-mail list. For the 2010 primary, the Secretary of State allowed each county to choose whether to hold either all-mail elections, use voter centers, or hold traditional in-person voting at precincts. Figure 1 contains a map showing the mode of election selected by each county. While urban areas generally selected all-mail elections, many rural counties chose to use in-person voting.

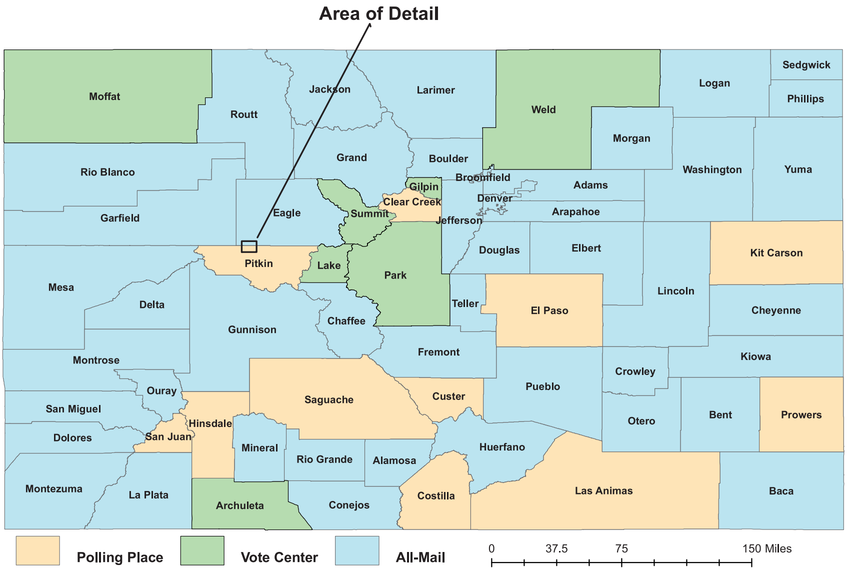

Figure 1. County map of Colorado with model of voting and location of Basalt highlighted.

The different modes of voting chosen by different counties allow us to study how election mode affects voter turnout. However, given that counties are responsible for election administration and were able to select their preferred mode of election in the 2010 primary, comparisons across counties might be invalidated by unobserved confounding or heterogeneity. To minimize this concern, we looked for some location where a town or city is split by a county border, where one county uses all-mail voting, while the other county uses in-person voting. We found that one town in the southwestern part of the state, Basalt, was split exactly in this fashion.

Figure 1 shows the location of the town of Basalt. According to the 2010 census, Basalt has a total population of 3857. The population is largely White, and about 20% of it identifies as Hispanic. The town is close to the resort city of Aspen, and using property sale records, we found that the median house price in 2010 was over $600,000. Figure 2 contains a map showing the town in greater detail. The central part of Basalt is split by the county border which defines mode of election, and this part of the town contains the main shopping district, residential areas, and schools. While the county border splits the town, the entire area is within the same school district. Moreover, all residents of Basalt attend the same set of public schools which are located within the central part of the town. Although property taxes in Colorado have a county component, property taxes are based on five different tax zones with school district contributing the most to the overall property tax burden.

Primary elections often hold little interest for voters, since primary races are often uncompetitive. The 2010 Colorado primary, however, had three high-profile elections on the ballot. In the Republican gubernatorial primary, a Tea Party insurgent beat Scott McInnis, a six-term U.S. representative, after it came to light that McInnis plagiarized a water study

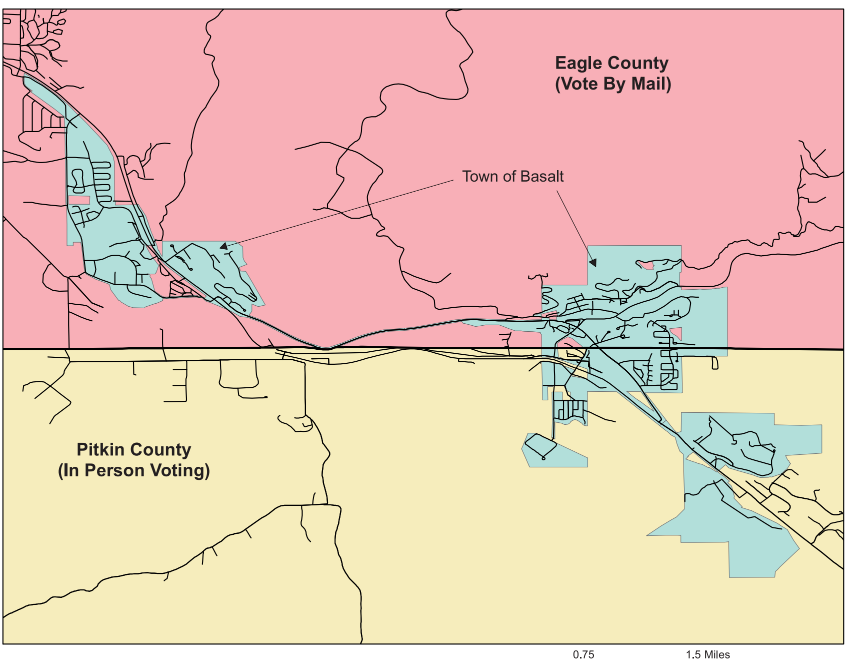

Figure 2. The county discontinuity: Basalt split by county boundary.

he was paid to conduct. In the Democratic US Senate primary, the candidate endorsed by then-President Obama narrowly beat a more liberal candidate endorsed by former president Bill Clinton. In the Republican US primary Ken Buck, a Tea Party candidate, beat Jane Norton—the candidate endorsed by the Colorado Republican party establishment. Results from the primary received national coverage and were featured on the front page of the *New York Times*.

3. Estimation and Inference without Interference

We adopt the potential outcomes framework for causal inference (see Holland 1986; Rubin 2005) assuming first that voters do not interfere with one another. We generalize the framework to allow for interference in Section 4.

We conduct our analysis with individual-level voter data. We use the binary variable $D_i \in [0, 1]$ to denote treatment status for resident i, with $D_i = 1$ if i resides in Eagle County and is assigned to an all-mail election, and $D_i = 0$ if i resides in Pitkin county and may vote in-person on Election Day. Each resident has several potential outcomes, only one of which is realized by the assignment of treatment. There are also k predetermined covariates for each resident, which we denote by $\boldsymbol{X}_i$. The observed data are $\{Y_i, D_i, \boldsymbol{X}_i\}_{i=1}^n$, which we assume is an i.i.d. random sample from a larger population. We collect all treatment indicators in the n-vector $\boldsymbol{D}$, and let $Y_i(\boldsymbol{D})$ be the potential outcome of resident i. We denote the observed outcome by $Y_i \equiv Y_i(\boldsymbol{d})$, where $\boldsymbol{d}$ is the realized treatment assignment vector. In general, if we let the treatment status of every resident affect the potential outcome of every other resident, every i will have one distinct potential outcome for every value that the treatment vector $\boldsymbol{D}$ might take, which is 2^n. We start by entirely simplifying this problem and assuming that there is no interference between residents, as formalized in the following assumption.

Assumption 1 (No Interference). The potential outcome of each unit depends only on the treatment received by that unit and not on the treatment assigned to any other unit: for all $\dot{\boldsymbol{D}} \neq \ddot{\boldsymbol{D}},\ Y_i(\dot{\boldsymbol{D}}) = Y_i(\ddot{\boldsymbol{D}})$ if $\dot{D}_i = \ddot{D}_i$, for $i = 1, 2, \ldots, n$.

Under Assumption 1, we can write $Y_i(\boldsymbol{D}) = Y_i(D_i)$, since i's potential outcome only depends on the treatment received by i. In this case, the observed outcome simplifies to $Y_i = Y_i(1)D_i + Y_i(0)(1 - D_i)$. The quantity $\tau_i = Y_i(1) - Y_i(0)$ captures the effect of all-mail voting for the ith voter. Our interest is estimating the average treatment effect on the treated (ATT), $\tau = \mathbb{E}[Y_i(1) - Y_i(0)|D_i = 1]$, which is only defined under the assumption of no interference.

