

# A Precinct Too Far: Turnout and Voting Costs

Enrico Cantoni\*

August 28, 2016

Latest version available at <http://economics.mit.edu/files/11936>

## Abstract

I study the effects of voting costs through a novel, quasi-experimental design based on geographic discontinuities. I compare parcels and census blocks located near borders between adjacent voting precincts. Units on opposite sides of a border are observationally identical, except for their assignment to different polling locations. The discontinuous assignment to polling places produces sharp changes in the travel distance voters face to cast their ballots. In a sample of nine municipalities in Massachusetts and Minnesota, I find that a 1-standard deviation (.245 mile) increase in distance to the polling place reduces the number of ballots cast by 2% to 5% in the 2012 presidential, 2013 municipal, 2014 midterm, and 2016 presidential primary elections. During non-presidential elections, effects in high-minority areas are three times as large as in low-minority areas, while no significant difference emerges from the 2012 presidential election. Finally, I use my estimates to simulate the impact of various counterfactual assignments of voters to polling places. I find that erasing the effect of distances to polling places would increase turnout by 1.6 to 4 percentage points and reduce minority participation gaps in non-presidential elections by 11% to 13%. By contrast, the optimal feasible counterfactual boundaries, holding polling locations constant, would result in small changes in the minority participation gap.

---

\*Department of Economics, Massachusetts Institute of Technology. E-mail: [cantoni@mit.edu](mailto:cantoni@mit.edu).

# 1 Introduction

The legitimacy of every modern democracy rests on the notion that representatives of the people are chosen by the people. In this perspective, each vote is a small contribution to the legitimacy of the democratic process and to the representativeness of elected officials and public policies. However, a single vote is rarely, if ever, consequential for electoral outcomes, raising the possibility that even small changes in voting costs can have a sizable impact on voter participation. In this constant tension between voting as the source of democratic legitimacy and individual votes that carry no tangible consequences lies the Achilles' heel of democracy.

The peril that low participation poses for democratic legitimacy is particularly acute in the United States (Lijphart, 1997), where the choice of an unparalleled number of legislative, executive, and judicial officeholders lies in the hands of American voters (Taylor et al., 2014). In stark contrast to the many opportunities for electoral participation, turnout in the United States lags behind the vast majority of advanced democracies (e.g., Pintor and Gratschew, 2002), and most, if not all, non-presidential elections attract less than half of the voting-age population. Especially in less salient local and off-year elections, low participation spells unequal participation, reinforcing concerns for the representativeness of electoral outcomes (Avery and Peffley, 2005; Avery, 2015; Franko et al., 2016; Hajnal and Trounstine, 2005; Hajnal, 2009; Hill and Leighley, 1992; Lijphart, 1997).

The contrast between voting as primary source of democratic legitimacy and inconsequential individual ballots raises the question of whether seemingly small and innocuous changes in voting costs contribute to America's turnout problem, and, if so, if they equally impact voters across socio-economic groups. To shed light on these issues, I examine the effects of an overlooked source of voting costs: sharp changes in distance to the polling place that result from the almost arbitrary shape of precinct boundary lines. Within municipalities, I compare parcels of land and census blocks located in close proximity to borders between adjacent voting precincts. Parcels and blocks on opposite sides of a precinct border share the same school and electoral districts, as well as observationally identical characteristics. Yet, voters living on opposite sides are assigned to vote in

different polling locations, thereby creating plausibly exogenous variation in distance to the polling place.

Because higher voting costs can reduce the likelihood of registration (e.g., dissuading eligible voters from registering and/or through the purging of inactive voters), I aggregate outcomes by parcels and census blocks and implement three empirical designs at these levels. The first specification follows the boundary discontinuity design pioneered by [Black \(1999\)](#). I draw samples of parcels and census blocks from narrow bands around precinct borders and use boundary fixed effects to control for constant correlates of voter participation. Lending credibility to identification, the large sample size allows to focus on units located in extreme proximity (i.e., as little as .05 mile, or 80.5 meters) to precinct borders while preserving good statistical power. Unlike earlier geographic studies (e.g., [Black, 1999](#); [Lavy, 2010](#)) that feature only discontinuous, cross-boundary treatment variation, the distance-to-polling-place treatment varies *both* discontinuously across boundaries *and* continuously within each side of a boundary. Motivated by this observation, the second specification augments boundary fixed effects with local polynomials in latitude and longitude (as in, e.g., [Dell, 2010](#); [Dell et al., 2015](#)). The joint inclusion of boundary fixed effects and latitude-longitude controls isolates the specific effect of sharp treatment changes across boundaries, at the cost of losing some precision to estimate additional parameters. The third and last specification builds on recent work by [Keele et al. \(2014\)](#). I use proximity-based matching to create pairs of matched parcels (and census blocks) that span across precinct borders. Matched pair fixed effects absorb observable and unobservable characteristics that are constant within pairs, while assignment to different polling locations leaves plausibly exogenous variation in distance to polling places.

I apply my empirical designs to 2012, 2013, 2014, and 2016 voter records from nine municipalities in Massachusetts and Minnesota, which I integrate with a vast amount of property-level geographic data (including assessors information, such as land use codes, land and building values, building and residential area), 2010 decennial census data by block and block group (e.g., racial composition of the resident population, median household income), and several electoral and

non-electoral maps (e.g., precincts, polling locations, school and election district boundaries). I focus on a sample of urban municipalities in Massachusetts and Minnesota because of their ready availability of extremely detailed GIS data for properties and electoral maps, and due to the high accessibility of their voter-level information. The choice of this sample is additionally motivated by how voters in the two states could cast their ballots in the elections considered. Whereas Massachusetts voters were required to provide an excuse to vote absentee throughout the sample period, in Minnesota this requirement existed in 2012 and 2013 but was dropped starting from the 2014 midterm election. Thus, I compare changes over time in state-specific impact estimates to see if no-excuse absentee voting attenuates the detrimental effect of distance to the polling place on participation. I find it does not.

I find that a 1-standard deviation (.245 mile) increase in distance to the polling place reduces the number of ballots cast by parcel residents by 2% to 5%, which translates to a 1-to-3 percentage-point decrease in turnout. Turnout effects are paralleled by sizable impacts on voter registration, which suggest that voter-level studies ignoring endogenous registration might suffer from sample selection bias.

Interacted specifications reveal that the negative impact of distance to the polling place is concentrated disproportionately in high-minority areas during non-presidential elections, while no heterogeneity emerges in the 2012 presidential election. Specifically, a 1-mile increase in distance to the polling place reduces the number of ballots cast by parcel residents of high-minority areas by 33, 27, and 19 log points in the 2013, 2014, and 2016 elections, respectively. Corresponding proportional effects in low-minority areas are one third as large (i.e., 11, 6, and 5 log points) and mostly insignificant. A comparable, though less stark contrast emerges from interactions with census block group income and car availability. In a similar vein, although the paucity of Republican voters in my sample rules out clear-cut conclusions, Democrats and unaffiliated voters appear to be more affected by distance to the polling place than registered Republicans in the 2014 and 2016 elections, but not in 2012 and 2013.

Finally, I combine optimization tools from location science with my econometric estimates to

simulate the turnout effects of efficiently redrawing precinct lines. In the interest of realism, I fix polling locations and precinct capacities to their actual values as of the 2012 presidential election. Then, I model the reprecincting problem faced by election administrators as an optimal allocation problem subject to the 2012 polling locations and precinct capacities constraints. Solving the model numerically, I show that the efficient redrawing of precinct lines can increase voter turnout by .2 to .4 percentage points (i.e., between .4% and .7%). In practice, election administrators are expected to have some discretion over the choice of *both* precinct boundaries *and* polling locations. But because I rule out any freedom to choose polling locations, the resulting solution represents an absolute lower bound to the effectiveness of optimal reprecincting. By contrast, a hypothetical policy that erased the impact of distance to the polling place would increase overall turnout by 1.6 to 4 percentage points (i.e., between 4.2% and 7.7%) and reduce participation gaps across low- and high-minority areas in non-presidential elections by 11.4% to 12.8%.

The next section discusses the institutional setting. I detail my empirical strategy in Section 3 and discuss data, summary statistics, and tests of the identification assumption in Section 4. Section 5 presents main effects at the parcel- and census block-levels. Section 6 reports interacted effects by SES, party, and state. Section 7 describes the efficient reprecincting algorithm. Section 8 concludes.

## **2 Institutional Background**

Election administration in the United States is regulated by an intricate network of norms and conventions that are issued or enforced by government at the federal, state, and local levels. While an in-depth review of these practices is beyond the scope of this research, this section summarizes specific aspects of administering American elections that are key to understand my empirical strategy and findings.

## 2.1 Precincts, Census Blocks and Polling Places

In most OECD countries and in twenty U.S. states, the only legitimate way voters can cast their ballots is visiting a predesignated polling place.<sup>1</sup> In these twenty states, precincts determine assignment to Election-Day polling locations, while the same polling place can serve multiple precincts. Precincts represent the basic geographic unit for administering elections in cities and towns across the country.<sup>2</sup> They partition municipalities and constitute the building blocks of every geographic aggregation used for election purposes, including wards, congressional districts and state legislative districts. Unlike larger electoral subdivisions, precinct boundaries are typically drawn by nonpartisan local administrators. This also applies to Massachusetts and Minnesota<sup>3</sup>, the states of my sample municipalities and the subjects of the rest of this section.

By the first statewide election after the federal census, cities and towns in Massachusetts revise precinct boundaries, establish new precincts or consolidate old ones to ensure that extant precincts contain an equal share of the population ( $\pm 5\%$ ), not to exceed 4,000 inhabitants.<sup>4</sup> Except for the lack of a hard-set population ceiling, reprecincting in Minnesota follows a similar process.<sup>5</sup> Precinct boundaries are then left unchanged until the next decennial census. Both states recommend (MN) or outright prescribe (MA) that precincts be bounded by census block boundaries.

Census blocks are the smallest geographic unit used by the U.S. Census Bureau for tabulation of decennial census data from all houses. Because they are bounded on all sides by visible features (e.g., streets, railroads, power lines, streams of water) or nonvisible boundaries (e.g., city limits), urban census blocks typically coincide with city blocks (U.S. Census Bureau, 2010). Their boundaries are not delineated based on population, so census blocks that coincide with parks or empty

---

<sup>1</sup>These twenty states are classified by the National Conference of State Legislatures (NCSL) as requiring an excuse to vote absentee. Some of these states allow forms of early in-person voting. Source: <http://www.ncsl.org/research/elections-and-campaigns/absentee-and-early-voting.aspx> (accessed June 17, 2016).

<sup>2</sup>However, state laws typically exempt the smallest municipalities from splitting in multiple precincts. For example, Massachusetts towns with fewer than 6,200 residents are exempted from dividing into precincts (Massachusetts General Laws – MGL – ch.54 §9).

<sup>3</sup>MGL ch.54 §2 and MN Statutes 204B.14, respectively.

<sup>4</sup>The requirement to establish ward and precinct lines after the decennial census does not apply to the towns of Nantucket and Rockport, and to the city of Boston (1982 Mass. Acts ch. 605, section 3).

<sup>5</sup>State reprecincting guidelines merely suggest that “precincts containing more than 2,000 to 3,000 registered voters can become unwieldy for election administration purposes.” (Secretary of State, 2011).

lots of land have no inhabitants. Still, thanks to the small area they cover, urban census blocks encompass relatively few residents. As of 2010, census blocks in my sample municipalities had an average of 100.1 inhabitants and 82.6 residents aged 18 or older. To preserve confidentiality at this extreme level of disaggregation, only basic demographic data are released at the census block-level; that is, head counts by race, ethnicity, and age groups. Proxies for SES, including income and educational attainment, are nonetheless available by block group.<sup>6</sup>

Polling places can change over time for two reasons. First, a census block (and residents of the parcels therein) can be assigned to a new precinct as result of postcensal reprecincting. Second, a variety of practical or logistical reasons can lead to routine turnover in polling locations without simultaneous changes in precinct boundaries (e.g., closure or renovation of a building that functions as polling site).

## 2.2 Voter List Maintenance

Three pieces of legislation provide the bulk of existing federal requirements for the maintenance of voter lists: the Help America Vote Act (HAVA) of 2002, the National Voter Registration Act of 1993 (NVRA) and the Voting Rights Act (VRA) of 1965. Among others, these laws prescribe a minimum set of requirements that states should follow to maintain current and accurate voter lists. In particular, the NVRA prescribes that a voter can be “purged” from voter rolls if she does not respond to an address confirmation notice *and* fails to vote in two consecutive statewide elections. Interestingly, the U.S. Election Assistance Commission ([U.S. Election Assistance Commission, 2015](#)) reports that inactivity was the leading nationwide cause of voter registration removal between 2012 and 2014, accounting for 35.5% of the 14.8 million voters removed in these years.

With a few procedural differences, both states in my sample purge voters because of electoral inactivity. Massachusetts voters who fail to respond to the annual municipal census for two consecutive years are flagged as “inactive”. Subsequent failure to vote in two consecutive biennial

---

<sup>6</sup>Census blocks are grouped into block groups, which are in turn statistical divisions of census tracts and the smallest geographic aggregation at which the U.S. Census Bureau releases income and schooling data. Block groups typically contain between 600 and 3,000 people.

statewide elections results in removal from voter rolls (MGL ch.51 §37A). Similarly, Minnesota voters who do not vote for four years become inactive and must re-register before voting again (MN Statutes 201.171). Between 2012 and 2014, Massachusetts and Minnesota removed, respectively, 51,452 and 197,667 voters for failure to vote, accounting for 9.2% and 64.3% of the total removals in the two states ([U.S. Election Assistance Commission, 2015](#)).

Failure to vote can thus cause removal from the voter rolls, thereby suggesting that turnout studies based on samples of registered voters might suffer from sample selection bias. This becomes more likely if the putative treatment persists over time (as purging is triggered by failure to vote in multiple elections), if voters exhibit habit formation (because a one-time shock in voting costs can cause abstention in multiple elections), or if registered voters are compared across states that follow different purging procedures. In my context, some voters might not even register to the voter rolls if they anticipate their polling places being too far. This channel seems particularly plausible in Minnesota, where eligible voters have the option of registering directly at the polling station on Election Day. For these reasons, I use the number of votes cast by parcel residents as my main dependent variable, rather than turnout as a percentage of registered voters.

### **2.3 Election Calendars**

The electoral calendars of the two states were remarkably similar in 2012 through 2016. Both held U.S. Senate elections in 2012 and 2014, voted for governor in 2014, held presidential primaries on March 1, 2016, and the most populous cities in the two states (Boston and Minneapolis) both held mayoral elections in November 2013.

In the 2012 presidential election, Massachusetts and Minnesota allowed no form of early voting and required a valid excuse to vote absentee by mail.<sup>7</sup> Thus, the only legitimate way most voters had to cast their ballots was by traveling to their assigned polling places on Election Day.<sup>8</sup> Unlike

---

<sup>7</sup>Valid excuses in MA being: absence on Election Day for any reason, physical disability, or religious beliefs (M.G.L. ch.50 §1; M.G.L. ch.54 §86; M.G.L. ch.54 §89).

<sup>8</sup>Massachusetts state law prescribes stiff penalties for those who make a false absentee ballot application: a fine of up to \$10,000 and up to five years in prison (M.G.L. ch.56 §5). Although it is difficult to assess the extent of illegal absentee voting, practical reasons suggest that to be limited. Most importantly, casting an absentee ballot is far from automatic. An application needs to be mailed or hand-delivered to the elections office before each election. The office



Massachusetts, which required an excuse throughout the sample period, since June 2014 registered voters in Minnesota no longer need an excuse to vote absentee.

## 2.4 Existing Evidence on Distance to the Polling Place and Turnout

A huge literature researches the electoral and post-electoral effects of different “mechanisms of voting”, and its review falls largely beyond the scope of this paper. A smaller literature, to which this paper relates closely, looks specifically at the effects on turnout inequality. Gerber et al. (2013) document that the introduction of all-mail elections in Oregon increased participation by two to four percentage points, an effect that was larger among less active registrants than for frequent voters. On a similar note, Hodler et al. (2013) show that postal voting in Switzerland increased voter turnout, especially among low-educated voters. Interestingly, they also find that postal voting was related to lower government welfare expenditures and lower business taxation. Fujiwara (2015) reports that electronic voting in Brazil fostered the enfranchisement of less educated citizens. In turn, this induced elected officials to increase health care spending, whose primary beneficiaries were the newly enfranchised citizens.

A few papers study the disenfranchising effect of distance to the polling place. Using a sample of 300 precincts in Maryland, Gimpel et al. (2004) uncover a non-linear relationship between distance to polling and the average precinct-level turnout in the 2000 presidential election. Haspel and Knotts (2005) are the first to employ geocoded data at the voter level to regress turnout on distance to the polling place. Using the Atlanta 2001 mayoral election as their testing ground, the authors find that the negative turnout effect of distance to the polling place is concentrated in areas where most residents have no cars. In the 2002 midterm election in Clark County, NV, Dyck and Gimpel (2005) observe that distance to the polling site displays stronger correlation with the method voters choose to cast their ballots (i.e., in-person vs. absentee) than with turnout. A recent paper by Amos et al.