3.1 A geographic identification strategy

We estimate the effects of all-mail voting on voter turnout using a geographic natural experiment. Under this identification strategy, a geographic or administrative boundary splits units into two adjacent areas, one of which receives a treatment, $\mathcal{A}^t$, and the other of which receives control, $\mathcal{A}^c$, and analysts make the case that the assignment of units into treated and control areas occurs in an as-if random fashion (Keele and Titiunik 2015, 2016). Researchers make comparisons between units in the treated and control areas to infer the effect of the treatment on an outcome of interest, relying on the spatial proximity of each unit to the border between $\mathcal{A}^c$ and $\mathcal{A}^t$, and on the fact that the treatment changes abruptly along this boundary—that is, $D_i = 1$ if unit i is located in $\mathcal{A}^t$, and $D_i = 0$ if i is located in $\mathcal{A}^c$. Applying this strategy to our application, we assume that around the county border that divides the town of Basalt into all-mail and in-person voting regimes, individuals choose their residence on either side of the county boundary on an as-if random fashion, possibly after conditioning on predetermined covariates.

In essence, the assumptions behind a geographic natural experiment require that the placement of each unit on either side of the geographic boundary between $\mathcal{A}^c$ and $\mathcal{A}^t$ be as-if random or, in other words, that units cannot precisely sort or self-select to one side of the boundary based on unobserved factors that are also correlated with the outcomes of interest. A consequence of this assumption is that observable predetermined covariates should be similar in expectation within some narrow area around the border of interest. A weaker assumption is that treatment assignment is as-if randomized for those who live near the border, after conditioning on a set of observable covariates (Keele et al. 2015). Given the need to condition on covariates, such designs have been characterized as geographic-quasi experiments (GQEs) (Galiani et al. 2017; Keele et al. 2017). Since we are interested in the ATT, we adopt a version of this assumption that only restricts the average potential outcome under control: we assume that there exists a small neighborhood around the boundary that separates both areas where the average potential outcome under control is mean independent of the treatment given the covariates. We state it formally below.

Assumption 2 (Conditional Mean Independence in Local Neighborhood). For all units that reside in a narrow band around the boundary that separates $\mathcal{A}^c$ and $\mathcal{A}^t$, $\mathbb{E}[Y_i(0)|\boldsymbol{X}_i, D_i = 1] = \mathbb{E}[Y_i(0)|\boldsymbol{X}_i, D_i = 0]$.

Note that Assumption 2 implicitly invokes Assumption 1, which reduces the set of potential outcomes to $Y_i(1)$ and $Y_i(0)$. Moreover, since our boundary of interest is simultaneously the boundary of multiple institutional, administrative, or political units, and we wish to make inferences about the effect of only one of these treatments, we must assume that the treatment of interest is the only treatment that affects potential outcomes, i.e. that there are no compound treatments (Keele and Titiunik 2015). In particular, in our application we must assume that no other county-level factor affects voter turnout other than the administration of voting.

Under these assumptions, the ATT is identified by

$$\begin{aligned}\tau &= \mathbb{E}[Y_i(1) - Y_i(0)|D_i = 1]\\ &= \mathbb{E}[\tau(\boldsymbol{X}_i)|D_i = 1]\\ &= \mathbb{E}[\mathbb{E}(Y_i|\boldsymbol{X}_i, D_i = 1) - \mathbb{E}(Y_i|\boldsymbol{X}_i, D_i = 0)|D_i = 1],\end{aligned}$$

where we have defined $\tau(\boldsymbol{X}_i) \equiv \mathbb{E}(Y_i(1) - Y_i(0)|D_i = 1, \boldsymbol{X}_i)$, the ATT conditional on covariates.

We estimate effects of interest using least squares methods. To motivate our estimation strategy, we first assume that treatments were randomly assigned. We can express the observed outcome as:

$$Y_i = \mu + \tau \cdot D_i + u_i, \tag{1}$$

where

$$\mu = \mathbb{E}[Y_i(0)|D_i = 0]$$
$$u_i = (1 - D_i) \cdot \{Y_i(0) - \mathbb{E}[Y_i(0)|D_i = 0]\} + D_i \cdot \{Y_i(1) - \mathbb{E}[Y_i(1)|D_i = 1]\},$$

and τ is the ATT defined above. Random assignment would imply $\mathbb{E}[Y_i(0)|D_i = 0] = \mathbb{E}[Y_i(0)|D_i = 1]$, which in turn would lead to $\mathbb{E}[u_i \cdot D_i] = \mathbb{E}[u_i|D_i = 1] = 0$. Thus, under random assignment, the coefficients μ and τ can be consistently estimated using least squares methods, simply regressing the voter turnout binary outcome on an indicator variable for treatment.

In our application, however, we do not believe that the condition $\mathbb{E}[Y_i(0)|D_i = 0] = \mathbb{E}[Y_i(0)|D_i = 1]$ is plausible. Instead, we assume Assumption 2 holds, where mean independence holds conditional on $\boldsymbol{X}_i$. For simplicity, to be able to apply least-squares methods to this case, we assume that we can condition on covariates in a linear fashion. In this case, we can express the observed outcome as:

$$\begin{aligned} Y_i &= \mathbb{E}[Y_i(0)|D_i = 0, \boldsymbol{X}_i] + \tau \cdot D_i + \tilde{u}_i \\ &= \boldsymbol{X}_i{}'\beta + \tau \cdot D_i + \tilde{u}_i \end{aligned}, \tag{2}$$

where

$$\begin{aligned} \tilde{u}_i = {} & D_i \cdot (\tau(\boldsymbol{X}_i) - \tau) + (1 - D_i) \cdot (Y_i(0) - \mathbb{E}[Y_i(0)|D_i = 0, \boldsymbol{X}_i]) \\ & + D_i \cdot (Y_i(1) - \mathbb{E}[Y_i(1)|D_i = 1, \boldsymbol{X}_i]), \end{aligned}$$

and we imposed $\mathbb{E}[Y_i(0)|D_i = 0, \boldsymbol{X}_i] = \boldsymbol{X}_i{}'\beta$ to obtain the second line. Now, under Assumption 2, $\mathbb{E}[\tilde{u}_i|D_i = 0, \boldsymbol{X}_i] = \mathbb{E}[\tilde{u}_i|\boldsymbol{X}_i, D_i = 1] = 0$, so τ can again be consistently estimated with least squares methods.

4. Interference between Units in a GQE

In geographical natural experiments generally, and the GQE in particular, the research strategy is based on a comparison of units that are spatially proximate, under the assumption that units very close to one another but with opposite treatment status can provide valid counterfactuals for each other. However, in some cases, this focus on spatially proximate units may introduce the possibility of spillovers or interference between units. In our application, we are concerned about indviduals who reside in the all-mail-voting (treated) area being influenced by individuals who reside in the in-person-voting (control) area because in the latter area Election Day still acts as a focal point, and the act of voting is socially coordinated and shared. This might make the election generally more salient in the in-person area in the weeks before the election, affecting the propensity to vote of all-mail residents on the other side of the boundary. We assume that interference between units depends on geographic proximity to units of opposite treatment status. This is similar to the approach in Verbitsky-Savitz and Raudenbush (2012), although we do not focus on contiguity but rather on the density of control units that reside within a pre-specified distance of a given treated unit. Since we are not similarly concerned about individuals in the in-person-voting area being affected by individuals in the all-mail-voting area, we assume that interference is one-sided, from control to treated—but the setup is generalizable to two-sided patterns of interference.

Thus far, we assumed that resident i's treatment status did not depend on the treatment status of any other resident, allowing us to write potential outcomes as $Y_i(\boldsymbol{D}) = Y_i(D_i)$, since i's potential outcome only depended on i's own treatment status. If we allow for any pattern of interference, we must work with the full vector $Y_i(\boldsymbol{D})$, which allows individual i's treatment status to depend on the treatment status of every other individual. However, this level of generality is unworkable, as the large number of causal effects per subject makes it difficult to summarize the data in any interpretable way.

To add structure and reduce the dimensionality of our problem, we assume that each individual's potential outcome depends on the individual's own treatment status, D_i, and also on the number of individuals of the opposite treatment status who reside within a specified distance of i's location. To introduce the necessary notation, we first define the function $g_i(\boldsymbol{D}; \delta, \eta)$ for fixed values of the scalars δ and η, as follows:

$$g_i(\boldsymbol{D}; \delta, \eta) = \mathbb{1}(D_i = 1) \cdot \mathbb{1}(N_{i\delta} \geq \eta),$$

where

$$N_{i\delta} = \sum_{j=1}^{n} \mathbb{1}(d(i,j) \leq \delta) \cdot \left(D_j \neq D_i\right),$$

$\mathbb{1}(\cdot)$ is the indicator function, $d(i, j)$ is a measure of distance between i's and j's locations and $\delta, \eta \in \mathbb{R}$.

The function $g_i(\cdot)$ is an indicator for whether individual i receives interference, taking a value of 1 or 0 for every individual. If individual i is treated (i.e., resides in the all-mail-voting area), $g_i(\boldsymbol{D}; \delta, \eta) = 1$ if there are at least η control individuals who reside within δ meters of i's location, and $g_i(\boldsymbol{D}; \delta, \eta) = 0$ if there are less than η control individuals in a δ radius around i's location. If individual i is control (resides in the in-person-voting area), $g_i(\boldsymbol{D}; \delta, \eta) = 0$ regardless of how many treated individuals reside close to i—because we have assumed one-sided interference only.

Thus, we capture our model of one-sided geographic-based interference by letting the potential outcomes depend on the full vector of treatment assigments $\boldsymbol{D}$ in a restricted way; in particular, we let individual i's potential outcome depend on its own treatment status D_i and on the treatment status of other individuals only through the function $g_i(D; \delta, \eta)$. Thus, in our interference framework, $Y_i(\boldsymbol{D}) = Y_{i;\delta,\eta}(D_i, g_i(\boldsymbol{D}; \delta, \eta))$, reducing the number of arguments in each individual's potential outcome from n to 2. This is similar to the approach in Hong and Raudenbush (2006) and Verbitsky-Savitz and Raudenbush (2012), where the dependence of i's potential outcomes on the treatment assignment of all other units is also modeled via a scalar function that substantially reduces the range of possible potential outcomes that may occur.

Under our specific assumption of geographic-based interference, every unit has three potential outcomes:

- $Y_{i;\delta\eta}(1, 1)$: Individual i resides in the all-mail-voting area and receives spillovers, i.e. there are at least η control individuals in the in-person-voting area within δ meters of i's location.

- $Y_{i;\delta\eta}(1,0)$: Individual i resides in the all-mail-voting area and does not receive spillovers, i.e. there are less than η control individuals in the in-person-voting area within δ meters of i's location.
- $Y_{i;\delta\eta}(0,0)$: Individual i resides in the in-person-voting area. Since $g_i(\boldsymbol{D};\delta,\eta)=0$ for all individuals in the in-person-voting area, we can simply write $Y_{i;\delta\eta}(0,0)=Y_i(0)$.

We define the ATT in the absence of interference,

$$\tau_{\mathrm{T},\delta\eta}=\mathbb{E}\left[Y_{i;\delta\eta}(1,0)-Y_i(0)|D_i=1\right].$$

For brevity, we refer to this effect as the 'interference-free treatment effect'. In the context of our application, $\tau_{\mathrm{T},\delta\eta}$ captures the average effect of all-mail voting when individuals in the all-mail-voting area are geographically far from densely populated areas in the in-person-voting area and thus receive no spillovers. Since, under our model of interference, residents in the all-mail-voting area receive no spillovers when they are geographically distant from dense control areas, and there are no spillovers for control individuals, the two potential outcomes in $\tau_{\mathrm{T},\delta\eta}$ reflect the potential outcomes that would be observed under treatment and control in the absence of interference.

We also define an additional parameter,

$$\tau_{\mathrm{S},\delta\eta}=\mathbb{E}\left[Y_{i;\delta\eta}(1,1)-Y_{i;\delta\eta}(1,0)|D_i=1\right],$$

which captures the average effect of interference or spillovers on treated units. For brevity, we refer to this parameter as the 'interference effect'. In our application, $\tau_{\mathrm{S},\delta\eta}$ compares the average potential outcomes under all-mail voting for residents in the all-mail-voting area who are geographically close to the (densely populated parts of the) in-person-voting area, to the average potential outcome under all-mail voting for residents in the all-mail-voting area who are relatively isolated from the in-person area. Note that, in the absence of interference, $Y_{i;\delta\eta}(1,1)=Y_{i;\delta\eta}(1,0)=Y_i(1)$ for all i, which implies $\tau_{\mathrm{S},\delta\eta}=0$. Thus, a test of interference can be based on a test of the null hypothesis $H_0:\tau_{\mathrm{S},\delta\eta}=0$.

Both $\tau_{\mathrm{T},\delta\eta}$ and $\tau_{\mathrm{S},\delta\eta}$, however, depend on three different potential outcomes, and only one of those is observed for every i. We now investigate assumptions that, in the particular context of geographic natural or quasi experiments, could be invoked to identify these parameters. We first note that, letting Y_i denote the observed outcome for individual i, we have the following equalities between observed and potential outcomes:

- $Y_i=Y_i(0)$ if $D_i=0$
- $Y_i=Y_{i;\delta\eta}(1,1)$ if $D_i=1$ and $g_i(\boldsymbol{D};\delta,\eta)=1$
- $Y_i=Y_{i;\delta\eta}(1,0)$ if $D_i=1$ and $g_i(\boldsymbol{D};\delta,\eta)=0$.

We now consider the following assumption,

Assumption 3 (As-if random geographic location within interference areas).

$$\mathbb{E}\left[Y_{i;\delta\eta}(1,1)|D_i=1,g_i(\boldsymbol{D};\delta,\eta)=1,\boldsymbol{X}\right]=\mathbb{E}\left[Y_{i;\delta\eta}(1,1)|D_i=1,g_i(\boldsymbol{D};\delta,\eta)=0,\boldsymbol{X}\right]$$
$$\mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i=1,g_i(\boldsymbol{D};\delta,\eta)=1,\boldsymbol{X}\right]=\mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i=1,g_i(\boldsymbol{D};\delta,\eta)=0,\boldsymbol{X}\right]$$

A sufficient condition for Assumption 3 is that each unit is randomly assigned to a geographic location in the combined treated area, so that whether they fall in the interference region ($g_i(\boldsymbol{D}; \delta, \eta) = 1$) or the non-interference region ($g_i(\boldsymbol{D}; \delta, \eta) = 0$) is unrelated to their potential outcomes. Our assumption is weaker than this, since it requires only that conditional on pretreatment covariates, falling in the interference region in the treatment area is mean independent of potential outcomes, though it still is a strong assumption. However, as we discuss in detail below, under the type of treatment assignment that is typical of geographic natural or quasi experiments, this assumption might be plausible if the neighborhood around the boundary that separates treated and control areas is small enough and enough pretreatment covariates are available.

Under Assumptions 2 and 3, we have:

$$\begin{aligned}\tau_{\mathrm{T},\delta\eta}(\boldsymbol{X}) &\equiv \mathbb{E}\left[Y_{i;\delta\eta}(1,0) - Y_i(0)|D_i = 1, \boldsymbol{X}\right]\\ &= \mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 0, \boldsymbol{X}\right] - \mathbb{E}[Y_i(0)|D_i = 0, \boldsymbol{X}]\\ &= \mathbb{E}[Y_i|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 0, \boldsymbol{X}] - \mathbb{E}[Y_i|D_i = 0, \boldsymbol{X}]\end{aligned}$$

and

$$\begin{aligned}\tau_{\mathrm{S},\delta\eta}(\boldsymbol{X}) &\equiv \mathbb{E}\left[Y_{i;\delta\eta}(1,1) - Y_{i;\delta\eta}(1,0)|D_i = 1, \boldsymbol{X}\right]\\ &= \mathbb{E}\left[Y_{i;\delta\eta}(1,1)|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 1, \boldsymbol{X}\right] - \mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 0, \boldsymbol{X}\right]\\ &= \mathbb{E}[Y_i|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 1, \boldsymbol{X}] - \mathbb{E}[Y_i|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 0, \boldsymbol{X}].\end{aligned}$$

These results, which express $\tau_{\mathrm{T},\delta\eta}(\boldsymbol{X})$ and $\tau_{\mathrm{S},\delta\eta}(\boldsymbol{X})$ exclusively in terms of observable data, allow us to estimate and make inferences about the treatment effect in the absence of interference, $\tau_{\mathrm{T},\delta\eta}$, and the interference effect, $\tau_{\mathrm{S},\delta\eta}$.