---

proceeds to mail the ballot to the voter, who eventually needs to mail the ballot back in time to be counted. Anecdotal media evidence also highlights how illegal absentee voting appears to be (i) a fairly stigmatized practice, and (ii) mostly concentrated among high-propensity voters; see, e.g., Marty Walsh’s campaign encouraging its staffers to vote absentee ahead of the 2013 Boston mayoral election, as reported by David S. Bernstein. 2015. “Guess How Many of Marty Walsh’s Campaign Staffers Voted Illegally on His Election Day?” *Boston*. June 24. <http://www.bostonmagazine.com/news/blog/2015/06/24/marty-walsh-staffers-voted-illegally/> Accessed August 6, 2015.

(2016) explores the effects of precinct consolidation in Manatee County, Florida, before and after the 2014 midterm election. Controlling for a host of individual-level covariates, the authors find that, relative to other racial groups, Hispanic voters are significantly more likely to abstain from voting as a result of reassignment to a new precinct.

In general, the papers reviewed above are mostly mute about endogeneity concerns, so the estimated effects could be partly due to (potentially) unobserved variables correlating with both distance and turnout. These concerns are taken seriously by [Brady and McNulty \(2011\)](#), who study the consolidation of voting precincts in Los Angeles County ahead of California's 2003 gubernatorial election. Differently from voters in unconsolidated precincts, those affected by the consolidation were assigned to a new polling location for the 2003 election. The authors thus match registered voters in the two groups and find that a 1-mile increase in distance to the polling place reduced in-person voting by as much as 4 percentage points. Alas, estimates gauged on a sample of registered voters might hold limited external validity if distance to the polling place affects voter registration. I substantiate this point in Sections [3.1](#) and [5.2](#).

### **3 Empirical Framework**

Correlational estimates between distance to the polling place and voter turnout incur two potential sources of bias: endogeneity and sample selection. On the one hand, voters who live far from their polling places (e.g., suburban voters) might not be a valid counterfactual for voters who live relatively closer (e.g., inner-city voters). Indeed, the former are likely to be wealthier, more educated and less likely to belong to a minority group, so the raw correlation between distance to the polling place and turnout confounds the causal effect of interest with SES. On the other hand, voter registration could itself be an outcome. Because voter lists are routinely purged of inactive voters, a negative effect on turnout could reduce the likelihood of registration, thereby adding a second source of bias to voter-level correlations. Similarly, some voters might refrain from registering altogether if they anticipate their polling place being too far. This section discusses how I make progress on these two issues.

### 3.1 Level of Analysis

To address the endogeneity of voter registration, the main analysis is conducted at the parcel-level instead of the voter-level. That is, I match every voter with the parcel containing her house and use the total number of votes cast by parcel residents as the main outcome variable. Because parcels are included in the analysis independently of the number of registered voters they contain, this parcel-level sample is robust to so-called “endogenous registration bias”.<sup>9</sup>

In some regressions, I further aggregate the data by census blocks. Whereas there are no estimates of the resident population by parcel, census blocks feature the voting-age population (VAP) as a natural, albeit imperfect (McDonald and Popkin, 2001), denominator of turnout and voter registration. Despite this advantage, census blocks are a coarser geographic aggregation than parcels which could make measurement of distance to the polling place less accurate.

### 3.2 Boundary Fixed Effects

Discontinuous changes in assignment to polling locations across adjacent precincts provide plausibly exogenous treatment variation. As detailed in Section 2, the precinct in which voters live determines where they vote. Parcels within close proximity to each other but located in different precincts should share similar observable and unobservable characteristics, on average. But if they are assigned to different polling locations, distance to the polling place changes discontinuously across the precinct boundary. This motivates a boundary discontinuity design *à la* Black (1999):

$$y_i = \delta_{b(i)} + \beta dist_i + \varepsilon_i, \tag{1}$$

---

<sup>9</sup>Interestingly, the issue of endogenous voter registration in micro-level studies has until recently received limited scholarly attention; see Erikson (1981); Timpone (1998) for two early exceptions, while a recent paper by Nyhan et al. (2015) describes a sensitivity test to gauge the potential bias induced by endogenous registration in voter studies with binary treatments. Nickerson (2015) examines the effects of voter registration campaigns by randomly assigning entire city streets to receive face-to-face visits encouraging voter registration or a control group exposed to no registration information. Similarly to my context, the author implements the analysis by street to make sure that the resulting outcomes (i.e., counts of newly registered voters and the number of ballots they cast) are not affected by endogenous registration bias.

where  $i$  and  $b(i)$  denote, respectively, a generic parcel and the precinct boundary closest to it;  $y_i$  is the number of votes cast by residents of parcel  $i$ ;  $\delta_{b(i)}$  is a full set of precinct boundary fixed effects;  $dist_i$  is distance from parcel  $i$  to its assigned polling place. The same notation extends unchanged to census block regressions, except for  $i$  denoting a generic census block instead of a parcel. I refer to equation 1 as the within-boundary specification.

Notice that boundary fixed effects alone leave two sources of within-boundary treatment variation. One is the abrupt change in distance to the polling place occurring at the boundary, while the other is the continuous variation within each side of the boundary. The presence of two sources of variation contrasts with existing geographic discontinuity designs (e.g., [Black, 1999](#); [Dell, 2010](#); [Dell et al., 2015](#); [Keele et al., 2014](#); [Keele and Titiunik, 2015, 2016](#); [Lavy, 2010](#)), which only feature sharp changes in treatment assignment across boundary sides.

Within-boundary-side variation could be problematic if it correlates with unobserved determinants of voter participation. I address this potential concern in two ways. First, in the hope that boundary fixed effects control for determinants of turnout that change continuously over space, I restrict attention to parcels (or census blocks) that are extremely close to precinct borders. In the analysis that follows, I report balancing tests and results based on samples drawn from three narrow bands around precinct boundaries: .15, .10, and .05 mile (241.4, 160.9, and 80.5 meters, respectively). Second, I show that all results are unchanged when using alternative specifications – detailed in sections 3.3 and 3.4 – that rely exclusively on the discontinuous variation across boundary sides. Unless specified otherwise, standard errors are clustered by boundary throughout.

Equation 1 also requires that distance to the polling place is the only determinant of voter participation that changes discontinuously across precinct boundaries. [Keele and Titiunik \(2015\)](#) call this assumption “compound treatment irrelevance”. A precinct boundary that overlaps the border between school assignment zones is an example of compound treatments. Notice that the mere overlap between precinct and school assignment boundaries is not problematic. In fact, the effect of interest will be confounded only if distance to the polling place correlates with school characteristics across the school boundary and, in their turn, school characteristics also affect voter

participation. For example, the boundary side that is on average assigned to a farther polling location also features better schools, and families who opt for better schools are more likely to turn out. Although it is not obvious how compound treatments could be an issue in the present context, I conservatively exclude precinct boundaries overlapping other institutional or geographic discontinuities. Section 4.2 details these sample restrictions.

Figure 1 illustrates the identification strategy using two precinct boundaries from Cambridge, MA. The small polygons and the thick black lines represent parcels and precinct borders, respectively. Colored parcels are closer than .10 mile to either of two precinct boundaries. Parcels of the same color share (on either side) the same precinct boundary, while different shades of the same color denote relative, within-boundary proximity to the polling place. For example, parcels colored with dark shades of red are closer to their respective polling locations than parcels in light red. Uncolored parcels are excluded from the sample for one or more of the following reasons: they are farther than .10 mile to the nearest precinct boundary, their precinct boundary overlaps other discontinuities<sup>10</sup>, and/or they are non-residential lots.

A visual analysis of the green boundary in Figure 1 reveals the two sources of within-boundary treatment variation. The polling place of parcels on the south (resp., north) side of the boundary is denoted by the green dot at the center (resp., top-center) of the figure. Most parcels on the south side are closer to their polling location than parcels that fall right on the north side, so they display darker shades of green. This is the discontinuous, cross-boundary variation. However, moving eastward along the south side of the boundary increases distance to the polling place relative to other southwest-side parcels, as reflected by lighter shades of green. This is the continuous, within-boundary-side variation.

### 3.3 Boundary Fixed Effects with Latitude-Longitude Interaction

Specification 1 can be modified to rely (almost) exclusively on the discontinuous change in distance to the polling place that occurs at the precinct borders. Following the semi-parametric

---

<sup>10</sup>For example, parcels in the west part of the figure are excluded because their precinct boundary coincides with the Fresh Pond water reservoir.

RD originally described in [Dell \(2010\)](#) and recent work by [Gelman and Imbens \(2014\)](#), I augment regression 1 with boundary-specific linear polynomials in latitude and longitude:

$$y_i = \delta_{b(i)} + \gamma_{b(i)}^{lat} latitude_i + \gamma_{b(i)}^{long} longitude_i + \beta dist_i + \varepsilon_i, \quad (2)$$

where  $\gamma_{b(i)}^{lat}$  and  $\gamma_{b(i)}^{long}$  denote the boundary-specific coefficients on parcel  $i$ 's latitude and longitude, respectively. I refer to equation 2 as the interacted specification. Table A1 shows that the simultaneous inclusion of boundary fixed effects and their linear interaction with latitude and longitude leaves essentially no residual variation in distance to the polling place, except at the discontinuities.

Boundary-specific interactions with latitude and longitude are the RD polynomial, which controls for relevant factors (besides the treatment) that vary smoothly across precinct boundaries. Because of the disaggregated level of analysis, the RD polynomial arguably plays a limited role in my setting compared to existing studies based on some version of equation 2. In my context, all boundaries are shorter than 1 mile and the large sample size allows to restrict attention to parcels located within .05 mile of the nearest precinct boundary. Thus, there is limited geographic space for substantial within-boundary variation of correlates of voter participation other than distance to the polling place. By contrast, [Dell \(2010\)](#); [Dell et al. \(2015\)](#); [Ferwerda and Miller \(2014\)](#); [Fontana et al. \(2016\)](#) compare observations that are several kilometers apart from each other and that are located on either side of boundaries spanning multiple provinces or regions. Corroborating the limited role that the RD polynomial plays in my design, Section 4.3 shows that balancing tests and main results are substantively unaffected by the inclusion of boundary-specific linear polynomials in latitude and longitude.

Because of the larger level of aggregation, the average precinct boundary in the census block sample contains far fewer observations than the average boundary in the parcel-level sample. Thus, to avoid issues of multicollinearity, the census block-counterpart of regression 2 interacts latitude and longitude with city (instead of boundary) fixed effects.

### 3.4 Matching

Following [Keele et al. \(2014\)](#), I use distance-based nearest-neighbor matching (with replacement) as a third method to identify the impact of distance to the polling place. Each residential parcel (denoted by  $i$ ) is matched to the nearest residential parcel (denoted by  $j$ ) that satisfies two conditions:  $j$ 's precinct is assigned to a different polling location than  $i$ 's, and the two precincts are not separated by any of the institutional or geographic discontinuities detailed in Section 4.2. A generic parcel  $i$  and its match  $j$  are called a matched pair (denoted by  $p$ ) henceforth. With this matching sample I estimate equations of the following form:

$$y_{ip} = \delta_p + \beta dist_i + \varepsilon_i, \quad (3)$$

where  $y_{ip}$  denotes that parcels are repeated for all pairs they are part of, and  $\delta_p$  is a full set of matched-pairs fixed effects. Similarly to the boundary discontinuity design, I report results based on samples of parcels within .15, .10, and .05 mile of their matches. Matching distance measures the length of the straight line connecting the centroids of the two parcels within a pair.

Figure 2 shows the sample of matched parcels in the same geographic area of Figure 1. Colored parcels are within .10 mile of their matches. Green parcels share either side of the border between two precincts, and purple parcels do likewise with a different border. Each color appears in two shades, which denote the two sides of a border. Orange lines connect pairs of matched parcels.

A comparison of Figures 1 and 2 reveals that the .10-mile matching sample is limited to parcels that are closer to the nearest precinct border than the .10-mile boundary sample. This is because the former includes parcels that are within .10 mile of the nearest *match*, while the latter includes parcels located within .10 mile of the nearest *border*.

Figure 1 also shows that a particular parcel can appear in multiple pairs. This can occur because the same parcel is included once as a “treated” and once as a “control” unit (as the matching algorithm loops through every parcel), or because multiple parcels have the same match (as matching is with replacement). As discussed in a similar context by [Dube et al. \(2010\)](#), the presence of a single

unit in multiple pairs along a boundary induces mechanical correlation in the residuals across pairs, and potentially along an entire boundary. To address this issue, the authors argue that standard errors should be clustered by boundary, as I already do.<sup>11</sup> Moreover, a few parcels appear in more than one boundary, thereby creating a second source of potential correlation among the residuals (i.e., across boundaries). Thus, to avoid bias in estimation of the standard errors, regressions run on the matching sample rely on two-way clustering by boundary and precinct (Cameron et al., 2011).

The matching procedure and equation 3 apply without changes to the census block sample. However, because of their size, few census blocks are within .05 mile of their matches.<sup>12</sup> Thus, balancing tests and results performed on matched census blocks are based on the following distances instead: any distance between matches, .15 and .10 mile.

## 4 Data

### 4.1 Data Sources

This project relies on three main types of data: voter information, GIS maps, and census data. Municipal election offices and the Minnesota Secretary of State provided lists of registered voters and turnout files for, respectively, eight municipalities in Massachusetts<sup>13</sup> and the city of Minneapolis, MN. As of the 2010 census, these nine municipalities encompassed a total population of more than 1.5 million residents. Separate voter lists, complete with residential address, date of birth, gender, and party affiliation<sup>14</sup>, were collected, along with the respective turnout files, for the 2012 presidential, 2013 municipal, 2014 midterm, and 2016 presidential primary elections.

The sample for the November 4, 2013, municipal elections only includes the cities of Boston, Fall River, Lowell, and Minneapolis. Moreover, the sample for the March 1, 2016, presidential pri-

---

<sup>11</sup>Dube et al. (2010) compare contiguous pairs of counties located in different U.S. states to estimate the effects of minimum wages on earnings and employment. Because the variation they exploit varies at the state level, standard errors based on the county-pair sample are clustered both by state and boundary.

<sup>12</sup>For example, Minneapolis has only four pairs of matched census blocks within .05 mile of each other.

<sup>13</sup>The Massachusetts municipalities are the cities of Boston, Cambridge, Fall River, Lowell, Newton, Quincy, Somerville, and the town of Brookline.

<sup>14</sup>Because Minnesota does not record voters party affiliation, this variable is not available for Minneapolis.



mary is limited to the eight Massachusetts municipalities, since Minnesota featured party caucuses for which the Secretary of State collected no voter-level information. I received the 2014 and 2016 voter lists updated as of Election Day, whereas lists for the 2012 and 2013 elections were requested and obtained between November 2013 and August 2014. Unfortunately, this implies that the 2012 voter lists were already purged of inactive voters who failed to vote in the 2010 and 2012 statewide elections and, more generally, they might differ somewhat from the actual lists used on Election Day.

GIS data come from municipal, county, and state GIS offices. The Massachusetts Office of Geographic Information (MassGIS), the Boston Redevelopment Authority (BRA), and the Hennepin County GIS Office (Hennepin GIS) provided shapefiles of address points and land parcels, along with basic assessors information (e.g., parcel type, lot size, land value, value of buildings, etc.). Shapefiles of school assignment zones, as well as precinct boundaries and polling locations, were obtained from the BRA (Boston), municipal GIS offices (other MA municipalities), and Hennepin GIS (Minneapolis). Finally, I collected maps of State House, State Senate, and Congressional districts from MassGIS and Hennepin GIS.

To link parcels with the most disaggregated census data available, I intersect parcel centroids with 2012 TIGER/Line<sup>®</sup> census block shapefiles. I then use census block identifiers to retrieve: (i) population counts and racial makeup by census block, median household income, the proportion of occupied residential units without a car, and the fraction of high-school noncompleters by block groups.<sup>15</sup>

## 4.2 Sample Construction

Because my analysis is at the parcel level, precisely geocoding voter addresses is crucial to obtain reliable data. In fact, an imprecise address locator<sup>16</sup> could amass groups of geocoded addresses

---

<sup>15</sup>Block-level total population (resp., population 18 and over) by race and ethnicity comes from Table P9 (resp., P11) of the 2010 Federal Census Summary File 1. Block group median household income, the proportion of occupied residential units without a car, and the fraction of high-school noncompleters come from, respectively, Tables B19013, B25044, and B15003 of the 2009-2013 American Community Survey (ACS) 5-year data.