As above, we outline an estimation strategy using least squares methods. To simplify the notation, let $G_i \equiv g_i(\boldsymbol{D}; \delta, \eta)$. Under Assumptions 2 and 3, we can express the observed outcome as:

$$\begin{aligned}Y_i &= \mathbb{E}[Y_i(0)|D_i = 0, \boldsymbol{X}_i] + \tau_{\mathrm{T},\delta\eta} \cdot D_i + \tau_{\mathrm{T},\delta\eta} \cdot G_i \cdot D_i + \tilde{\epsilon}_i\\ &= \boldsymbol{X}_i{}'\gamma + \tau_{\mathrm{T},\delta\eta} \cdot D_i + \tau_{\mathrm{S},\delta\eta} \cdot G_i \cdot D_i + \tilde{\epsilon}_i\end{aligned}, \tag{3}$$

where

$$\begin{aligned}\tilde{\epsilon}_i = {} & (Y_i(0) - \mathbb{E}[Y_i(0)|D_i = 0, \boldsymbol{X}_i]) \cdot (1 - D_i) +\\ & (Y_i(1,0) - \mathbb{E}[Y_i(1,0)|D_i = 1, G_i = 0, \boldsymbol{X}_i]) \cdot D_i(1 - G_i) +\\ & (Y_i(1,1) - \mathbb{E}[Y_i(1,1)|D_i = 1, G_i = 1, \boldsymbol{X}_i]) \cdot D_i G_i +\\ & D_i\left(\tau_{\mathrm{T},\delta\eta}(\boldsymbol{X}_i) - \tau_{\mathrm{T},\delta\eta}\right) + D_i \cdot G_i \cdot \left(\tau_{\mathrm{S},\delta\eta}(\boldsymbol{X}_i) - \tau_{\mathrm{S},\delta\eta}\right)\end{aligned},$$

and we again impose the linear specification $\tilde{\alpha} \equiv \mathbb{E}[Y_i(0)|D_i = 0, \boldsymbol{X}_i] = \boldsymbol{X}_i{}'\gamma$. Given our assumptions, $\mathbb{E}[\tilde{\epsilon}_i|D_i, G_i, \boldsymbol{X}_i] = 0$, and the parameters can be consistently estimated by least-squares. We note, however, that our identification assumptions do not rely on linearity assumptions and other, more flexible estimators could be employed to estimate the parameters of interest.

4.1 The treatment assignment mechanism in geographic natural and quasi experiments

We now discuss the plausibility of Assumptions 2 and 3, in particular the latter. Of course, in the absence of a concrete research design, both assumptions are exceedingly strong. But these assumptions may be more plausible when inferences are based on a GQE, where comparisons are made between units on one side or the other of the boundary, perhaps after conditioning on predetermined covariates. In a succesful GQE without interference, in a sufficiently small neighborhood around the boundary, we would have $\mathbb{E}[Y_i(0)|D_i = 0, \boldsymbol{X}] = \mathbb{E}[Y_i(0), \boldsymbol{X}]$, as stated in Assumption 2. But the random or as-if random assignment of units to treated or control areas does not imply Assumption 3, the assumption on which our derivations under interference were based. We now offer some discussion on the scenarios under which the assumption can be expected to hold.

If we consider the geographic location of every unit as fixed, an experiment where every unit has the same probability of receiving treatment might result in each unit having a different probability of receiving spillovers. This arises from the fact that when the locations of units are fixed, units that are spatially isolated and have no other units near them may have a small or 0 probability of receiving spillovers. In contrast, units that are in close proximity to other units may have a positive and larger probability of receiving spillovers. In this case, estimation of the parameters of interest may require weighting the observations according to each unit's probability of receiving spillovers (Gerber and Green 2012, Ch. 8).

However, keeping units' locations fixed and randomly locating the boundary may not be the most plausible way to conceptualize treatment assignment in our application. In many geographic designs, we might view the boundary as fixed, but we assume that, within a narrow band around this boundary, units randomly choose their geographic location. Units that happen to choose a location in the treated area receive the treatment, and units that happen to choose a location within the control area receive the control condition. In this sense, the assignment of treatment is seen as a result of units' location decisions, and therefore units' locations are not seen as fixed. Under this assignment mechanism, if every unit is equally likely to select any location within a fixed band around the boundary, each unit is equally likely to receive spillovers and *ex ante*, within the fixed band, the units' potential outcomes with and without interference are equally likely to be revealed.

Importantly, if we think this form of assignment mechanism is in operation, Assumption 3 holds naturally, provided the band around the boundary is sufficiently narrow. Why might we think this is true in the context of the quasi experiment in Colorado? First, the boundary is a county border, which has existed for decades, so it is natural to think of the boundary as fixed and individual location decisions near the border as random. Secondly, our focus on a town where residents on each side of the county border share the same city amenities, gives us the basis to assume that the choice of residence on each side of the boundary is unrelated to the (control) turnout potential outcomes, once we condition on predetermined covariates. Nonetheless, we must emphasize that however plausible this form of assignment mechanism appears to be, this is an untestable and strong assumption, and our conclusions about the extent of interference depend on its validity.

5. Exploring Interference Effects and the Extent of Interference

Before turning to the exploration of interference in our application, we use the framework introduced above to explore some features of the interference pattern in more detail. We explore two specific issues. First, we relax the assumption that all individuals in the interference area receive interference. So far, our framework assumed that all treated individuals near enough control individuals—that is, treated individuals with $g_i(\boldsymbol{D}; \delta, \eta) = 1$—received spillovers, and all other treated individuals were free of interference. We consider a generalization of this condition where only a proportion of the treated individuals near populated control areas receive interference—while individuals with $g_i(\boldsymbol{D}; \delta, \eta) = 0$ are still assumed to receive no interference. Secondly, we consider how the results from an analysis that mistakenly ignored interference and proceeded to simply compare the outcomes among treated and control units as if there were no spillovers would differ from the interference-free effect in an analysis that took interference into account.

5.1 Sensitivity of interference effect when geographic spillovers affect a subset of units

Above, the effect of interference was captured by the parameter $\tau_{S,\delta\eta}$, which assumed that every individual in the interference area was in fact affected by control individuals. We now investigate what happens when only a fraction of the individuals who reside in the interference area are affected by spillovers.

We assume that treated individuals with $g_i(\boldsymbol{D}; \delta, \eta) = 1$ receive interference from the control area with probability $0 < q \leq 1$ instead of with certainty. For given values of δ and η, the expected observed outcome for treated individuals with $g_i(\boldsymbol{D}; \delta, \eta) = 1$ is now

$$\mathbb{E}[Y_i|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 1, \boldsymbol{X}] = q \cdot \mathbb{E}\left[Y_{i;\delta\eta}(1,1)|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 1, \boldsymbol{X}\right] \\ + (1-q) \cdot \mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 1, \boldsymbol{X}\right].$$

Under this generalization, $\mathbb{E}[Y_i|D_i=1, g_i(\boldsymbol{D};\delta,\eta)=1,\boldsymbol{X}] \neq \mathbb{E}\left[Y_{i;\delta\eta}(1,1)|D_i=1, g_i(\boldsymbol{D};\delta,\eta)=1,\boldsymbol{X}\right]$, and $\tau_{S,\delta\eta}(\boldsymbol{X})$ can no longer be identified. However, under Assumptions 2 and 3, we have:

$$\begin{aligned}\mathbb{E}[Y_i|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 1, \boldsymbol{X}] &= q\mathbb{E}\left[Y_{i;\delta\eta}(1,1)|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 1, X\right] \\ &\quad +(1-q)\mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 1, \boldsymbol{X}\right] \\ &= q\mathbb{E}\left[Y_{i;\delta\eta}(1,1)|D_i = 1, \boldsymbol{X}\right] + (1-q)\mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i = 1, \boldsymbol{X}\right], \quad \text{and} \\ \mathbb{E}[Y_i|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 0|\boldsymbol{X}] &= \mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 0, \boldsymbol{X}\right] = \mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i = 1, \boldsymbol{X}\right],\end{aligned}$$

leading to

$$\begin{aligned}&\mathbb{E}[Y_i|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 1, \boldsymbol{X}] - \mathbb{E}[Y_i|D_i = 1, g_i(\boldsymbol{D}; \delta, \eta) = 0, \boldsymbol{X}] \\ &\quad = q\mathbb{E}\left[Y_{i;\delta\eta}(1,1)|D_i = 1, \boldsymbol{X}\right] + (1-q)\mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i = 1, \boldsymbol{X}\right] \\ &\qquad -\mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i = 1, \boldsymbol{X}\right] \\ &\quad = q\left\{\mathbb{E}\left[Y_{i;\delta\eta}(1,1)|D_i = 1, \boldsymbol{X}\right] - \mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i = 1, \boldsymbol{X}\right]\right\} \\ &\quad = q\tau_{S,\delta\eta}(\boldsymbol{X})\end{aligned}$$

Therefore, the interference effect, $\tau_{\mathrm{S},\delta\eta}$, is now a function of q. This effect, which we denote by $\tau^{q}_{\mathrm{S},\delta\eta}(\boldsymbol{X})$, is given by:

$$\tau^{q}_{\mathrm{S},\delta\eta}(\boldsymbol{X}) = (1/q)(\mathbb{E}[Y_i|D_i = 1, g_i(\boldsymbol{D};\delta,\eta) = 1, \boldsymbol{X}] - \mathbb{E}[Y_i|D_i = 1, g_i(\boldsymbol{D};\delta,\eta) = 0, \boldsymbol{X}]).$$

Thus, whereas before all treated individuals in close proximity to (enough) control individuals were assumed to receive spillovers, now only $q\%$ of them have their outcomes affected by interference from individuals in the control area, while the remaining $(1-q)\%$ have the same outcome they would have had if they had been in the interference-free area where $g_i(\boldsymbol{D};\delta,\eta) = 0$.

This analysis reveals that the differences in observed average treated outcomes between the interference-free and the interference areas are equal to the interference effect $\tau^{q}_{\mathrm{S},\delta\eta}$ when $q = 1$. But when $q < 1$, the true effect of interference will be larger than the observed difference in outcomes by a factor $1/q > 1$. Thus, assuming that interference affects all units in the interference area ($q = 1$) gives a lower bound on the interference effect.

5.2 Characterizing the extent of interference

We now investigate the degree to which the conclusions from an analysis that ignored interference when interference was in fact present would lead to incorrect conclusions. Recall that, in the absence of interference, we defined the parameter $\tau = \mathbb{E}[Y_i(1) - Y_i(0)|D_i = 1]$. This parameter, however, is undefined in the presence of interference. When we allowed for interference, we focused instead on the parameters $\tau_{\mathrm{T},\delta\eta} = \mathbb{E}\left[Y_{i;\delta,\eta}(1,0) - Y_i(0)|D_i = 1\right]$ and $\tau_{\mathrm{S},\delta\eta} = \mathbb{E}\left[Y_{i;\delta,\eta}(1,1) - Y_{i;\delta,\eta}(1,0)|D_i = 1\right]$.

If units interfered with each other but an analyst made the incorrect assumption that the study is interference-free, the analyst would proceed to estimate the mean outcome differences between all units in the treatment area and all units in the control area—after conditioning on pretreatment covariates if the assumptions introduced above were invoked. Under our setup, however, the average among treated outcomes conditional on $\boldsymbol{X}$ would not equal $\mathbb{E}[Y_i(1)|D_i = 1, \boldsymbol{X}]$ but rather the weighted average of $\mathbb{E}\left[Y_{i;\delta,\eta}(1,1)|D_i = 1, \boldsymbol{X}\right]$ and $\mathbb{E}\left[Y_{i;\delta,\eta}(1,0)|D_i = 1, \boldsymbol{X}\right]$. Letting $p_{\mathrm{I}} = \Pr[g_i(\boldsymbol{D};\delta,\eta) = 1]$, the estimand estimated by the analyst that ignored interference would be

$$\begin{aligned}
\mathbb{E}[Y_i|D_i = 1, \boldsymbol{X}] - \mathbb{E}[Y_i|D_i = 0, \boldsymbol{X}] = \\
&= \{p_{\mathrm{I}} \cdot \mathbb{E}[Y_i|D_i = 1, \boldsymbol{X}, g_i(\boldsymbol{D};\delta,\eta) = 1] + (1-p_{\mathrm{I}}) \cdot \mathbb{E}[Y_i|D_i = 1, \boldsymbol{X}, g_i(\boldsymbol{D};\delta,\eta) = 0]\}. \\
&\quad -\mathbb{E}[Y_i(0)|D_i = 1, \boldsymbol{X}]
\end{aligned}$$

The estimand is now a comparison between the control units and a weighted average of treated units—some of which are subject to interference and some not. Writing the overall estimand this way we observe that, when p_{I} is small, the approach that ignored interference and pooled all treated observations would closely approximate the interference-free treatment effect, $\tau_{\mathrm{T},\delta\eta}$, since in this case, under the assumptions of our framework, we would have:

$$\begin{aligned}
&p_{\mathrm{I}} \cdot \mathbb{E}[Y_i|D_i = 1, \boldsymbol{X}, g_i(\boldsymbol{D};\delta,\eta) = 1] + (1-p_{\mathrm{I}}) \cdot \mathbb{E}[Y_i|D_i = 1, \boldsymbol{X}, g_i(\boldsymbol{D};\delta,\eta) = 0] \\
&\approx \mathbb{E}[Y_i|D_i = 1, \boldsymbol{X}, g_i(\boldsymbol{D};\delta,\eta) = 0] = \mathbb{E}\left[Y_{i;\delta\eta}(1,0)|D_i = 1, \boldsymbol{X}\right],
\end{aligned},$$

the first term in $\tau_{\mathrm{T},\delta\eta}$.

Moreover, when the potential outcomes are bounded as they are in our voting application where they take values equal to 0 (non voting) or 1 (voting), given a value of p_{I}, we can calculate the maximum value of the term $p_{\mathrm{I}} \cdot \mathbb{E}[Y_i|D_i = 1, \boldsymbol{X}, g_i(\boldsymbol{D}; \delta, \eta) = 1]$. In our case, this upper bound is p_{I}.

This discussion illustrates that it is not only of interest to calculate the effect of interference but also to establish whether interference affects a large enough proportion of units. If it does not—that is, if p_{I} is small—in the case of bounded outcomes, we may be able to assert that an analysis that ignored interference when in fact interference is present would produce treatment effect estimates that would approximate the effects under no interference.

6. Application to the 2010 Colorado Primary

We now apply the framework described above to analyze the 2010 Colorado primary, starting with a description of the data. Our main source of information is the Colorado voter registration file, the database of registered voters maintained by the state of Colorado for administrative purposes. We acquired these data from a private vendor. The administrative data from the state contain a limited number of covariates including date of birth, gender, voting history, voters' addresses, and the legislative districts in which each voter's address is included. The private vendor also includes an additional variable estimating the voter's likely race. In this region, most voters are White with a substantial minority of Hispanic voters. To determine voter locations, we geocoded each voter's location using the address in the voter file.

We first restrict our data to include only those individuals who in 2010 lived in the central area of Basalt that is split by the border between Eagle and Pitkin counties—the border that determines all-mail or in-person voting. Within the central area of Basalt, our covariate-adjusted results condition on the set of pre-treatment covariates that we have available: age, gender, whether the individual is Hispanic, voting history for 2008 and 2006, party affiliation as declared in the registration file, and an indicator for whether the individual's registration status is considered active by the state. Since our main data source is the Colorado registration file, all our turnout measures—including both pre-treatment turnout shares and the turnout share in the 2010 primary—are constructed as the proportion of individuals in the registration file that turn out to vote in a given election. This means that our turnout measures condition on registration status. We discuss the potential methodological complications associated with such measures in Section 6.2.

Before presenting the estimation results, we examine the observed covariates in our sample within central Basalt, to asses whether the treated and control areas inside this small region are already comparable in terms of these characteristics. Table 1 contains sample means and the absolute standardized differences in means (difference in means divided by the pooled standard deviation between groups before matching) for three demographic characteristics, voter registration status, and turnout in the last four elections. While geographic proximity produces acceptable balance on residents' Hispanic ethnicity, gender, age, and active registration status, there are larger differences in turnout in past elections, with standardized differences in these variables exceeding 0.20. These differences suggest that unadjusted comparisons between the groups cannot be interpreted as causal effects of all-mail elections on 2010 voter turnout.