<sup>16</sup>The address locator is the dataset containing address attributes and geographic coordinates (typically, latitude and longitude) and that serves as a crosswalk between addresses and geographic coordinates.

on the same parcel (e.g., consecutive house numbers on the same street) instead of assigning them to their actual, distinct lots. To maximize geocoding accuracy, I use a procedure called “address-point matching”.<sup>17</sup> I start by standardizing voter addresses following the conventions used by MassGIS and Hennepin GIS for their address point shapefiles.<sup>18</sup> To identify the parcels where address points are located, I intersect address points and parcels shapefiles. I then match voters with the intersected address-points/parcels shapefile using address and precinct number. This produces a perfect match for more than 96% of voter addresses. Finally, I geocode unmatched addresses with Esri® ArcGIS 2013 address locator and use Google StreetView to manually review and correct the location of the resulting output. Distances between polygons (e.g., a parcel and a polling place) are computed as the Euclidean, straight-line distance between the polygon centroids. Distances between parcels (or census blocks) and precinct boundaries are computed as the shortest straight-line distance from the parcel (or census block) centroid to the boundary.<sup>19</sup>

Analysis samples satisfy several restrictions. First, samples of parcels are limited to residential lots whose area does not exceed 70,000 square feet.<sup>20</sup> Second, my analysis is restricted to census blocks (and the parcels therein) that had at least one resident at the 2010 decennial census. Boundary discontinuity samples further exclude parcels and blocks whose precinct boundaries span multiple school zones, State House, State Senate, or Congressional districts. I similarly exclude parcels and census blocks assigned to precinct boundaries delineated by ponds, streams of water, highways, railroads, large parks, reservations, cemeteries, and railroads. I also exclude boundaries between precincts assigned to the same polling location. To preserve sample comparability

---

<sup>17</sup>For a review of the superior precision of address-point matching relative to alternative geocoding techniques, see [Zandbergen \(2008\)](#).

<sup>18</sup>For instance, I replace all abbreviations of street types (“ST”, “AVE”, etc.), as well as cardinal prefixes and suffixes (“N”, “S”, “E”, “W”) with their respective spelled-out versions.

<sup>19</sup>Precisely in the context of distance to the polling place, [McNulty et al. \(2009\)](#) argue that Euclidean distance is preferable to more complicated measurement methods (e.g., Manhattan block grid or street distance). All methods examined by the authors display high correlation with one another, with Euclidean distance being easier to compute and interpret.

<sup>20</sup>I determine residential type using land use codes from assessors files. I exclude overly large parcels to avoid the inclusion of huge residential projects and to make sure that distance from parcel centroids to polling places reliably proxies the distance voters face on Election Day. For comparison, an American football field covers an area of 57,600 square feet, inclusive of the two end zones. All results are substantively unaffected by alternative choices of the area threshold or by dropping the threshold altogether.

across elections, I restrict attention to boundaries whose precincts were assigned to vote at the same polling location during every election included in the sample. Finally, samples of census blocks exclude blocks where the number of cast ballots in one or more elections exceeds the 2010 VAP.<sup>21</sup> Similar restrictions apply to matching samples, which are thus limited to residential parcels smaller than 70,000 square feet, census blocks with one or more residents, and precincts that maintained the same polling location over the sample years.

Figure 3 plots the distribution of distance to the polling place in the parcel sample. The average residential parcel has a distance of .356 mile to its polling place, with a standard deviation of .245. Because the sample consists of densely populated urban areas, the overwhelming majority of parcels are assigned to polling locations that are less than .5 mile away.<sup>22</sup>

### 4.3 Summary Statistics and Balancing Exercises

Balance exercises based on predetermined parcel characteristics provide simple tests for the identification assumption described in Section 3. I implement these tests using parcel attributes as outcomes in regressions 1, 2 and 3. If the exploited treatment variation is exogenous, coefficients on distance to the polling place should be small and insignificant. Table 1 reports test results (columns 2 through 11) and summary statistics (column 1). Each cell reports estimates and standard errors from a separate regression. Rows correspond to parcel characteristics. Columns are combinations of regression specifications and samples. Samples for columns 1 and 2 consist of all residential parcels smaller than 70,000 square feet lying in census blocks with one or more residents. Samples for columns 3-11 are subject to the additional parcel and precinct boundary restrictions detailed in Section 4.2.

The average residential parcel measures 6,440 square feet, has 1.74 dwelling units, an external height of 2.06 full stories, 9.79 rooms, and a total building area of 4,295 square feet, of which 2,938

---

<sup>21</sup>These are typically census blocks that contain large residential buildings constructed after 2010 (i.e., the year the decennial census was published).

<sup>22</sup>A regression of distance to the polling place on boundary dummies yields a residual standard deviation of .17 mile. Adding boundary-specific linear polynomials in latitude and longitude reduces the residual standard deviation to .12 mile. Similarly, the residual standard deviation in the full matching sample is approximately .15 mile.

for residential purposes. The mean value of buildings and land are 306 and 175 thousand dollars, respectively, while 71% of the parcels in Boston and Minneapolis are owner occupied.<sup>23</sup>

Column 2 reports coefficients from bivariate regressions of parcel attributes on distance to the polling place measured in miles. Parcels that are relatively farther to polling places are larger and overall less intensively developed than closer ones. Out of nine estimates reported in column 2, eight are significant at the 1% level. By contrast, regressions that control for boundary fixed effects (columns 3 through 5), boundary dummies and their interactions with latitude and longitude (columns 6 through 8), or matched pair fixed effects (columns 9 through 11) appear to do a good job of eliminating correlations between distance to the polling place and parcel characteristics. All regression coefficients are tightly centered around zero, only one is significant at the 5% level, and three are significant at the 10% level. Importantly, their insignificance is not driven by restricting the sample to parcels near precinct boundaries or closer to their matches. Indeed, most standard errors are unaffected by the choice of distance to the nearest boundary or between matched units.<sup>24</sup>

Table 2 reports summary statistics (column 1) and balancing exercises (columns 2 through 11) for census block samples. The average census block has 82.6 residents aged 18 or older, of which 52.3 are non-Hispanic White, 10.8 are non-Hispanic Black, and 8.3 are Hispanic. Thanks to the affluent Boston suburbs of Brookline and Newton, the median household income is about \$71,000 (measured in 2013 inflation-adjusted dollars), which is higher than corresponding figures at the national (\$53,046) and state levels (respectively, \$66,866 and \$59,836 for MA and MN). 19% of occupied residential units in the average census block have no cars and 12% of residents 25 or older never completed high school.

Like for parcel attributes, bivariate regressions of census block characteristics on distance to the

---

<sup>23</sup>Assessors data for Minneapolis only report information on lot size, land value, the value of buildings, and owner occupancy. Other than Minneapolis, owner occupancy is only available for Boston. Boston assessors data do not report the number of residential units, while information on residential area is not available from Fall River and Quincy.

<sup>24</sup>It may seem puzzling that the number of boundaries increases as distance to the boundary decreases (e.g., comparing the number of clusters in columns 4 and 5). This signals that parcels in the .05-mile boundary discontinuity sample are not a subset of those in the .10-mile sample. The reason is that larger bands around a precinct boundary are more likely to intersect multiple school assignment zones, causing all parcels assigned to that boundary to be excluded from the sample. The strong and significant correlation between parcel attributes and distance to the polling place persists in bivariate regressions based on the samples from columns 3-11. Results available upon request.

polling place (column 2) reveal a pattern of significant correlations. Consistent with the intuition that isolated polling locations serve higher-SES populations, greater distance to the polling place is associated with fewer census block residents, higher income, fewer residential units without cars, and a lower concentration of minorities. Reassuringly, none of these correlations persists systematically after controlling for boundary fixed effects (columns 3 through 5), boundary fixed effects and city-specific linear polynomials in latitude and longitude (columns 6 through 8), or using matched pair fixed effects in the matching sample. F-tests of joint significance similarly support the conditional uncorrelation of distance to the polling place.

## 5 Main Results

In this section, I estimate the reduced-form effects of distance to the polling place on voter participation and registration. I start by examining the effect on the number of ballots cast by parcel residents in the 2012 presidential, 2013 municipal, 2014 midterm, and 2016 primary elections. Next, I report impacts on census block turnout and registration, measured as fractions of the census block VAP. The treatment is defined as distance in miles to the polling place throughout.

### 5.1 Effects on Parcel-Level Ballots Cast

Table 3 reports effects on parcel-level voter participation. Each panel corresponds to a different election, and each cell reports estimates from a separate regression. Rows within panels denote estimates from regressions without (row 1) or with (row 2) parcel and census block covariates.<sup>25</sup> Columns 1-3 report estimates from boundary fixed effects specifications, that is, equation 1, which

---

<sup>25</sup>Parcel covariates comprise lot size, value of buildings, land value, owner occupancy, and separate sets of dummies based on the following variables: quartiles of residential and building areas, number of residential units (categories: 1, 2, 3, more than 3), assessed external full-story height (categories: 0, 0 to 1, 1 to 2, 2 to 3, taller than 3; right bounds are included), and number of rooms (categories: 0, 0 to 4, 4 to 7, 7 to 11, more than 11; right bounds are included). Census block and block group covariates are: total population aged 18 or older, percentage of non-White or Hispanic block residents, median household income, percentage of occupied residential units without a car, and percentage of adults aged 25 or older without a high-school diploma.

I report again here in the interest of readability:

$$y_i = \delta_{b(i)} + \beta dist_i + \varepsilon_i.$$

Columns 4-6 reports estimates from interacted specifications, that is:

$$y_i = \delta_{b(i)} + \gamma_{b(i)}^{lat} latitude_i + \gamma_{b(i)}^{long} longitude_i + \beta dist_i + \varepsilon_i.$$

Finally, samples for columns 7-9 are based on pairs of matched parcels. The matching specification takes the following form:

$$y_{ip} = \delta_p + \beta dist_i + \varepsilon_i.$$

Outcomes  $y_i$ 's are spelled out as: "counts of votes cast by residents of parcel  $i$  in a given election".

As shown in Panel A, 2012 parcel-level impact estimates range from  $-.247$  to  $-.427$ . The treatment standard deviation is approximately 1/4 mile, so bounds on estimated effects per standard deviation are  $-.061$  and  $-.107$ . This compares to an average of around two ballots cast per parcel, so a 1-standard deviation increase in distance to the polling place reduces counts of votes cast in 2012 by 3% to 5%. All estimates are significant at the 10% level. They are also robust to the inclusion of covariates and to sample restrictions based on proximity to precinct boundaries or matched units. In terms of precision, simple within-boundary regressions (columns 1-3) dominate more sophisticated alternatives (columns 4-9). This is perhaps not surprising in view of the smaller amount of treatment variation exploited by less parsimonious specifications.

Results from municipal, midterm, and primary elections reveal an overall similar picture. Ranges of impact estimates are  $[-.306, -.087]$ ,  $[-.328, -.107]$ , and  $[-.223, -.013]$  in, respectively, 2013 (Panel B), 2014 (Panel C), and 2016 (Panel D). All estimates are virtually unaffected by the inclusion of parcel and census block controls. Estimated effects per Compared to alternative parametrizations, boundary fixed effects regressions appear to be more robust to the choice of the estimation sample. Matching and within-boundary estimates display comparable magnitudes, while the latter dominate in terms of precision. 17 out of 18 (resp., 13 out of 18) coefficients

estimated using within-boundary (resp., within-matched-pair) variation are significant at the 5% level. Estimates from interacted specifications reveal a noisier and more nuanced picture. Most of these estimates are smaller in magnitude than corresponding coefficients from within-boundary regressions, and, partly because of the large imprecision, only 6 out of 18 coefficients achieve conventional levels of significance.

The choice of sample and specification is somewhat arbitrary, so it is important to show that conclusions are not driven by these choices. To this end, Figure A1 plots estimates from within-boundary, interacted, and matching specifications across values of distance to boundary/match. To fit three specifications without overcrowding the plot, the graph only reports point estimates without confidence intervals. The figure shows that the large negative effect of distance to the polling place does not vary substantively based on the choice of sample and specification. The relative magnitude of the effects across the four elections (roughly proportional to the salience of the election itself) is also largely unaffected by the choice of sample and specification. To give a sense of the noise around estimates, Figure A2 plots coefficients from within-boundary regressions along with 95% confidence intervals. Virtually all estimates support a “statistically significant” effect of distance to the polling place on parcel-level voter participation.

## **5.2 Effects on Census Block Turnout and Registration**

Because parcel-level regressions use ballot counts (instead of turnout rates) as dependent variables, results are potentially sensitive to the presence of outliers (e.g., large parcels with many residents). That estimation samples are restricted to parcels smaller than 70,000 square feet assuages but does not eliminate this concern. A second and perhaps secondary concern with parcel-level regressions is that, even within boundaries or pairs of matched parcels, distance to the polling place correlates negatively with parcel population. As I do not observe counts of voting-eligible parcel residents, this is a possibility I cannot rule out directly. On the one hand, looking at the number of registered voters does not help, because voter registration is itself a potential outcome. On the other hand, Table 1 reveals no partial correlations between distance to the polling place and parcel

characteristics, which would be expected if the VAP was unbalanced within boundaries. Similarly, census block balancing tests reported in Table 2 show no significant partial correlations between distance to the polling place and the VAP.

To decisively rule out both concerns, I now examine census block regressions. Census blocks are larger than parcels, so measurement error is more worrisome at this level of aggregation. However, data from the 2010 decennial census contain the census block VAP, which I use to construct outcomes whose scale of measurement is independent of the resident population. I start by defining census block turnout as the number of votes cast by block residents divided by the VAP. Since I exclude census blocks that, in any of the sample elections, cast more ballots than the 2010 resident VAP, turnout ranges from 0 to 1.

Estimates from block turnout regressions, reported in Table 4, are broadly comparable with corresponding parcel-level coefficients, both in terms of significance and relative (to outcome means) magnitudes. Results are unaffected by the inclusion of block and block group covariates and, for a given specification, by the choice of the estimation sample. Contrary to what happened in parcel-level regressions, latitude-longitude interactions (columns 4-6) never meaningfully affect coefficients from within-boundary regressions (columns 1-3). Of course, this is partly explained by block-level regressions interacting latitude and longitude with city instead of boundary fixed effects.

Depending on the combination of sample and specification, a 1-mile increase in distance to the polling place reduces turnout by 8.4 to 11.9 percentage points in the 2012 presidential (Panel A) and 2.7 to 7.5 percentage points in the 2013 municipal elections (Panel B). Ranges of estimated effects in the 2014 midterm and 2016 primary elections are, respectively,  $[-7.5, -4.7]$  and  $[-4.9, -1.7]$  percentage points (Panels C and D). Relative to average outcomes, census block estimates are broadly consistent with corresponding figures from parcel regressions.<sup>26</sup> This suggests measurement error is not of critical concern at the block level.<sup>27</sup>

---

<sup>26</sup>For example, take the mid-ranges of 2012 block- (i.e.,  $\frac{-.084-.119}{2} = -.101$ ) and parcel-level estimates (i.e.,  $\frac{-.247-.427}{2} = -.337$ ). These two figures compare to, respectively, an average census block turnout of .57 (i.e.,  $\frac{-.10.15}{.57} = -.179$ ) and 2.03 ballots cast per parcel (i.e.,  $\frac{-.337}{2.02} = -.17$ ).

<sup>27</sup>A complementary explanation is that the effects are relatively more pronounced in extremely large parcels. In



Figure A3 assesses the sensitivity of turnout effects to the choice of distance to boundary/match. Estimates from the three specifications appear even less sensitive to the choice of distance to boundary/match than corresponding parcel-level estimates. Across bandwidths, point estimates from within-boundary and interacted specifications are indistinguishable from one another, while matching estimates display slightly smaller magnitudes. Like with parcel-level regressions, relative magnitudes reflect the salience of the four elections, so estimated effects are more pronounced in 2012 than in non-presidential elections. Figure A4 focuses on within-boundary effects and plots point estimates along with 95% confidence intervals. The statistical significance of the estimates is supported by virtually every cutoff of distance to boundary.

Inactive voters are routinely purged from voter rolls. Hence, through its sizable effects on voter participation, distance to the polling place could also reduce the probability that voting-eligible individuals appear on voter lists. Additionally, potential voters might refrain from registering altogether if they know their polling place will be too far. To explore this channel, I use the percentage of registered VAP as outcome for census block regressions. I limit my attention to the 2014 midterm election. I do not report results using 2012 and 2013 registration because data for these elections were obtained in (and are updated as of) early 2014, so inactive voters who did not vote on Election Day might have already been purged. By contrast, 2014 voter rolls are exact copies of the lists used by election officials on Election Day. I also do not use the 2016 primary election, as I do not have electoral data for the Minnesota presidential caucus. Table 5 reports effects on voter registration from boundary fixed effects (columns 1-3), interacted (columns 4-6), and matching specifications (columns 7-9). In the interest of space, I only report coefficients from regressions that include block and block group covariates, although their exclusion leaves results unchanged.