Table 1. Covariate balance between treated and control areas around the geographic border

	Mean treated	Mean control	Std. diff.
Hispanic	0.07	0.08	0.03
Age	48.4	45.6	0.19
Female	0.49	0.50	0.01
Active Registration	0.24	0.29	0.10
2008 General Election Turnout	0.71	0.60	0.24
2008 Primary Election Turnout	0.07	0.02	0.27
2006 General Election Turnout	0.48	0.36	0.24
2006 Primary Election Turnout	0.06	0.01	0.22

Notes: Total sample size is 977 treated (all-mail) voters and 620 control (in-person) voters. Std. diff. = absolute standardized difference. Means for turnout are proportion of registered individuals voting in that election.

Table 2. Estimated average effect of all-mail voting on turnout on treated area under no interference, 2010 Colorado primary election

	Unadjusted	Covariate-adjusted
Difference in turnout rates	0.010	−0.059
95% confidence interval	(−0.027, 0.047)	(−0.091,−0.027)
Control turnout rate	0.155	

Notes: Total sample size is 977 treated (all-mail) voters and 620 control (in-person) voters. Turnout shares calculated as proportion of registered individuals voting in that election.

We estimate Equation (2) by least-squares using the variables in Table 1 as covariates. Table 2 contains point estimates of τ and associated 95% confidence intervals. We also report the unadjusted least-squares estimator of τ corresponding to Equation (1), which is simply the unadjusted difference in turnout rates between treated and control areas. This unadjusted estimate is reported only for completeness, but we do not believe it can be interpreted as a causal effect due to the observable pre-treatment differences between the areas.

We find that, in a narrow band around the boundary between the treated and control areas and after conditioning on a set of observed characteristics, the voter turnout rate was 6 percentage points lower in the all-mail county, with a 95% confidence interval ranging from −0.091 to −0.027. Given our assumptions, this difference is the average all-mail voting effect on turnout for the residents in the treated area. This analysis assumes that outcomes of an individual do not depend on the treatment status of other individuals, a constraint that we relax in the following section.

6.1 Empirical results allowing for spatial interference

We now re-analyze the effects above based on our GQE. Under our setup and given the assumptions introduced above, we calculate the interference-free treatment, $\tau_{\mathrm{T},\delta\eta}$, and the interference effect, $\tau_{\mathrm{T},\delta\eta}$, for different values of η and δ.

To implement estimation of these quantities, we calculate, given δ and η, the interference set, $\mathcal{I}_{\delta\eta}$, which is simply the collection of all treated individuals in our data for whom

$g_i(\boldsymbol{D}; \delta, \eta) = 1$. Once we form $\mathcal{I}_{\delta\eta}$, we can estimate $\tau_{S,\delta\eta}$ through a comparison of treated individuals that are in $\mathcal{I}_{\delta\eta}$ to those who are not, since this set identifies all treated individuals that are spatially proximate to enough control voters and may have been subject to interference. Similarly, we estimate $\tau_{T,\delta\eta}$ through a comparison of control voters to treated voters not in $\mathcal{I}_{\delta\eta}$. To form $\mathcal{I}_{\delta\eta}$, we must locate the distance between each treated unit and each control unit, calculate whether any control units reside within δ meters of each treated unit, and then count the number of control units within δ distance. We calculate this set using the following algorithm.

Algorithm 6.1: Computing the Interference Set($\mathcal{I}_{\delta\eta}$,δ,η)

for $i \leftarrow 1$ to m treated units

do $\begin{cases} \text{Calculate distance from treated } i \text{ to all controls} \\ \text{Locate all controls within } \delta \text{ distance of treated } i \\ N_{i\delta} \leftarrow \text{ number of control units that are } \delta \text{ distance from treated } i \\ \text{Place treated unit } i \text{ in } \mathcal{I}_{\delta\eta} \text{ if } N_{i\delta} \geq \eta \end{cases}$

return ($\mathcal{I}_{\delta\eta}$)

The size of $\mathcal{I}_{\delta\eta}$ depends on the values we choose for δ and η. As we make δ larger and η smaller, we allow for interference to become more severe. Once we have formed the set $\mathcal{I}_{\delta\eta}$, calculation of the quantities of interest is straightforward.

We apply these methods to the data from the 2010 primary in Colorado. For a set of η and δ values, we estimate $\tau_{S,\delta\eta}$ and $\tau_{T,\delta\eta}$. Different values of η and δ allow for more or less stringent definitions of spatial interference. For example, if we set δ to 250 m and η to 1, our model of interference asserts that all treated units who have at least one control unit within a 250-m radius of their location were subject to interference. We set δ to 250 and 100 m, and η to 1, 5, and 10 individuals. We estimate $\tau_{S,\delta\eta}$ and $\tau_{T,\delta\eta}$ for each set of values to observe whether these quantities change as a function of both distance and the density of control units that treated individuals are near to.

Table 3 contains estimates for $\tau_{S,\delta\eta}$ and $\tau_{T,\delta\eta}$ (which we denote as $\hat{\tau}_{S,\delta\eta}$ and $\hat{\tau}_{T,\delta\eta}$) along with 95% confidence intervals for each set of η and δ values. The table also reports the proportion of treated voters in the interference set (i.e., an estimate of p_I). All results in the table are covariate-adjusted using least-squares estimates from Equation (3), employing heteroscedasiticty-robust standard errors.

First, we consider the estimates when $\delta = 100$ (top panel). When $\eta = 1$, we let interference affect all treated units who live within 100 m of at least one control unit. Under this interference scenario, 9.6% of the treated observations are affected by interference, and estimate of $\tau_{T,\delta\eta}$ is -0.059 (95% confidence interval from -0.092 to -0.026), which is very similar to our covariate-adjusted estimate of τ in Table 2—the ATT estimated under no interference. In contrast, the estimate for $\tau_{S,\delta\eta}$ is -0.0056, much closer to 0 and not significantly different from 0 (95% confidence interval ranges from -0.061 to 0.050). Thus, for $\delta = 100$ and $\eta = 1$, there is not a statistically significant difference in the patterns of voter turnout between the interference and the interference-free area. When η is 5 or 10, the proportion of treated voters inside the interference area is naturally smaller. In this case, the estimates of $\tau_{S,\delta\eta}$ are, respectively, -0.012 and -0.054, both indistinguishable from 0. Since

Table 3. Estimated average effects of all-mail voting on turnout on treated area under difference interference scenarios, 2010 Colorado primary election

	$\eta=1$	$\eta=5$	$\eta=10$
		$\delta=100$ meters	
Interference-free effect ($\hat{\tau}_{T,\delta\eta}$)	−0.059	−0.059	−0.058
	(−0.092, −0.026)	(−0.091, −0.027)	(−0.090, −0.026)
Interference effect ($\hat{\tau}_{S,\delta\eta}$)	−0.0056	−0.012	−0.054
	(−0.061, 0.050)	(−0.109, 0.085)	(−0.155, 0.047)
$\hat{p}_{\mathrm{I}}$	0.096	0.045	0.031
		$\delta=250$ meters	
Interference-free effect ($\hat{\tau}_{T,\delta\eta}$)	−0.069	−0.062	−0.058
	(−0.104, −0.034)	(−0.095, −0.029)	(−0.091, −0.025)
Interference effect ($\hat{\tau}_{S,\delta\eta}$)	0.033	0.012	−0.011
	(−0.0078, 0.0738)	(−0.037, 0.061)	(−0.060, 0.038)
$\hat{p}_{\mathrm{I}}$	0.294	0.177	0.156

Notes: Total sample size is 977 treated (all-mail) voters and 620 control (in-person) voters. Turnout shares calculated as proportion of registered individuals voting in that election.

the proportion of treated units in the interference set is very small ($p_{\mathcal{I}}$ is equal to 0.045 and 0.031 for $\eta=5$ and $\eta=10$, respectively), the covariate-adjusted estimate for $\tau_{\mathrm{T},\delta\eta}$ remains very similar to the analogous estimate for $\eta=1$—and also to the covariate-adjusted effect under no interference reported in Table 2.