A 1-mile increase in distance to the polling place appears to reduce the fraction of VAP registered to the November 2014 voter rolls by 1.6 (column 3) to 5.2 (column 1) percentage points, on average, which compares to an outcome mean of approximately 81. Taken at face value, Table 5

---

fact, whereas lots larger than 70,000 square feet are excluded from all parcel samples, their census blocks may be used for census block regressions. This explanation is broadly consistent with (1) very large parcels being mostly public housing projects and (2) impact estimates being larger in lower-SES areas (see Section 6).

suggests that roughly two thirds of the turnout effect of distance to the polling place is mediated by a voter registration effect. That is, every three voters dissuaded by distance to the polling place from voting in the 2014 midterm election, about two were purged prior to Election Day or failed to register knowing their polling place would be hard to reach. Matching estimates fall short of statistical significance, but are otherwise similar in magnitude to corresponding coefficients from within-boundary and interacted specifications. Though impact estimates on registration should be taken with caution because of their imprecision, Table 5 provides suggestive evidence of a causal link between the convenience of getting to the polling place and the percentage of the VAP registered to voter rolls. Magnitude of Selection Bias

### 5.3 Magnitude of Selection Bias

Effects on voter registration raise a question about the severity of selection bias in *voter-level* turnout regressions. Section 5 answered this question indirectly by showing that 2014 registration effects at the block level are two thirds as large as corresponding turnout effects.

Table A2 takes a more direct look at endogenous voter registration. Columns 1 and 2 regress the 2014 fraction of *registered* voters who turned out on Election Day on distance to the polling place. Column 1 includes all residential parcels with one or more registered voters and uses no controls other than distance to the polling place. Column 2 focuses on residential parcels within .10 mile to the nearest precinct border (containing at least one registered voter) and controls for boundary fixed effects. For comparison's sake, column 3 reports the estimate from a within-boundary regression of ballots cast by parcel residents on distance to the polling place (i.e., the same estimate reported in Panel C of Table 3).

The estimate in column 1 is affected by both endogeneity – because SES confounds the effect of distance to the polling place – and selection bias – because the relevant regression ignores potential effects on voter registration. The joint severity of these two biases is such that the resulting point estimate is indistinguishable from zero.

By contrast, the regression in column 2 addresses endogeneity (by focusing on parcels close

to precinct boundaries and controlling for boundary fixed effects), but leaves endogenous voter registration unsolved. A 1-mile increase in distance to the polling place is thus found to reduce parcel-level turnout by 3.8 percentage points, which compares to an average turnout of 54%.

In column 3, a 1-mile increase in distance to the polling place is found to reduce the number of votes cast at the parcel level by .215, from an average of 1.43. Estimates from columns 2 and 3 can be compared to each other by dividing for the respective sample means (i.e.,  $-.038/.54 \approx -7\%$  and  $-.215/1.43 \approx -15\%$ ). However, the estimate from column 2 gives the effect on the *registered* population and thus requires appropriate rescaling. As an approximation, I do so by multiplying the relative turnout effect from column 2 (i.e.,  $-.038/.54 \approx -7\%$ ) by the average census block registration rate (i.e., approximately .8). This gives  $-5.6\%$ , which, consistently with the registration effects discussed in Section 5, is surprisingly close to one third the size of the relative effect on the number of votes cast by parcel residents from column 3 (i.e.,  $-15\%$ ).

## 6 Heterogeneous Effects

### 6.1 Effects by Block and Block Group Characteristics

Does ballot box accessibility disproportionately affect lower-SES voters? To answer this question, I explore heterogeneous effects by census block minority presence, block group income, and block group car availability. For each of these proximate measures of SES, I use the sample median to classify parcels into two groups (i.e., above or below the median value).<sup>28</sup> For brevity, I refer to these groups as high- or low-minority/income/car-availability parcels. Since average outcomes likely vary by SES, I estimate proportional effects using the following Poisson fixed effects model:

$$E[y_i|\mathbf{X}_i] = \exp \left( \delta_{b(i)} + \sum_{s \in \{L,H\}} \mathbf{1}(SES_{c(i)} = s) \times (\theta^s + \eta^s dist_i) + X'_{ic(i)} \boldsymbol{\beta} \right), \quad (4)$$

---

<sup>28</sup>Sample medians of the fraction minority, median household income, and car availability are .186, \$62,885, and .131, respectively. Percentage minority is defined as the proportion of non-White or Hispanic residents in the census block.

where  $y_i$  is the number of votes cast by residents of parcel  $i$ ;  $\delta_{b(i)}$  is a full set of precinct boundary fixed effects;  $dist_i$  is distance in miles to the polling place;  $X_{ic(i)}$  denotes parcel  $i$  and census block  $c(i)$  covariates;  $\mathbf{1}(SES_{c(i)} = L)$  and  $\mathbf{1}(SES_{c(i)} = H)$  are dummies identifying, respectively, parcels with lower-than- and higher-than-median-SES proxies; and  $\mathbf{X}_i$  succinctly refers to the whole set of right-hand-side variables.<sup>29</sup> Table 6 reports estimates based on interactions with minority (Panel A), income (Panel B), and car availability (Panel C). Results from alternative parcel-level specifications are reported in Appendix Tables A3 and A4, while heterogeneous turnout effects at the census block are shown in Appendix Tables A5, A6, and A7.

Except for the 2012 presidential election, Panel A reveals significant heterogeneity by minority presence. Effects on high-minority (resp., low-minority) parcels are  $-16.7$ ,  $-33.3$ ,  $-27.2$ , and  $-19.1$  (resp.,  $-15.6$ ,  $-11$ ,  $-6.1$ , and  $-5$ ) log points in the 2012, 2013, 2014, and 2016 elections, respectively. Within-year F-tests of equal effects across low- and high-minority parcels (i.e.,  $H_0 : \eta^H = \eta^L$ ) reject the null hypotheses of identical impacts in the 2013, 2014, and 2016 elections, while no significant difference emerges in 2012. A joint F-test further rejects the null hypothesis that, across all elections, proportional effects are the same across low- and high-minority parcels.

Similarly, estimated proportional effects on low-income parcels have larger magnitudes than corresponding effects on high-income parcels (Table 6, Panel B). However, differences between the two impact estimates are smaller than analogous differences from minority interactions, leading to never reject the null hypothesis of equal effects across low- and high-income parcels. Like with minority interactions, the 2012 presidential election offers no evidence of a disproportionate effect of distance to the polling place on lower-SES/lower-income parcels.

Finally, the pattern of heterogeneous effects by car availability, reported in Panel C of Table 6, is indistinguishable from the one depicted by minority interactions. Except in the 2012 presidential election, parcels in less-motorized areas are disproportionately impacted by distance to the polling place relative to their counterparts with higher access to cars. Differential effects vanish in the

---

<sup>29</sup>Equation 4 is estimated by Poisson conditional maximum likelihood using  $B \sum_{b=1}^B y_{ib(i)}$  as a sufficient statistic for the boundary fixed effects  $\delta_{b(i)}$ . The consistency of the maximum likelihood estimator for  $\beta$  despite the presence of the incidental parameters  $\delta_{b(i)}$  is a special result of the Poisson fixed effects model. For an in-depth review of Poisson fixed effects estimators, see chapter 9 of Cameron and Trivedi (2013).

higher-salience 2012 election, while almost the entire effect in 2016 concentrates among parcels with lower-than-median car availability.

## 6.2 Effects by Party Affiliation

Given the tight relationship between SES and party identification, larger effects in low-SES areas suggest that distance to the polling place could disproportionately affect more liberal voters. I test this hypothesis in my subsample of Massachusetts municipalities. Unlike Minnesota, Massachusetts features partisan voter registration, so every registered voter can be identified as a Republican, Democrat, independent, or third-party voter. Thus, separately for each election, I define three parcel-level outcomes: votes cast by registered Republicans, votes cast by registered Democrats, and votes cast by unaffiliated or third-party voters. In 2016, I also know who participated in the Democratic and Republican primaries, which lets me identify (at least indirectly) the political orientation of unaffiliated voters who turned out on Election Day. To exploit this extra information, outcomes for the 2016 presidential primaries are defined as the number of votes cast in the Republican and Democratic primaries. I then use Poisson equivalents of within-boundary specification 1 to regress these outcomes on distance to the polling place. Table 7 reports the results.

In every election, proportional effects on votes cast by Democrats and unaffiliated/third-party voters share similar magnitude and precision. Their point estimates are roughly 15 log points, implying that a 1-mile increase in distance to the polling place reduces the number of ballots cast by Democrats and unaffiliated/third-party voters by approximately 15%. This contrasts with a small (or even positive, in 2014 and 2016) and mostly insignificant effect on votes cast by Republicans.

Of course, very few voters in urban Massachusetts identify with the Republican party, resulting in only one vote cast by registered Republicans for every 10 cast by Democrats. This ideological imbalance is only partially attenuated in the 2016 election, whose outcomes are defined based on participation in party primaries (for every ballot cast in the Republican primary, there are 5.57 ballots cast in the Democratic primary). It is thus hardly surprising that estimates on votes cast

by Republicans are much noisier than those based on Democratic or unaffiliated voters. With this admittedly important caveat in mind, I can reject equality of effects at the 5% level in 2014 and at the 10% level in 2016, while a joint test of equal proportional effects across the four elections is marginally significant.

### 6.3 Effects by State

Does absentee voting alleviate the negative turnout effect of distance to the polling place?<sup>30</sup> To answer this question, I compare changes over time in Minneapolis-specific impact estimates with corresponding changes in Massachusetts-specific effects. Both Massachusetts and Minnesota required a valid excuse to vote absentee in 2012 and 2013. While Minnesota lifted this requirement in August 2014, Massachusetts retained it throughout 2016. Thus, assuming that changes in the effect of distance to the polling place in the Massachusetts subsample are a valid counterfactual for corresponding changes in Minneapolis, the effect of no-excuse absentee voting can be estimated via a Differences-in-Differences (DD) design. Separately for each election held 2012 through 2014, I estimate Poisson regressions of the following form:

$$E[y_i|\mathbf{X}_i] = \exp\left(\delta_{b(i)} + \gamma^{MA}dist_i + \gamma^{MN}dist_i + X'_{ic(i)}\beta\right), \quad (5)$$

where  $\gamma^{MA}$  and  $\gamma^{MN}$  denote state-specific proportional effects.<sup>31</sup> DD estimates, reported in Table 8, are then computed as  $(\gamma^{MN} - \gamma^{MA})_{14} - (\gamma^{MN} - \gamma^{MA})_{BaselineYr}$ , where subscripts denote election years and *BaselineYr* is either of the two elections (2012 and 2013) in which both states required a valid excuse to vote absentee.

---

<sup>30</sup>Existing evidence on the turnout effects of absentee voting is largely inconclusive. [Karp and Banducci \(2001\)](#) use individual-level data from the National Election Studies to document a small, positive correlation between turnout and the availability of universal absentee voting. Using state-level panel data, [Gronke et al. \(2007\)](#) find no significant correlations between turnout and forms of convenience voting, including no-excuse absentee and early voting. By contrast, a more recent paper by [Larocca and Klemanski \(2011\)](#) detect a positive association between no-excuse absentee voting and turnout in data from the Current Population Survey. [Meredith and Endter \(2015\)](#) document that Texas voters receiving quasi-random stimulation to vote absentee in 2008 remain more likely to vote absentee in 2012. However, equal turnout rates across “stimulated” and “non-stimulated” voters suggest that absentee voting merely replaces in-person voting.

<sup>31</sup>Boundary fixed effects are defined within city, so they already incorporate the states main effects.

Estimated proportional effects in the Massachusetts municipalities are remarkably stable across elections (respectively,  $-17.7$ ,  $-18.1$ , and  $-15$  log points in 2012, 2013, and 2014). By contrast, Minneapolis estimates are larger in lower-salience municipal ( $-35.3$  log points) and midterm ( $-21.5$  log points) elections than in the 2012 presidential election ( $-11.2$  log points). Despite the different magnitudes, I can never reject the hypothesis that, within each year, the effects are the same across the two states.

Proportional effects in Massachusetts are roughly constant across the three elections, while in Minneapolis they are larger in 2013 than in the other years. Thus, signs of DD estimates depend on whether 2012 or 2013 is used as reference year. That is, the Minnesota-minus-Massachusetts difference in 2014 impact estimates (i.e.,  $-.215 + .150 = -.065$ ) is more pronounced than the corresponding 2012 gap (i.e.,  $-.112 + .177 = .065$ ), but less so than the 2013 difference (i.e.,  $-.353 + .181 = -.172$ ). Of course, elections in the two states potentially differed along a number of dimensions (e.g., intensity of party mobilization efforts, coincident ballot measures or minor races) that could have affected the relative salience of the distance effect. Additionally, Minnesota voters had no prior experience with no-excuse absentee voting and little time to learn about its availability. With these caveats in mind, insofar as no-excuse absentee voting does not appear to significantly mitigate the negative effect of distance to the polling place, I find inconclusive evidence of the short-run turnout-enhancing potential of this form of convenience voting.

## **7 Efficient Redrawing of Precinct Boundaries**

Every 10 years, most cities and towns in Massachusetts and Minneapolis draw new precinct lines. They use decennial census data to make sure that every precinct has an equal number of people. Perhaps expecting routine turnover in polling place availability over time, state laws do not explicitly include proximity of polling sites to voters among the objectives of reprecincting. Motivated by this observation, I examine if efficiently redrawing existing precinct lines can reduce distance to the polling place, on average.

I formalize the reprecincting problem faced by election administrators as a generalized assign-

ment problem (GAP, Fernández and Landete, 2015; Kundakcioglu and Alizamir, 2009). In each city, a finite set of census blocks,  $J = \{1, \dots, j, \dots, n\}$ , must be optimally allocated to a finite, predetermined set of polling places,  $I = \{1, \dots, i, \dots, m\}$ . The set of census blocks assigned to a specific polling site constitutes a precinct. Let  $d_j$  denote the service demand of census block  $j \in J$ . Associated with each polling site  $i \in I$ ,  $q_i$  denotes its maximum capacity. For each  $i \in I$  and  $j \in J$ ,  $c_{ij}$  is the cost of serving census block  $j$  through polling place  $i$ . To make the problem realistic and consistent with the regulations discussed in Section 2.1, I make the following assumptions:

1. **Aggregation units:** as implicit in the notation above, precincts must be constructed from aggregations of census blocks.
2. **Polling locations:** polling locations  $i \in I$  are those used in the November 2012 election. If  $x \geq 1$  precincts were assigned to vote at the same polling site, this site appears  $x$  times in the set of facilities. This establishes a one-to-one relationship between polling places and precincts, so I use indistinctly the two terms. It also ensures that the resulting number of precincts  $m$  equals the number of precincts actually used in the 2012 presidential election and all elections thereafter.
3. **Demands weights:** census block  $j$ 's demand,  $d_j$ , is given by the total resident population as of the 2010 decennial census.
4. **Capacity constraints:** the maximum capacity of precinct  $i$ ,  $q_i$ , corresponds to the total population actually assigned to  $i$  after the 2010 decennial reprecincting.
5. **Service costs:** the cost of assigning census block  $j$  to polling station  $i$  is equal to the population-weighted travel distance from block  $j$  to station  $i$ . That is:  $c_{ij} = d_j \times dist(i, j)$ , where  $dist(i, j)$  denotes the  $j$ -to- $i$  distance.

For each combination of block  $j \in J$  and polling station  $i \in I$ , I define the following decision



variable:

$$x_{ij} = \begin{cases} 1 & \text{if census block } j \text{ is assigned to precinct } i \\ 0 & \text{otherwise.} \end{cases}$$

The integer programming formulation for the reprecincting problem is as follows:

$$\text{minimize } \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} \quad (6)$$

$$\text{subject to } \sum_{i \in I} x_{ij} = 1 \quad j \in J \quad (7)$$

$$\sum_{j \in J} d_j x_{ij} \leq q_i \quad i \in I \quad (8)$$

$$x_{ij} \in \{0, 1\} \quad i \in I, j \in J. \quad (9)$$

Constraints 7 and 9 guarantee that each census block is entirely assigned to exactly one precinct, while constraint 8 ensures that precinct capacities are not exceeded. Notice that this setup is quite restrictive. In particular, since I have no direct knowledge of where election administrators might want to locate additional polling sites, if at all, existing polling locations and precinct capacities are taken as given. This creates potentially stringent limits to how much the optimal reprecincting problem can improve on existing precinct boundaries. Overall, I reckon my problem setup to be conservative, in the sense that it privileges realistic assumptions over the achievement of larger, but perhaps infeasible, efficiency gains.