Next, we consider the case of $\delta=250$ (bottom panel of Table 3). When $\eta=1$, the estimated interference-free effect is −0.069 (confidence interval ranging from −0.104 to −0.034), similar to our previous estimates. In addition, turnout is about 3.3 percentage points higher for treated voters in the interference area relative to the turnout of treated voters in the interference-free area, but this effect is again statistically indistinguishable from 0. Note that in this scenario interference is assumed to be much more prevalent, with 29% of treated voters inside the interference area. When we set η to either 5 or 10, the estimated effect of interference, $\hat{\tau}_{\mathrm{S},\delta\eta}$, continues to be indistinguishable from 0. In contrast, the confidence intervals including for $\tau_{\mathrm{T},\delta\eta}$ are $[-0.095, -0.029]$ for $\eta=5$ and $[-0.091, -0.025]$ for $\eta=10$, very similar to the confidence intervals under $\delta=100$ and also to confidence intervals for τ ignoring interference reported in Table 2.

In sum, our analysis shows that, under our assumptions, the interference effect is indistinguishable from 0 in all cases. Moreover, except under the most extreme scenario of interference with $\delta=1$ and $\eta=250$ where 29% of treated voters are in the interference set, the proportion of treated voters in the interference set tends to be small (between 17 and 3%), leading to an interference-free effect that is similar to the effect estimated assuming that interference is not present. These results suggest that all-mail voting reduced voter turnout in the 2010 primary election by an average of about 6 percentage points.

6.2 Sensitivity to differential voter registration

A potential complication with our analysis is that our data source is the Colorado voter registration file, and thus our measure of voter turnout is constructed as the proportion of registered citizens who turn out to vote. If the decision to register is itself affected by the mode of voting, our reported results could misrepresent the true turnout effects. Eagle county's decision to adopt all-mail voting was announced in 2010 before the primary election was held, making it possible for citizens to adjust their election registration decisions in response to the upcoming change in the mode of voting.

The ideal solution would be to obtain a list of the total voting eligible population in the treated and control areas at the moment of the 2010 primary election. Unfortunately, such data are unavailable. An alternative is to follow the approach in Nyhan et al. (2017) and explore how much differential registration between the treated and control groups could occur before our observed turnout effects—which construct turnout shares as total voters over total registration—became consistent with a zero effect on the true turnout share—the share of voters to the total voting eligible population. Generalizing the sensitivity analysis in Nyhan et al. (2017) to include covariates, for every estimated effect reported in Table 3, we report the differential registration factor $k^\star$—the amount of differential registration between treated and control groups that would be required to produce the difference in turnout-to-registration rates we observed if the true turnout effects were equal to 0.

Given a treated and a control or reference group, the differential registration factor is estimated by simply dividing the turnout-to-registration share in the control group (T_c^{Reg}) over the turnout-to-registration share in the treatment group (T_t^{Reg}). For estimation of the $k^\star$ associated with $\tau_{T,\delta\eta}$, T_t^{Reg} includes all units in the treatment group outside of the interference region ($D_i = 1$ and $g_i(\boldsymbol{D}; \delta, \eta) = 0$), and T_c^{Reg} includes all control units ($D_i = 0$). For estimation of the $k^\star$ associated with $\tau_{S,\delta\eta}$, T_t^{Reg} includes all units in the treatment group inside the interference region ($D_i = 1$ and $g_i(\boldsymbol{D}; \delta, \eta) = 1$), and T_c^{Reg} includes units in the treatment group outside of the interference region ($D_i = 0$ and $g_i(\boldsymbol{D}; \delta, \eta) = 0$). To incorporate covariates, we estimate T_c^{Reg} with the average predicted values from the linear model for all observations in the treatment group, but with the corresponding treatment indicator set to 0.

The results of the sensitivity analysis are shown in Table 4, where we report, for each of the three values of η combined with $\delta = 100$ or $\delta = 250$, the differential registration factor for both $\tau_{T,\delta\eta}$ and $\tau_{S,\delta\eta}$. For each combination of η and δ, the first two columns report T_c^{Reg} and T_t^{Reg}, the estimated values of the turnout-to-registration shares for each group—the difference between these values are the effects reported in Table 3. The third column reports the differential registration factor associated with each effect. For example, for $\eta = 1$ and $\delta = 250$, $k^\star$ is 1.447, showing that the rate of registration in the treatment group would have

Table 4. Sensitivity of average effects of all-mail voting to differential registration, 2010 Colorado primary election

	$\eta = 1$			$\eta = 5$			$\eta = 10$		
	T_t^{Reg}	T_c^{Reg}	$k^\star$	T_t^{Reg}	T_c^{Reg}	$k^\star$	T_t^{Reg}	T_c^{Reg}	$k^\star$
				$\delta = 100$ meters					
Interference-free effect ($\hat{\tau}_{T,\delta\eta}$)	0.165	0.224	1.355	0.165	0.224	1.355	0.166	0.224	1.346
Interference effect ($\tau_{S,\delta\eta}$)	0.117	0.123	1.048	0.182	0.194	1.066	0.133	0.188	1.407
				$\delta = 250$ meters					
Interference-free effect ($\hat{\tau}_{T,\delta\eta}$)	0.155	0.224	1.447	0.163	0.224	1.378	0.166	0.224	1.345
Interference effect ($\tau_{S,\delta\eta}$)	0.171	0.137	0.805	0.156	0.144	0.921	0.132	0.143	1.083

Notes: Total sample size is 977 treated (all-mail) voters and 620 control (in-person) voters. Turnout shares calculated as proportion of registered individuals voting in that election.

to be 44.7 percentage points higher in the treatment group than in the control group to make the estimated interference free treatment effect $\hat{\tau}_{T,\delta\eta} = -0.069$ (reported in Table 3 and also obtained from Table 4 as 0.155–0.224) consistent with a 0 effect on true turnout rates. A 44.7 percentage point difference in registration rates is a very large effect, unlikely to occur in practice. This large value of $k^{\star}$ suggests that the interference-free effect is robust: the negative turnout effect would remain even with substantial differential registration between the treated and control groups.

In general, turnout-to-registration effects are more sensitive to differential registration whenever these effects are larger in absolute value and whenever the turnout-to-registration share in the baseline group is smaller. Since our estimated interference-free effects $\tau_{T,\delta\eta}$ are much larger in absolute value than the interference effects $\tau_{S,\delta\eta}$, and these effects are large relative to $T_{\mathrm{c}}^{\mathrm{Reg}}$, the pattern in Table 4 is consistent: the differential registration factor $k^{\star}$ is large for the interference-free treatment effect $\tau_{T,\delta\eta}$ in all cases, and small for the interference effects $\tau_{S,\delta\eta}$. Our conclusion is that the estimated interference effects $\tau_{S,\delta\eta}$, even if statistically distinguishable from 0, would be less robust to differential registration patterns. For example, for $\eta = 5$ and $\delta = 100$, the registration factor of 1.066 indicates that a difference in registration rates of 6.6 percentage points would be sufficient to make the observed interferecne effect $\tau_{S,\delta\eta} = -0.012$ consistent with a 0 effect on true turnout rates.

Finally, we note that all our conclusions about sensitivity to differential registration assume that the interference set is correctly calculated based on the registration file. This implies the assumption that individuals who reside in the control area and are not registered to vote do not affect the potential outcomes of treated individuals in the all-mail voting area. This may be plausible if we assume that residents who are not registered are not known to get-out-the-vote campaigns and are not actively involved in political activities.

7. Discussion

Policymakers seldom use randomized experiments to study the effects of different voting regulations such as modes of voter registration or convenience voting policies on voter participation. GQEs such as the one we examine here may therefore provide fruitful opportunities for researchers to study policy effects that would otherwise go unexplored. Such natural experiments are not only useful in social science applications—where the assumption of comparability on either side of the geographic border is always a strong assumption—but also in other disciplines where the potential confounders are more closely related to the physical characteristics of the terrain and thus more likely to be offset by spatial proximity (see Wonkka et al. 2015).

In our analysis of the 2010 Colorado primary, we find that vote-by-mail elections appear to suppress turnout. Using our framework, we find that, unless we assume a fairly strong patten of interference, the treatment effects estimated under a framework that ignores interference would not differ from the interference-free effect in our spatial interference framework. Moreover, our estimated interference effect could not be distinguished from 0 in any of the scenarios we considered. Thus, given our assumptions, interference between voters does not seem to be prevalent in our application, as was also found by Sinclair et al. (2012).