Over the years, numerous approximation algorithms have been proposed for solving the GAP (see Kundakcioglu and Alizamir, 2009 for a review), which is NP-hard. Here, I use Esri® ArcGIS Network Analyst Location-Allocation solver, which relies on a combination of heuristic (Teitz and Bart, 1968) and metaheuristic methods.<sup>32</sup> Column 1 of Table 9 reports the average census block-to-polling-place distance (in miles). Column 2 shows the average difference between

---

<sup>32</sup>The solutions reported here are based on StreetMap North America data and, specifically, on the 2012 vintage of the streets.rs network dataset. For further technical details on the optimization algorithm used by the location-allocation solver, see <http://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/algorithms-used-by-network-analyst.htm> Accessed: June 29, 2016.

distance to the polling place in 2012 and the simulated distance that results from solving the efficient reprecincting problem. Averages are computed over the full census block sample (Panel A), and separately by blocks in areas with below- and above-median values of minority presence (Panel B), income (Panel C), and car availability (Panel D). The remaining columns (3 through 14) are divided in four groups, each representing a different election. Within each group, the first column reports the average census block turnout. The second column shows simulated turnout under efficient reprecincting, while the third details simulated turnout under a benchmark policy that eliminates the effect of distance to the polling place (or, equivalently, that removes distance to the polling place for all blocks).<sup>33</sup>

The good but somewhat discouraging news is that efficiently redrawing precinct lines reduces distance to the polling place by .035 mile, on average (from its 2012 mean of .36 mile). Remember that, in the interest of realism, this estimate is based on extremely conservative assumptions. In practice, one would expect election administrators to jointly determine precinct boundaries, along with the number, location, and capacity of polling sites. By contrast, here I am taking polling site characteristics as given, so the only control variables are the shapes of precinct boundaries. Additionally, I am abstracting from the possibility of assigning larger weights to census blocks that are more likely to benefit from polling place proximity (e.g., high-minority, low-car-availability census blocks). In this sense, .035 mile is an absolute lower bound to the average decrease in distance to the polling place that can result from optimally reprecincting. Interestingly, even such small gain in efficiency would yield small but non-trivial increases in census block turnout; that is, an additional .4 percentage point in 2012 and .2 percentage point in the other elections.

The bad news is that this higher turnout would not narrow participation gaps across high- and low-SES areas. First, the magnitude of the efficiency gain is too small to disproportionately benefit specific subsets of voters. Second, and perhaps more importantly, low-SES voters already live in relative proximity to polling places, which might explain why redrawing precinct lines yields a

---

<sup>33</sup>The simulated turnout effects of the two policies are computed using census block point estimates from Table A5 times the average distances shown in columns 1 and 2 of Table 9. Results are unchanged when I exclude Boston, which, as mentioned in Section 2.1, is exempted from the decennial requirement to redraw precinct lines.

larger decrease in distance to the polling place for voters who live in high-SES areas.

What would turnout and the turnout gap be in a hypothetical world where distance to the polling place has no effect on voter participation? As shown in Table 9, turnout would be considerably higher (+4, +2.3, +2.5, and +1.6 percentage points higher in the four elections, respectively) and, at least in non-presidential elections, the minority and car-availability participation gaps would be smaller. For example, participation gaps between low- and high-minority areas would decrease by 1.7 percentage points in 2013 (starting from 13.3 p.p.), 2 percentage points in 2014 (from 17.6 p.p.), and 1.5 percentage points in 2016 (from 12.3 p.p.). In other words, the disproportionate impact of distance to the polling place in high-minority areas contributes to between 11% and 13% of the participation gap between low- and high-minority areas during non-presidential elections.

## **8 Conclusion**

In a sample of municipalities in Massachusetts and Minnesota, I use a novel, quasi-experimental design based on geographic discontinuities to study the turnout effects of voting costs. I compare parcels and census blocks located in close proximity to boundaries between adjacent voting precincts, which determine assignment to polling places. Geographic units that share (on either side) a precinct boundary also share observationally identical attributes. At the same time, the discontinuous assignment to polling places across boundary sides provides quasi-random treatment variation.

I find that a 1-standard deviation increase in distance to the polling place reduces average turnout by approximately 2% to 5% in the 2012 presidential, 2013 municipal, 2014 midterm, and 2016 primary elections. I also document a negative but imprecise effect on census block voter registration, which suggests that higher voting costs reduce registration directly, by dissuading eligible voters from registering, or indirectly through the removal of inactive voters from voter rolls. During non-presidential elections, the effects of distance to the polling place concentrate disproportionately in high-minority, low-income, and low-car-availability areas, while no differential impact emerges in the higher-salience 2012 election.

Drawing from the location science literature, I discuss a possible algorithm to redraw precinct lines while maintaining voters as close as possible to polling locations. Under very conservative assumptions, the algorithm reduces the average parcel-to-polling-place distance by approximately .03 mile. This would be enough to raise average turnout by .2 to .4 percentage point, but not enough to narrow the turnout gap that separates voters across different socio-economic strata. By contrast, a hypothetical benchmark policy that eliminated distance to the polling place would increase average turnout by 1.6 to 4 percentage points and narrow the turnout gap between low- and high-minority areas in non-presidential elections by as much as 11% to 13%. No zero-cost solution is readily available to erase the negative effects of the inconvenience of casting a ballot. However, the noticeable potential for higher turnout and lower turnout inequality – especially during less salient elections – should be both a memento and a goal for future research on the determinants of voter participation.

## References

- Amos, Brian, Daniel A Smith, and Casey Ste,** “Reprecincting and Voting Behavior,” *Political Behavior*, 2016.
- Avery, James M.,** “Does Who Votes Matter? Income Bias in Voter Turnout and Economic Inequality in the American States from 1980 to 2010,” *Political Behavior*, 2015, 37 (4), 955–976.
- **and Mark Peffley,** “Voter Registration Requirements, Voter Turnout, and Welfare Eligibility Policy: Class Bias Matters,” *State Politics & Policy Quarterly*, 2005, 5 (1), 47–67.
- Black, Sandra E,** “Do Better Schools Matter? Parental Valuation of Elementary Education,” *Quarterly Journal of Economics*, may 1999, 114 (2), 577–599.
- Brady, Henry E and John E McNulty,** “Turning Out to Vote: The Costs of Finding and Getting to the Polling Place,” *Am. Polit. Sci. Rev.*, feb 2011, 105 (1), 115–134.
- Cameron, Colin A. and Pravin K. Trivedi,** *Regression Analysis of Count Data*, 2 edition ed., Cambridge: Cambridge University Press, 2013.
- , **Jonah B. Gelbach, and Douglas L. Miller,** “Robust Inference With Multiway Clustering,” *Journal of Business & Economic Statistics*, 2011, 29 (2), 238–249.
- Dell, Melissa,** “The Persistent Effects of Peru’s Mining Mita,” *Econometrica*, 2010, 78 (6), 1863–1903.
- , **Nathan Lane, and Pablo Querubin,** “State Capacity, Local Governance, and Economic Development in Vietnam,” 2015.
- Dube, Arindrajit, William T. Lester, and Michael Reich,** “Minimum Wage Effects Across State Borders: Estimates Using Contiguous Counties,” *Review of Economics and Statistics*, 2010, 92 (4), 945–964.

- Dyck, Joshua J and James G Gimpel**, “Distance, Turnout, and the Convenience of Voting,” *Social Science Quarterly*, sep 2005, 86 (3), 531–548.
- Erikson, Robert S.**, “Why Do People Vote? Because They Are Registered,” *American Politics Quarterly*, 1981, 9 (3), 259–276.
- Fernández, Elena and Mercedes Landete**, “Location Science,” in Gilbert Laporte, Stefan Nickel, and Francisco Saldanha da Gama, eds., *Location Science*, 2015 editi ed., Springer International Publishing, 2015, chapter 3, p. 644.
- Ferwerda, Jeremy and Nicholas L. Miller**, “Political Devolution and Resistance to Foreign Rule: A Natural Experiment,” *American Political Science Review*, 2014, 108 (3), 642–660.
- Fontana, Nicola, Tommaso Nannicini, and Guido Tabellini**, “Historical Roots of Political Extremism: the Effects of Nazi Occupation of Italy,” 2016.
- Franko, William W., Nathan J. Kelly, and Christopher Witko**, “Class Bias in Voter Turnout and Income Inequality,” *Perspectives on Politics*, 2016, 14 (2), 351–368.
- Fujiwara, Thomas**, “Voting Technology, Political Responsiveness, and Infant Health: Evidence From Brazil,” *Econometrica*, 2015, 83 (2), 423–464.
- Gelman, Andrew and Guido W. Imbens**, “Why High-order Polynomials Should not be Used in Regression Discontinuity Designs,” 2014.
- Gerber, Alan S., Gregory a. Huber, and Seth J. Hill**, “Identifying the Effect of All-Mail Elections on Turnout: Staggered Reform in the Evergreen State,” *Political Science Research and Methods*, 2013, 1 (01), 91–116.
- Gimpel, James G., Joshua J. Dyck, and Daron R. Shaw**, “Registrants, voters, and turnout variability across neighborhoods,” *Political Behavior*, 2004, 26 (4), 343–375.
- Gronke, Paul, Eva Galanes-Rosenbaum, and Peter a. Miller**, “Early Voting and Turnout,” *PS: Political Science & Politics*, 2007, 40 (04), 639–645.

**Hajnal, Zoltan**, *America's Uneven Democracy*, 1 edition ed., New York: Cambridge University Press, 2009.

— **and Jessica Trounstein**, “Where turnout matters: The consequences of uneven turnout in city politics,” *Journal of Politics*, 2005, 67 (2), 515–535.

**Haspel, Moshe and Gibbs H. Knotts**, “Location, Location, Location: Precinct Placement and the Costs of voting,” *Journal of Politics*, may 2005, 67 (2), 560–573.

**Hill, Kim Quaille and Jan E. Leighley**, “The Policy Consequences of Class Bias in State Electorates,” *American Journal of Political Science*, 1992, 36 (2), 351–365.

**Hodler, Roland, Simon Luechinger, and Alois Stutzer**, “The Effects of Voting Costs on the Democratic Process and Public Finances,” *American Economic Journal: Applied Economics*, 2013, 7 (1), 141–171.

**Karp, Jeffrey A and Susan A Banducci**, “Absentee voting, mobilization, and participation,” *American Politics Research*, 2001, 29 (2), 183–195.

**Keele, Luke and Rocío Titiunik**, “Geographic Boundaries as Regression Discontinuities,” *Political Analysis*, 2015, 23 (1), 127–155.

— **and —**, “Natural Experiments Based on Geography,” *Political Science Research and Methods*, 2016, 4 (1), 65–95.

—, —, **and José R. Zubizarreta**, “Enhancing a geographic regression discontinuity design through matching to estimate the effect of ballot initiatives on voter turnout,” *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 2014.

**Kundakcioglu, Erhun O. and Saed Alizamir**, “Generalized Assignment Problem,” in Christodoulos A. Floudas and Panos M. Pardalos, eds., *Encyclopedia of Optimization*, Springer US, 2009, pp. 1153–1162.

- Larocca, R. and J. S. Klemanski**, “U.S. State Election Reform and Turnout in Presidential Elections,” *State Politics & Policy Quarterly*, 2011, *11* (1), 76–101.
- Lavy, Victor**, “Effects of Free Choice Among Public Schools,” *Review of Economic Studies*, 2010, *77* (3), 1164–1191.
- Lijphart, Arend**, “Unresolved Dilemma Unequal Participation : Democracy ’ s Presidential Address , American Political Science Association ,,” *The American Political Science Review*, 1997, *91* (1), 1–14.
- McDonald, Michael P. and Samuel L. Popkin**, “The Myth of the Vanishing Voter,” *American Political Science Review*, 2001, *95* (4), 963–974.
- McNulty, J E, Conor M Dowling, and Margaret H Ariotti**, “Driving Saints to Sin: How Increasing the Difficulty of Voting Dissuades Even the Most Motivated Voters,” *Political Analysis*, sep 2009, *17* (4), 435–455.
- Meredith, Marc and Zac Endter**, “Aging into Absentee Voting: Evidence from Texas,” 2015.
- Nickerson, David W.**, “Do Voter Registration Drives Increase Participation? For Whom and When?,” *The Journal of Politics*, 2015, *77* (1), 88–101.
- Nyhan, Brendan, Christopher Skovron, and Rocío Titiunik**, “Differential Registration Bias in Voter File Data: A Sensitivity Analysis Approach,” 2015.
- Pintor, Rafael López and Maria Gratschew**, “Voter Turnout Since 1945: A Global Report,” Technical Report, International Institute for Democracy and Electoral Assistance (International IDEA), Stockholm, Sweden 2002.
- Secretary of State, MN**, “2011 Minnesota Redistricting Guide,” 2011.
- Taylor, Steven L., Matthew Soberg Shugart, Arend Lijphart, and Bernard Grofman**, *A Different Democracy: American Government in a 31-Country Perspective*, New Haven & London: Yale University Press, 2014.



**Teitz, Michael B. and Polly Bart**, “Heuristic Methods for Estimating the Generalized Vertex Median of a Weighted Graph,” *Operations Research*, oct 1968, *16* (5), 955–961.

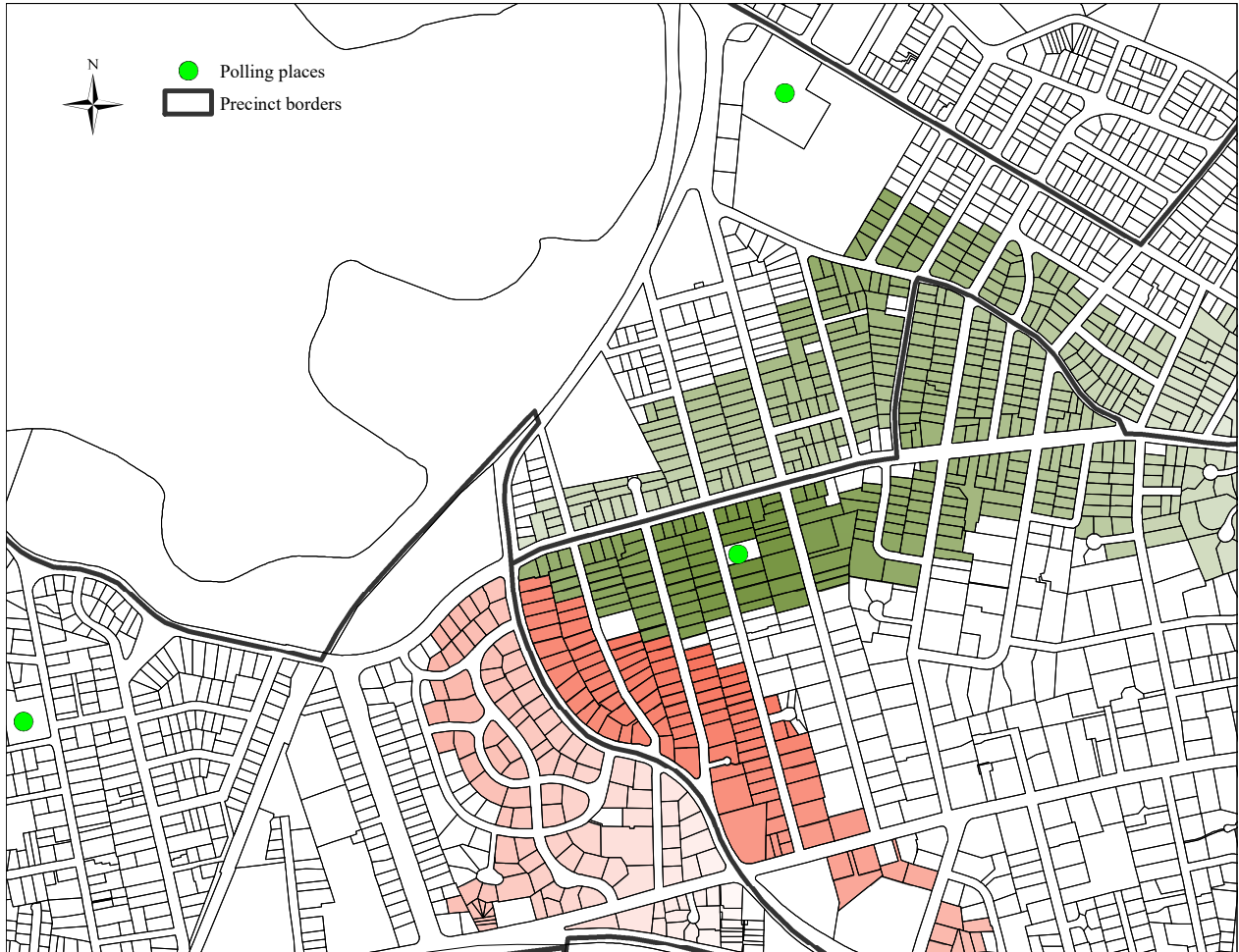
**Timpone, Richard J**, “Structure, behavior, and voter turnout in the United States,” *American Political Science Review*, 1998, *92* (1), 145–158.