The prior literature has found both positive and negative effects of all-mail reforms on voter turnout; our results are consistent with prior evidence of negative effects. The reasons behind the different conclusions across studies are hard to ascertain, as all studies, including our own, are non-experimental and potentially threatened by unobserved confounders driving the decision to adopt voting reforms. In addition to these threats to internal validity, the differences could be partially due to heterogeneity in the trade-off between convenience voting and the social aspects of voting in different political contexts. In those places where voting barriers affect a large proportion of potential voters, the lower barriers to voting as a result of convenience voting reforms may more than compensate the lower turnout that may be induced by removing the social aspect of voting. In contrast, in settings where most citizens can afford the costs of voting in the absence of convenience reforms, removing the social aspect and focal point of election day may lead to a decrease in turnout that is not compensated by making voting easier. Since we focus on a primary election, we are focusing on a subset of citizens who are highly committed and vote in most elections; moreover, in the 2010 Colorado primary, voters had three salient races on the ballot. This suggests that the scenario we study might belong to the latter category, where the elimination of the social and shared aspects of voting is relatively more costly.

Acknowledgements

This manuscript was prepared for the CESifo Economic Studies special issue on geocoding to follow our keynote lecture presentation at the 2016 CESifo Economic Studies Conference on the Use of Geo-coded Data in Economic Research, held in Munich in 18 and 19 November 2016. The authors thank Matias Cattaneo, Jake Bowers, Jeff Lewis, seminar participants at the University of California, Berkeley, and conference participants at the 2013 Causality in Political Networks Conference at the University of Chicago, the 2014 Southeastern Methods Meeting at the University of South Carolina, the 2015 Harvard University Conference on Political Geography, and the 2016 CESifo Economic Studies Conference on the Use of Geo-coded Data in Economic Research, for valuable comments and discussion. R.T. gratefully acknowledges financial support from the National Science Foundation (SES 1357561). A previous version of this manuscript was circulated under the title 'A Framework for Understanding Interference Between Units In Geographic Natural Experiments with an Application to All-Mail Voting in Colorado'.

References

Aronow, P. M. and C. Samii (2017), "Estimating Average Causal Effects Under Interference Between Units," *Annals of Applied Statistics* **11**, 1921–047.

Bergman, E. and P. A. Yates (2011), "Changing Election Methods: How Does Mandated Vote-by-Mail Affect Individual Registrants?", *Election Law Journal* **10**, 115–27.

Berinsky, A. J., N. Burns, and M. W. Traugott (2001), "Who Votes by Mail?: A Dynamic Model of the Individual-Level Consequences of Voting-by-Mail Systems", *Public Opinion Quarterly* **65**, 178–97.

Bowers, J., M. M. Fredrickson, and C. Panagopoulos (2013), "Reasoning about Interference between Units: A General Framework", *Political Analysis* **21**, 97–124.

Galiani, S., P. J. McEwan, and B. Quistorff (2017), *Regression Discontinuity Designs: Theory and Applications, Advances in Econometrics*, vol. **38**. In M. D., Cattaneo and J. C., Escanciano eds, Emerald Publishing Limited, Bingley, United Kingdom, pp. 195–236.

Gerber, A. S. and D. P. Green (2012), *Field Experiments: Design, Analysis, and Interpretation*. Norton, New York, NY.

Gerber, A. S., D. P. Green, and C. W. Larimer (2008), "Social Pressure and Voter Turnout: Evidence from a Large-Scale Field Experiment", *American Political Science Review* **102**, 33–48.

Gerber, A. S., G. A. Huber, and S. J. Hill (2013), "Identifying the Effect of All-Mail Elections on Turnout: Staggered Reform in the Evergreen State", *Political Science Research and Methods* **1**, 91–116.

Gronke, P., E. Galanes-Rosenbaum, P. A. Miller, and D. Toffey (2008), "Convenience Voting", *Annual Review of Political Science* **11**, 437–55.

Gronke, P. and P. Miller (2012), "Voting by Mail and Turnout in Oregon Revisiting Southwell and Burchett", *American Politics Research* **40**, 976–97.

Holland, P. W. (1986), "Statistics and Causal Inference", *Journal of the American Statistical Association* **81**, 945–60.

Hong, G. and S. W. Raudenbush (2006), "Evaluating Kindergarten Retention Policy", A Case Study of Causal Inference for Multilevel Observational Data," *Journal of the American Statistical Association* **101**, 901–10.

Hudgens, M. G. and M. E. Halloran (2008), "Toward Causal Inference with Interference", *Journal of the American Statistical Association* **103**, 832.

Karp, J. A. and S. A. Banducci (2000), "Going Postal: How All-Mail Elections Influence Turnout", *Political Behavior* **22**, 223–39.

Keele, L., S. Lorch, M. Passarella, D. Small, and R. Titiunik (2017), "An Overview of Geographically Discontinuous Treatment Assignments with an Application to Children's Health Insurance," in M. D., Cattaneo and J. C., Escanciano, eds, *Regression Discontinuity Designs: Theory and Applications, Advances in Econometrics*, vol. **38**. Emerald Publishing Limited, Bingley, United Kingdom, pp. 147–94.

Keele, L. and R. Titiunik (2015), "Geographic Boundaries as Regression Discontinuities", *Political Analysis* **23**, 127–55.

Keele, L., R. Titiunik, and J. Zubizarreta (2015), "Enhancing a Geographic Regression Discontinuity Design through Matching to Estimate the Effect of Ballot Initiatives on Voter Turnout", *Journal of the Royal Statistical Society: Series A* **178**, 223–39.

Keele, L. J. and R. Titiunik (2016), "Natural Experiments Based on Geography", *Political Science Research and Methods* **4**, 65–95.

Kousser, T. and M. Mullin (2007), "Does Voting by Mail Increase Participation? Using Matching to Analyze a Natural Experiment", *Political Analysis* **15**, 428–45.

Nyhan, B., C. Skovron, and R. Titiunik (2017), "Differential Registration Bias in Voter File Data: A Sensitivity Analysis Approach", *American Journal of Political Science* **61**, 744–60.

Richey, S. (2008), "Voting by Mail: Turnout and Institutional Reform in Oregon", *Social Science Quarterly* **89**, 902–15.

Rosenbaum, P. R. (2007), "Interference between Units in Randomized Experiments", *Journal of the American Statistical Association* **102**, 191–200.

Rubin, D. B. (2005), "Causal Inference Using Potential Outcomes", *Journal of the American Statistical Association* **100**, 322.

Sinclair, B., M. McConnell, and D. P. Green (2012), "Detecting Spillover Effects: Design and Analysis of Multilevel Experiments", *American Journal of Political Science* **56**, 1055–69.

Sobel, M. E. (2006), "What Do Randomized Studies of Housing Mobility Demonstrate? Causal Inference in The Face of Interference," *Journal of the American Statistical Association* **101**, 1398–407.

Southwell, P. L. and J. I. Burchett (2000), "The Effect of All-Mail Elections on Voter Turnout", *American Politics Research* **28**, 72–9.

Tchetgen Tchetgen, E. J. and T. J. VanderWeele (2012), "On Causal Inference in the Presence of Interference", *Statistical Methods in Medical Research* **21**, 55–75.

Vanderweele, T. (2008), "Ignorability and Stability Assumptions in Neighborhood Effects Research", *Statistics in Medicine* **27**, 1934–43.

VanderWeele, T. J., G. Hong, S. M. Jones, and J. L. Brown (2013), "Mediation and Spillover Effects in Group Randomized Trials: A Case Study of the 4Rs Educational Intervention", *Journal of the American Statistical Association* **108**, 469–81.

Verbitsky-Savitz, N. and S. W. Raudenbush (2012), "Causal Inference under Interference in Spatial Settings: A Case Study Evaluating Community Policing Program in Chicago", *Epidemiologic Methods* **1**.

Verdery, A. M., B. Entwisle, K. Faust, and R. R. Rindfuss (2012), "Social and Spatial Networks: Kinship Distance and Dwelling Unit Proximity in Rural Thailand", *Social Networks* **34**, 112–27.

Wonkka, C. L., W. E. Rogers, and U. P. Kreuter (2015), "Legal Barriers to Effective Ecosystem Management: Exploring Linkages between Liability, Regulations, and Prescribed Fire", *Ecological Applications* **25**, 2382–93.