**U.S. Census Bureau**, “Census 2010 Geographic Terms and Concepts,” 2010.

**U.S. Election Assistance Commission**, “The EAC 2014 Election Administration and Voting Survey Comprehensive Report,” Technical Report, U.S. Election Assistance Commission 2015.

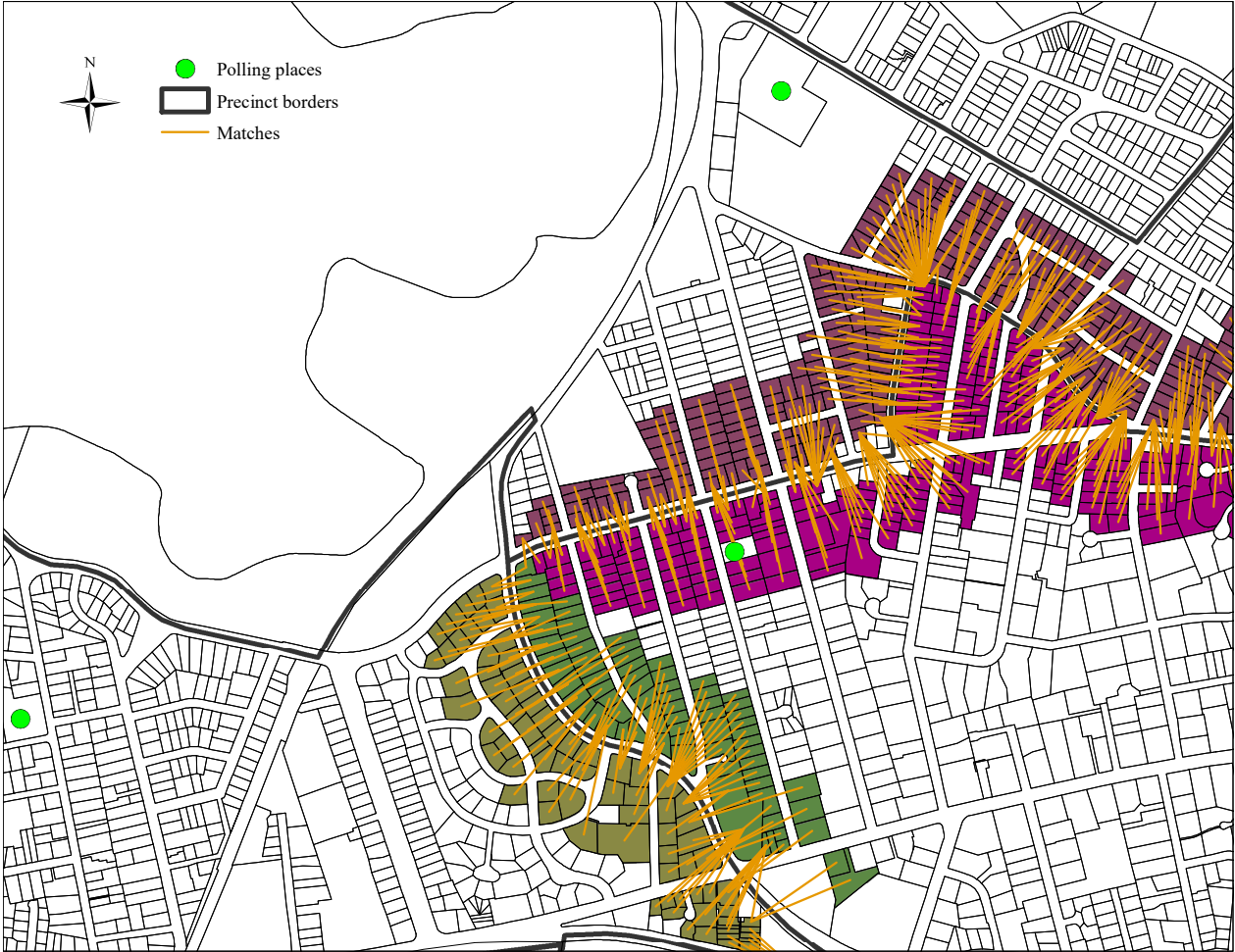
**Zandbergen, Paul A.**, “A comparison of address point, parcel and street geocoding techniques,” *Computers, Environment and Urban Systems*, 2008, *32* (3), 214–232.

Figure 1: Example of Boundary Fixed Effects



Notes: This map illustrates the within-boundary identification strategy using two precinct boundaries from Cambridge, MA. The small polygons and the thick black lines represent parcels and precinct borders, respectively. Colored parcels are closer than .10 mile to either of two precinct boundaries. Parcels of the same color share (on either side) the same precinct boundary, while different shades of the same color denote relative, within-boundary proximity to the polling place.

Figure 2: Example of Matching



This map shows the sample of matched parcels in the same geographic area of Figure 1. Colored parcels are within .10 mile of their matches. Green parcels share either side of the border between two precincts, and purple parcels do likewise with a different border. Each color appears in two shades, which denote the two sides of a border. Orange lines connect pairs of matched parcels.

Figure 3: Distance to the Polling Place

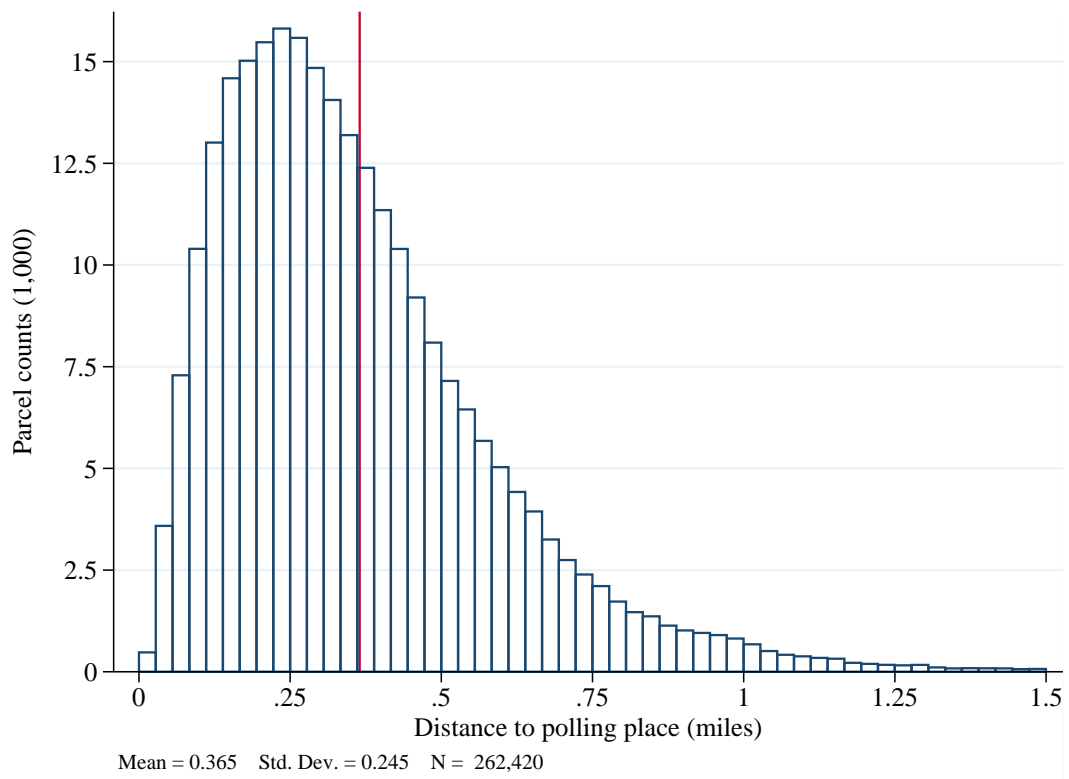


Table 1: Parcel Summary Statistics and Covariate Balance

Dist. to Boundary/Match:	Sample	Specification:									
	Mean	No controls	Boundary FEs			Bound. FEs+Lat-Long Inter.			Matched Pair FEs		
	Any (1)	Any (2)	<.15 mi (3)	<.10 mi (4)	<.05 mi (5)	<.15 mi (6)	<.10 mi (7)	<.05 mi (8)	<.15 mi (9)	<.10 mi (10)	<.05 mi (11)
Lot size (sq ft)	6,440	4,399 ** (509)	331 (228)	337 (207)	32 (196)	188 (374)	389 (369)	-9 (297)	105 (245)	-62 (229)	87 (214)
Building area (sq ft)	4,295	-658 ** (117)	-72 (144)	-164 (140)	-223 (151)	-152 (175)	-108 (164)	-135 (167)	-30 (202)	-51 (207)	-11 (201)
Residential area (sq ft)	2,938	-749 ** (155)	-21 (178)	-131 (177)	-35 (218)	-287 (229)	-325 (250)	-177 (247)	18 (272)	-7 (287)	-54 (290)
Value of buildings (\$1K)	306	-148 ** (30)	-10 (17)	-17 (18)	1 (25)	-17 (21)	-20 (21)	-16 (31)	-17 (35)	-25 (37)	-24 (41)
Value of land (\$1K)	175	-13 (18)	7 (6)	9 (7)	5 (8)	3 (7)	7 (7)	6 (7)	2 (7)	-1 (7)	4 (8)
Units	1.74	-.94 ** (.12)	-.07 (.11)	-.12 (.13)	-.09 (.15)	-.10 (.14)	-.09 (.15)	-.005 (.157)	.02 (.20)	-.07 (.22)	-.05 (.21)
Stories	2.06	-.59 ** (.05)	.03 (.04)	.02 (.03)	.01 (.03)	-.04 (.05)	.002 (.046)	.03 (.05)	.06 (.04)	.05 (.04)	.03 (.04)
Rooms	9.79	-2.22 ** (.28)	-.50 (.30)	-.55 ~ (.33)	-.37 (.33)	-.56 (.38)	-.42 (.40)	-.55 (.44)	-.22 (.50)	-.66 (.55)	-.64 (.56)
Owner occupied (fraction)	.71	.16 ** (.03)	-.02 (.03)	-.01 (.02)	-.003 (.027)	-.05 * (.02)	-.06 ~ (.03)	-.03 (.04)	.01 (.03)	.02 (.03)	.06 (.04)
F-test		38.56	1.41	1.30	.79	1.08	1.21	.57	.57	.86	.79
p		.00	.18	.23	.62	.37	.28	.82	.83	.56	.62
Clusters		1,323	374	383	388	374	383	388	438	432	420
N	262,420	262,420	71,581	59,990	36,061	71,581	59,990	36,061	192,480	133,644	57,128

Notes: This table describes the parcel samples. Column 1 provides summary statistics. Each cell in columns 2 through 11 reports estimates from a separate regression on distance to the polling place, measured in miles. Estimates reported in column 1 are from bivariate regressions of parcel characteristics (in rows) on distance to the polling place. Columns 3 through 5 control for boundary fixed effects. Columns 6 through 8 also control for boundary-specific linear interactions with latitude and longitude. The specification in columns 9 to 11 includes matched-pair fixed effects. Standard errors are clustered by precinct boundary (separately by precinct boundary and precinct in columns 9-11) and reported in parentheses.

\*\* p < 0.01, \* p < 0.05, ~ p < 0.10

Table 2: Census Block Summary Statistics and Covariate Balance

Dist. to Boundary/Match:	Sample	Specification:									
	Mean	No controls	Boundary FEs			Bound. FEs+Lat-Long Inter.			Matched Pair FEs		
	Any (1)	Any (2)	<.15 mi (3)	<.10 mi (4)	<.05 mi (5)	<.15 mi (6)	<.10 mi (7)	<.05 mi (8)	Any (9)	<.15 mi (10)	<.10 mi (11)
Adult population	82.6	-40.3 ** (5.2)	-12.5 (7.7)	-8.0 (7.3)	-5.6 (7.7)	-14.9 ~ (8.0)	-9.0 (7.7)	-9.3 (8.5)	-1.0 (6.4)	-8.3 (7.2)	-16.3 (11.6)
Non-Hispanic Whites	52.3	-19.6 ** (3.6)	-6.3 (5.1)	-4.7 (4.9)	-2.1 (4.8)	-7.0 (5.1)	-4.0 (5.0)	-3.5 (5.1)	-0.4 (3.9)	-2.0 (4.2)	-3.7 (5.5)
Non-Hispanic Blacks	10.8	-7.7 ** (1.3)	-3.4 (2.1)	-1.5 (1.9)	-2.2 (2.6)	-3.9 ~ (2.2)	-2.0 (2.0)	-3.0 (2.8)	-1.3 (1.9)	-3.3 (2.5)	-7.7 ~ (4.3)
Hispanics, all races	8.3	-7.4 ** (1.0)	-1.3 (1.5)	-.9 (1.4)	-1.5 (1.8)	-1.9 (1.5)	-1.4 (1.5)	-2.0 (1.9)	-.4 (1.1)	-2.0 (1.5)	-3.4 (2.4)
Nonwhites/Hispanics (fraction)	.32	-.10 ** (.02)	.02 (.02)	.02 (.02)	.03 (.03)	.01 (.01)	.01 (.02)	.02 (.02)	-.001 (.017)	-.005 (.016)	.01 (.02)
Median HH income (\$1K)	71.0	13.8 ** (4.6)	1.4 (3.4)	1.3 (3.2)	3.6 (3.3)	2.6 (3.3)	2.3 (2.9)	3.9 (2.6)	-.2 (3.6)	5.0 ~ (2.7)	3.5 (2.9)
Units w/o cars (fraction)	.19	-.15 ** (.02)	.002 (.017)	.002 (.016)	-.01 (.02)	-.003 (.016)	-.004 (.016)	-.02 (.02)	-.02 (.01)	-.02 (.01)	-.01 (.02)
HS noncompleters (fraction)	.12	-.01 (.01)	.01 (.01)	.01 (.01)	.01 (.01)	.004 (.014)	.01 (.01)	.01 (.01)	-.01 (.01)	.003 (.011)	.01 (.01)
F-test		15.18	.91	.51	.54	.93	.49	.74	.45	.85	1.12
p		.00	.51	.85	.82	.49	.86	.65	.89	.56	.35
Clusters		1,296	354	354	290	354	354	290	443	437	374
N	15,098	15,068	3,872	3,339	1,705	3,872	3,339	1,705	20,990	8,770	4,130

Notes: This table replicates the covariate balance tests from Table 1 on census block samples. Differently from Table 1, columns 6 through 8 control for interactions between latitude and longitude and municipality (instead of boundary) fixed effects.

\*\* p < 0.01, \* p < 0.05, ~ p < 0.10

Table 3: Effects on Parcel-Level Votes Cast

Dist. to Bound./Match:		Specification:								
		Boundary FEs			Bound. FEs+Lat-Long Inter.			Matched Pair FEs		
		<.15 mi	<.10 mi	<.05 mi	<.15 mi	<.10 mi	<.05 mi	<.15 mi	<.10 mi	<.05 mi
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<u>A. Votes Cast in 2012</u>										
Without covariates	(1)	-.324 ** (.096)	-.296 ** (.097)	-.340 ** (.121)	-.281 * (.132)	-.247 ~ (.146)	-.408 * (.193)	-.402 ** (.143)	-.412 ** (.148)	-.373 * (.160)
With covariates	(2)	-.372 ** (.099)	-.324 ** (.090)	-.330 ** (.109)	-.326 * (.128)	-.318 * (.133)	-.405 * (.183)	-.427 ** (.118)	-.391 ** (.117)	-.400 ** (.127)
Mean dep. var.		2.00	2.03	2.12	2.00	2.03	2.12	2.25	2.27	2.24
N		71,581	59,990	36,061	71,581	59,990	36,061	192,480	133,644	57,128
<u>B. Votes Cast in 2013</u>										
Without covariates	(1)	-.246 ** (.046)	-.188 ** (.049)	-.198 ** (.060)	-.149 * (.072)	-.087 (.080)	-.178 (.116)	-.294 ** (.079)	-.243 ** (.079)	-.150 ~ (.082)
With covariates	(2)	-.306 ** (.067)	-.209 ** (.041)	-.191 ** (.051)	-.217 * (.087)	-.114 ~ (.064)	-.167 (.102)	-.295 ** (.073)	-.234 ** (.070)	-.162 * (.073)
Mean dep. var.		.99	1.01	1.05	.99	1.01	1.05	1.02	1.04	1.08
N		55,719	45,705	27,292	55,719	45,705	27,292	139,052	96,084	39,892

(Continues)

		Specification:								
		Boundary FEs			Bound. FEs+Lat-Long Inter.			Matched Pair FEs		
Dist. to Bound./Match:		<.15 mi	<.10 mi	<.05 mi	<.15 mi	<.10 mi	<.05 mi	<.15 mi	<.10 mi	<.05 mi
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<u>C. Votes Cast in 2014</u>										
Without covariates	(1)	-.219 ** (.057)	-.215 ** (.061)	-.199 ** (.076)	-.156 * (.079)	-.107 (.087)	-.172 (.130)	-.305 ** (.108)	-.294 ** (.112)	-.188 (.117)
With covariates	(2)	-.245 ** (.058)	-.231 ** (.055)	-.194 ** (.069)	-.172 * (.069)	-.138 ~ (.076)	-.164 (.117)	-.328 ** (.086)	-.282 ** (.084)	-.212 * (.091)
Mean dep. var.		1.41	1.43	1.47	1.41	1.43	1.47	1.54	1.55	1.53
N		71,581	59,990	36,061	71,581	59,990	36,061	192,480	133,644	57,128
<u>D. Votes Cast in 2016</u>										
Without covariates	(1)	-.168 ** (.059)	-.151 * (.067)	-.149 ~ (.078)	-.077 (.073)	-.018 (.080)	-.044 (.096)	-.184 ~ (.110)	-.152 (.110)	-.067 (.113)
With covariates	(2)	-.177 ** (.052)	-.145 ** (.056)	-.147 * (.069)	-.052 (.056)	-.013 (.061)	-.050 (.087)	-.223 ** (.083)	-.171 * (.077)	-.108 (.075)
Mean dep. var.		1.38	1.39	1.43	1.38	1.39	1.43	1.54	1.55	1.51
N		49,447	42,939	27,047	49,447	42,939	27,047	140,128	99,082	45,326

Notes: This table reports estimates from regressions of counts of votes cast by parcel residents on distance to the polling place. Panels represent outcomes defined by different elections. Each column represents a different combination of sample and specification. Each cell reports results from separate regressions, which exclude (row 1 within each panel) or include (row 2 within each panel) parcel and census block covariates.

\*\* p < 0.01, \* p < 0.05, ~ p < 0.10



Table 4: Effects on Census Block Turnout

		Specification:								
		Boundary FEs			Bound. FEs+Lat-Long Inter.			Matched Pair FEs		
Dist. to Bound./Match:		<.15 mi	<.10 mi	<.05 mi	<.15 mi	<.10 mi	<.05 mi	Any	<.15 mi	<.10 mi
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<u>A. Census Block Turnout in 2012</u>										
Without covariates	(1)	-.118 ** (.030)	-.115 ** (.029)	-.119 ** (.037)	-.113 ** (.031)	-.111 ** (.031)	-.113 * (.044)	-.074 * (.035)	-.070 * (.030)	-.084 * (.040)
With covariates	(2)	-.115 ** (.030)	-.112 ** (.030)	-.118 ** (.038)	-.114 ** (.031)	-.111 ** (.031)	-.114 * (.044)	-.075 * (.034)	-.073 * (.029)	-.087 * (.040)
Mean dep. var.		.56	.57	.54	.56	.57	.54	.56	.55	.53
N		3,872	3,339	1,705	3,872	3,339	1,705	20,990	8,770	4,130
<u>B. Census Block Turnout in 2013</u>										
Without covariates	(1)	-.075 ** (.016)	-.063 ** (.021)	-.066 * (.031)	-.072 ** (.015)	-.059 ** (.020)	-.062 * (.031)	-.047 ** (.015)	-.034 (.023)	-.027 (.029)
With covariates	(2)	-.074 ** (.016)	-.062 ** (.021)	-.064 * (.030)	-.073 ** (.016)	-.061 ** (.021)	-.063 * (.031)	-.045 ** (.015)	-.039 ~ (.024)	-.031 (.029)
Mean dep. var.		.30	.30	.28	.30	.30	.28	.29	.29	.28
N		2,953	2,552	1,233	2,953	2,552	1,233	15,112	6,594	2,938

49

(Continues)

		Specification:								
		Boundary FEs			Bound. FEs+Lat-Long Inter.			Matched Pair FEs		
Dist. to Bound./Match:		<.15 mi	<.10 mi	<.05 mi	<.15 mi	<.10 mi	<.05 mi	Any	<.15 mi	<.10 mi
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
<u>C. Census Block Turnout in 2014</u>										
Without covariates	(1)	-.073 ** (.018)	-.075 ** (.018)	-.056 * (.028)	-.072 ** (.015)	-.072 ** (.015)	-.060 ** (.020)	-.058 * (.023)	-.047 * (.022)	-.047 (.029)
With covariates	(2)	-.070 ** (.016)	-.072 ** (.017)	-.054 * (.026)	-.071 ** (.014)	-.072 ** (.016)	-.061 ** (.019)	-.058 ** (.022)	-.049 * (.022)	-.047 (.029)
Mean dep. var.		.41	.41	.38	.41	.41	.38	.41	.39	.37
N		3,872	3,339	1,705	3,872	3,339	1,705	20,990	8,770	4,130
<u>D. Census Block Turnout in 2016</u>										
Without covariates	(1)	-.049 ** (.019)	-.044 ** (.017)	-.049 ~ (.028)	-.046 ** (.016)	-.038 * (.015)	-.048 * (.022)	-.047 * (.021)	-.026 (.020)	-.019 (.023)
With covariates	(2)	-.049 ** (.017)	-.045 ** (.016)	-.045 ~ (.025)	-.048 ** (.015)	-.041 ** (.015)	-.046 * (.021)	-.048 * (.020)	-.026 (.020)	-.017 (.022)
Mean dep. var.		.35	.35	.34	.35	.35	.34	.35	.34	.33
N		2,783	2,376	1,415	2,783	2,376	1,415	14,104	6,134	3,334

Notes: This table reports estimates from regressions of census block turnout on distance to the polling place. Panels represent outcomes defined by different elections. Each column represents a different combination of sample and specification. Each cell reports results from separate regressions, which exclude (row 1 within each panel) or include (row 2 within each panel) parcel and census block covariates.

\*\* p < 0.01, \* p < 0.05, ~ p < 0.10

Table 5: Effects on 2014 Census Block Voter Registration

Dist. to Bound./Match:	Specification:								
	Boundary FEs			Bound. FEs+Lat-Long Inter.			Matched Pair FEs		
	<.15 mi	<.10 mi	<.05 mi	<.15 mi	<.10 mi	<.05 mi	Any	<.15 mi	<.10 mi
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
Distance to polling place	-.052 *	-.049 *	-.021	-.049 *	-.044 *	-.016	-.038	-.035	-.046
	(.022)	(.020)	(.030)	(.022)	(.021)	(.030)	(.029)	(.029)	(.038)
Mean dep. var.	.81	.82	.81	.81	.82	.81	.80	.80	.79
Observations	3,872	3,339	1,705	3,872	3,339	1,705	20,990	8,770	4,130

Notes: This table reports estimates from regressions of 2014 census block registration on distance to the polling place. Registration is defined as the number of registered voters in the census block as of the 2014 midterm election divided by the adult population as of the 2010 federal census.

\*\* p < 0.01, \* p < 0.05, ~ p < 0.10

Table 6: Heterogeneous Effects by Census Characteristics

Election:	2012 Presidential		2013 Municipal		2014 Midterm		2016 Primary	
	Mean (1)	Effect (2)	Mean (3)	Effect (4)	Mean (5)	Effect (6)	Mean (7)	Effect (8)
<u>A. By % Minority</u>								
% minority $\leq$ median	1.88	-.156 * (.079)	1.04	-.110 * (.049)	1.42	-.061 (.044)	1.33	-.050 (.048)
% minority $>$ median	2.16	-.167 ** (.047)	.97	-.333 ** (.066)	1.43	-.272 ** (.048)	1.45	-.191 ** (.055)
F-test (within year)		.02		8.33		13.8		5.29
p		.89		.00		.00		.02
F-test (across years)		3.96						
p		.00						
N	59,990	59,990	45,705	45,705	59,990	59,990	42,939	42,939
<u>B. By Median HH Income</u>								
Income $\leq$ median	1.98	-.144 ** (.045)	.88	-.268 ** (.059)	1.28	-.205 ** (.055)	1.20	-.173 ** (.067)
Income $>$ median	2.08	-.178 * (.075)	1.17	-.184 ** (.053)	1.56	-.134 ** (.041)	1.54	-.091 * (.042)
F-test (within year)		.21		1.65		1.58		1.66
p		.65		.20		.21		.20
F-test (across years)		.97						
p		.42						
N	59,990	59,990	45,705	45,705	59,990	59,990	42,939	42,939
<u>C. By % Units w/o Cars</u>								
% w/o cars $\leq$ median	1.67	-.160 * (.077)	.92	-.137 * (.054)	1.29	-.093 * (.043)	1.20	-.019 (.042)
% w/o cars $>$ median	2.32	-.162 ** (.047)	1.08	-.313 ** (.063)	1.53	-.250 ** (.052)	1.51	-.238 ** (.058)
F-test (within year)		.00		4.83		6.38		13.42
p		.98		.03		.01		.00
F-test (across years)		4.91						
p		.00						
N	59,990	59,990	45,705	45,705	59,990	59,990	42,939	42,939

Notes: This table reports estimates from Poisson, boundary fixed effects regressions that interact distance to the polling place with dummies for lower- and higher-than-median values of census block minority presence (Panel A), census block group median income (Panel B), and block group percentage of residential units without cars (Panel C). The null hypothesis of within-year F-tests is that the effect of distance to the polling place is the same across parcels with higher-than-median and lower-than-median values of the interacting characteristic. The null hypothesis of across-years F-tests is that the effects are identical in every election.

\*\* p < 0.01, \* p < 0.05, ~ p < 0.10

Table 7: Heterogeneous Effects by Voter Party Identification

Party Affiliation/Primary:	Election: 2012 Presidential		2013 Municipal		2014 Midterm		2016 Primary	
	Mean	Effect	Mean	Effect	Mean	Effect	Mean	Effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Republican	.12	-.221 ~ (.121)	.06	-.164 (.148)	.08	.114 (.124)	.21	.024 (.085)
Democratic	1.25	-.162 ** (.055)	.73	-.164 ** (.051)	.87	-.184 ** (.048)	1.17	-.158 ** (.051)
Unaffiliated	.84	-.191 * (.075)	.41	-.155 * (.073)	.55	-.130 * (.056)		
F-test (within year)		.23		.01		3.36		3.26
p		.79		.99		.03		.07
F-test (across years)		1.95						
p		.06						
N	42,939	42,939	28,660	28,660	42,939	42,939	42,939	42,939

Notes: Each cell reports estimates from a separate Poisson, boundary fixed effect regression estimated on the subsample of Massachusetts parcels. Outcomes in columns 1-6 are defined as the number of votes cast by parcel residents of a given partisan affiliation. Outcomes in columns 7 and 8 are the number of votes cast by parcel residents in a given presidential primary. The null hypothesis of within-year F-tests is that proportional effects are identical across party affiliations/primaries. The null hypothesis of across-years F-tests is that the effects are identical in every election.

\*\* p < 0.01, \* p < 0.05, ~ p < 0.10

Table 8: Heterogeneous Effects by State

Election:	2012 Presidential		2013 Municipal		2014 Midterm	
	Mean (1)	Effect (2)	Mean (3)	Effect (4)	Mean (5)	Effect (6)
Massachusetts - $\gamma^{MA}$	2.21	-.177 ** (.060)	1.20	-.181 ** (.047)	1.50	-.150 ** (.044)
Minnesota - $\gamma^{MN}$	1.58	-.112 (.085)	.68	-.353 ** (.100)	1.23	-.215 ** (.075)
F-test (within year)		.388		2.505		.560
p		.533		.113		.454
$(\gamma^{MN} - \gamma^{MA})_{14} - (\gamma^{MN} - \gamma^{MA})_{12}$		-.130 (.085)				
$(\gamma^{MN} - \gamma^{MA})_{14} - (\gamma^{MN} - \gamma^{MA})_{12}$		.107 (.084)				
N	59,990	59,990	45,705	45,705	59,990	59,990

Notes: This table reports estimates from Poisson, boundary fixed effects regressions that interact distance to the polling place with state dummies. The null hypothesis of within-year F-tests is that the effects are identical across states. DD estimates of the effect of no-excuse absentee voting are reported below within-year F-tests.

\*\*  $p < 0.01$ , \*  $p < 0.05$ , ~  $p < 0.10$

Table 9: Simulated Turnout with Reprecincting and 0-Distance Scenarios

	Actual Dist.	$\Delta$ Dist. Actual-Simul.	2012 Turnout (%)			2013 Turnout (%)			2014 Turnout (%)			2016 Turnout (%)		
			Actual	Simulated		Actual	Simulated		Actual	Simulated		Actual	Simulated	
				Reprec. 0-Dist.	0-Dist.		Reprec. 0-Dist.	0-Dist.		Reprec. 0-Dist.	0-Dist.			
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	
<u>A. Average Census Block</u>														
All parcels	.360	.035	61.0	61.4	65.0	29.9	30.1	32.2	45.1	45.3	47.6	38.5	38.7	40.1
<u>B. By % Minority</u>														
% minority $\leq$ median	.382	.043	69.9	70.5	74.7	37.5	37.7	38.9	54.8	55.0	56.5	45.5	45.7	46.5
% minority $>$ median	.342	.029	53.9	54.1	57.2	24.2	24.4	27.3	37.2	37.5	40.8	33.2	33.4	35.7
Turnout gap (High-Low SES)			16.1	16.3	17.5	13.3	13.2	11.6	17.6	17.5	15.6	12.3	12.2	10.8
<u>C. By Median HH Income</u>														
Income $\leq$ median	.342	.032	50.1	50.5	54.2	22.2	22.4	24.8	33.8	34.1	37.3	27.8	28.0	30.1
Income $>$ median	.376	.039	71.5	71.9	75.5	41.1	41.2	42.6	55.5	55.7	57.2	46.8	46.9	47.9
Turnout gap (High-Low SES)			21.3	21.4	21.3	18.9	18.8	17.8	21.7	21.6	19.8	19.0	18.9	17.8
<u>D. By % Housing Units w/o Cars</u>														
% w/o cars $\leq$ median	.417	.043	69.3	69.7	73.7	35.2	35.3	37.0	54.3	54.5	56.4	44.7	44.8	45.6
% w/o cars $>$ median	.309	.029	53.8	54.2	57.6	25.7	25.9	28.3	37.0	37.3	40.1	34.3	34.5	36.8
Turnout gap (High-Low SES)			15.4	15.5	16.1	9.5	9.4	8.7	17.3	17.3	16.3	10.4	10.3	8.8

Notes: This table reports actual and estimated counterfactual values of census block distance to the polling place and turnout in the full census block sample (Parcel A), by minority presence (Panel B), by median household income (Panel C), and by the proportion of units without cars (Panel D). Column 1 reports averages of actual polling place distances. Column 2 reports average differences between actual and counterfactual distances to the polling place, where the latter comes from the efficient reprecincting algorithm described in the text. Simulated "Reprecincting" turnout is the expected census block turnout under efficient reprecincting. Simulated "0-Distance" turnout is the expected turnout assuming distance to the polling place is erased for every census block.

Figure A1: Sensitivity of Parcel Estimates to Distance to Boundary/Match

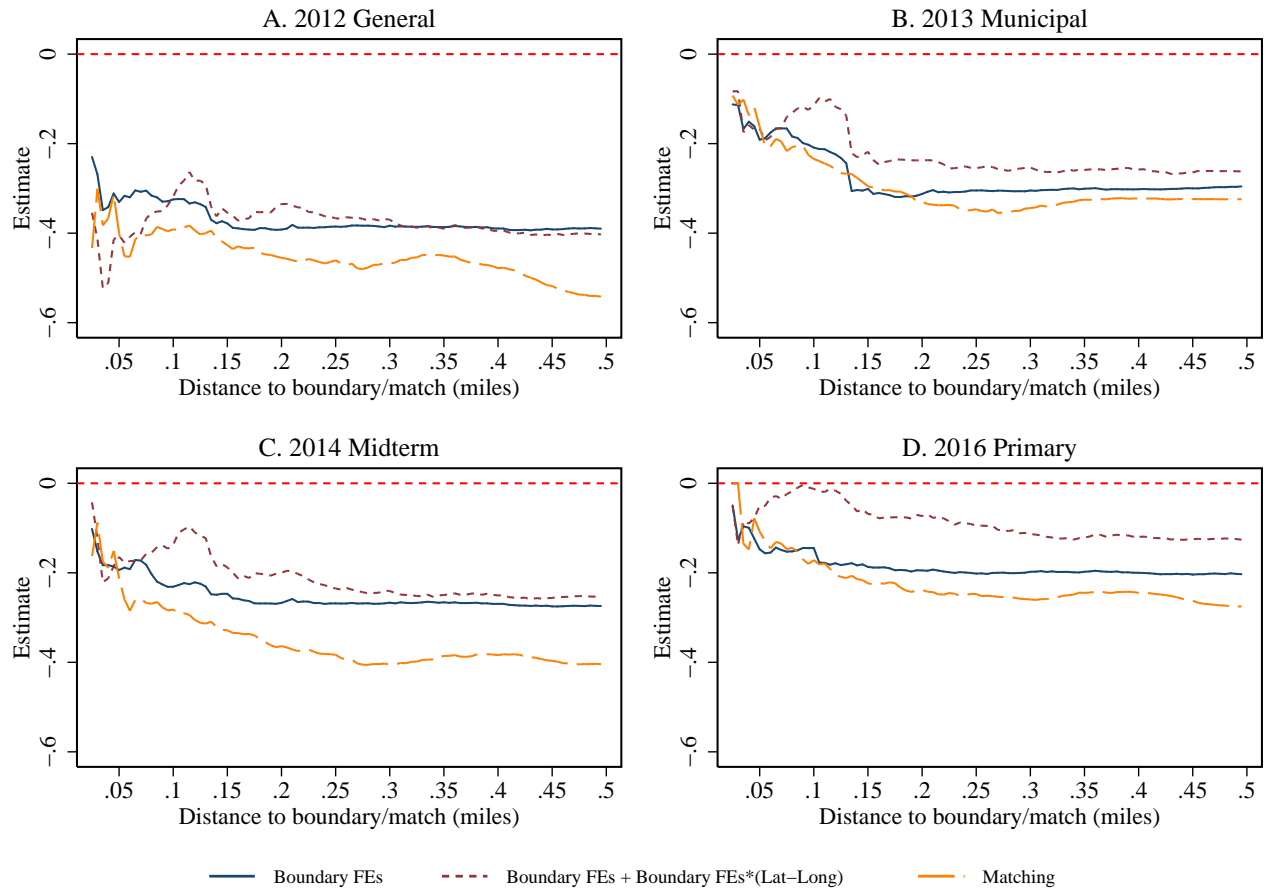




Figure A2: Within-Boundary Parcel-Level Estimates Across Distances to Boundary

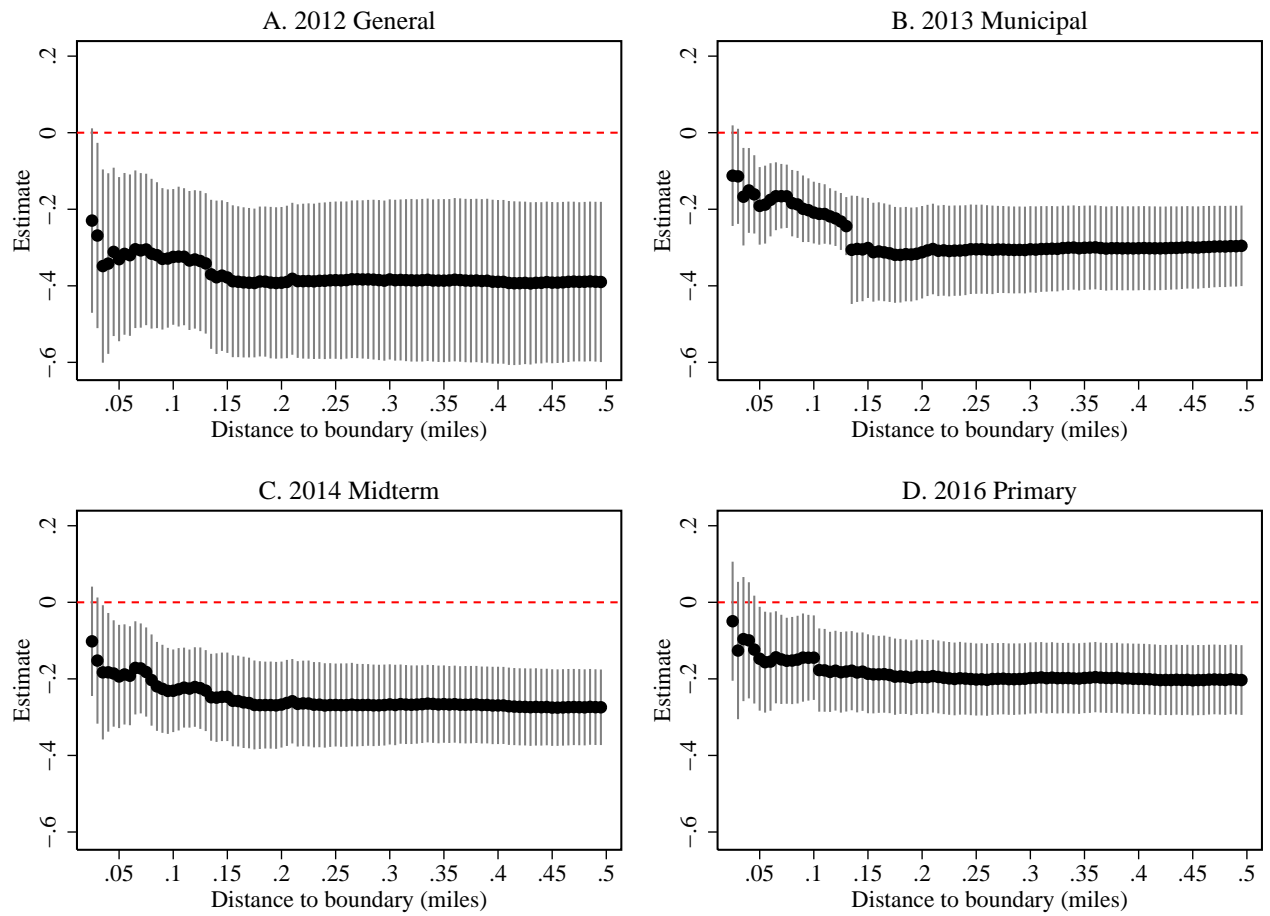


Figure A3: Sensitivity of Census Block Estimates to Distance to Boundary/Match

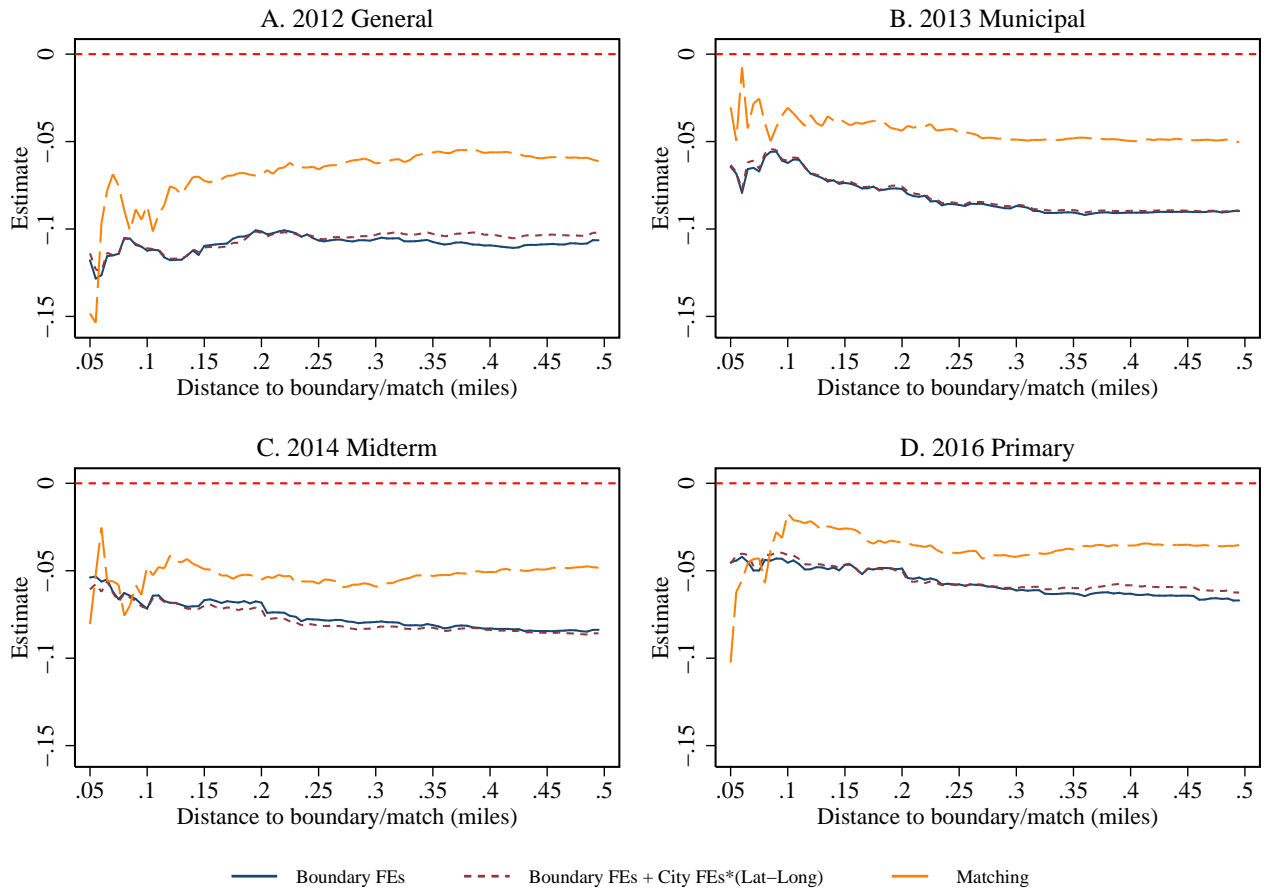


Figure A4: Within-Boundary Block-Level Estimates Across Distances to Boundary

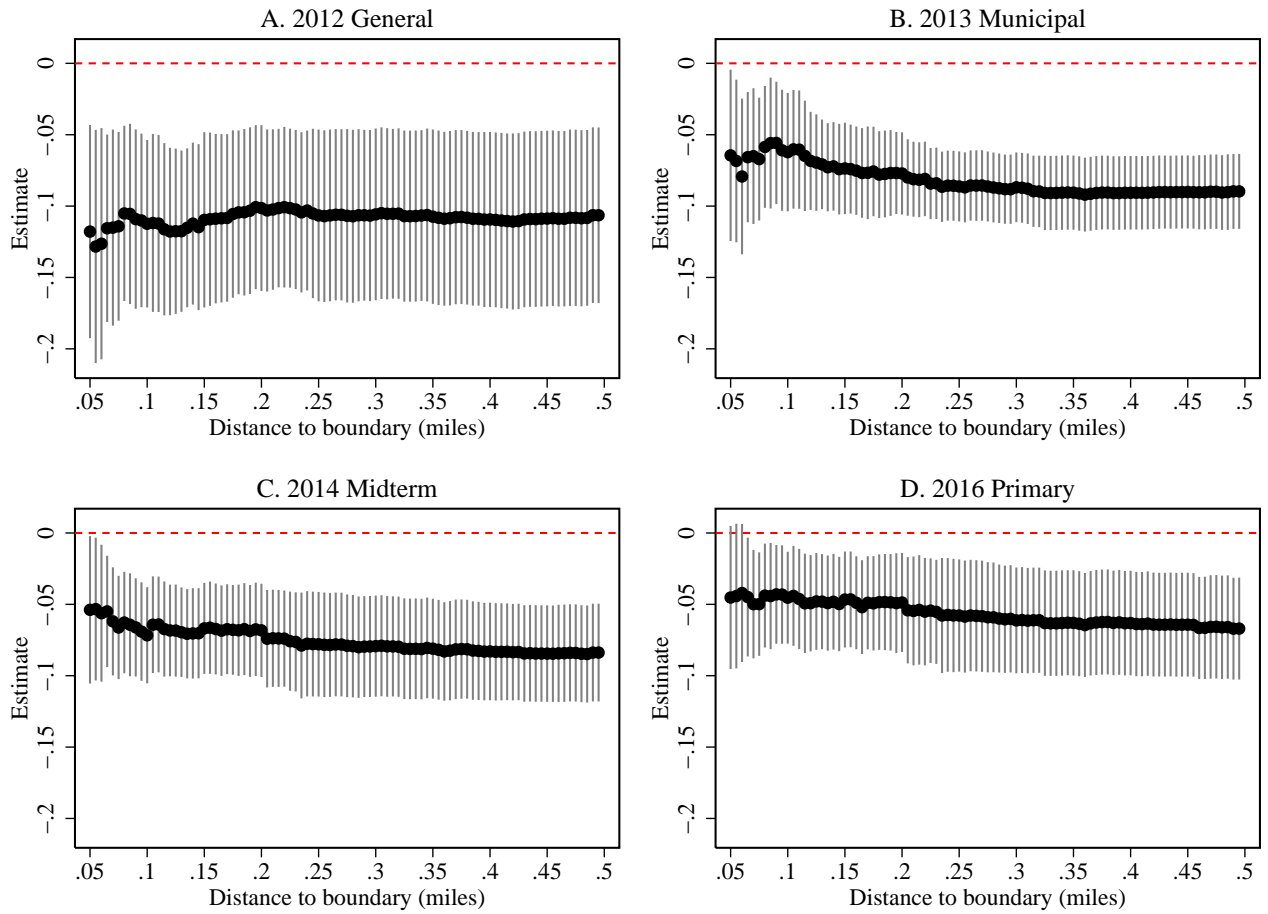


Table A1:  $R^2$  of Distance to the Polling Place on Geographic Controls

	<u>Discontinuity Sample</u>			<u>Placebo Sample</u>		
Dist. to Bound.:	<.15 mi	<.10 mi	<.05 mi	<.15 mi	<.10 mi	<.05 mi
	(1)	(2)	(3)	(4)	(5)	(6)
	<u>A. Boundary FEs</u>					
$R^2$	0.53	0.54	0.56	0.68	0.70	0.73
	<u>B. Boundary FEs + City FEs×(Lat-Long)</u>					
$R^2$	0.54	0.55	0.57	0.74	0.76	0.78
	<u>C. Boundary FEs + City FEs×(Lat-Long)<sup>2</sup></u>					
$R^2$	0.56	0.56	0.58	0.79	0.80	0.81
	<u>D. Boundary FEs + City FEs×(Lat-Long)<sup>3</sup></u>					
$R^2$	0.57	0.57	0.59	0.79	0.80	0.82
	<u>E. Boundary FEs + Boundary FEs×(Lat-Long)</u>					
$R^2$	0.79	0.79	0.79	0.95	0.96	0.97
N	71,581	59,990	36,061	39,914	33,443	20,633

Notes: This table reports the R-squared from parcel-level regressions of distance to the polling place on boundary fixed effects (Panel A), boundary fixed effects and municipality-specific polynomials in latitude-longitude (Panels B, C, and D), and boundary fixed effects interacted with latitude-longitude (Panel E). Columns 4 through 6 are based on precinct boundaries that do not induce discontinuities in assignment to polling places; that is, parcels on either side of each boundary are assigned to vote at the same polling location.

Table A2: Bias in Voter-Level Turnout Regressions

	2014 Turnout		2014 Votes Cast
	(1)	(2)	(3)
Distance to polling place	-.003 (.014)	-.038 * (.016)	-.215 ** (.061)
Sample	Full sample	<.10 mile	<.10 mile
Boundary FEs	NO	YES	YES
Mean dep. var.	.55	.54	1.43
N	238,505	54,858	59,990

Notes: Columns 1 and 2 report estimates from regressions of parcel-level 2014 turnout (i.e., votes cast by parcel residents over the number of registered voters) on distance to the polling place. The outcome of column 3 is the number of ballots cast by parcel residents in 2014.

\*\* p < 0.01, \* p < 0.05, ~ p < 0.10

Table A3: Heterogeneous Effects by Census Characteristics  
Interacted Specifications

Election:	2012 Presidential		2013 Municipal		2014 Midterm		2016 Primary	
	Mean (1)	Effect (2)	Mean (3)	Effect (4)	Mean (5)	Effect (6)	Mean (7)	Effect (8)
<u>A. By % Minority</u>								
% minority $\leq$ median	1.88	-.145 (.106)	1.04	-.048 (.077)	1.42	.005 (.048)	1.33	.040 (.051)
% minority $>$ median	2.16	-.117 ~ (.066)	.97	-.155 (.096)	1.43	-.151 * (.062)	1.45	.008 (.059)
F-test (within year)		.11		1.14		6.0		.21
p		.74		.28		.01		.65
F-test (across years)		1.75						
p		.14						
N	59,990	59,990	45,705	45,705	59,990	59,990	42,939	42,939
<u>B. By Median HH Income</u>								
Income $\leq$ median	1.98	-.101 (.063)	.88	-.127 (.086)	1.28	-.106 (.066)	1.20	-.003 (.062)
Income $>$ median	2.08	-.156 (.104)	1.17	-.063 (.072)	1.56	-.040 (.049)	1.54	.040 (.044)
F-test (within year)		.45		.81		.98		.57
p		.50		.37		.32		.45
F-test (across years)		.55						
p		.70						
N	59,990	59,990	45,705	45,705	59,990	59,990	42,939	42,939
<u>C. By % Units w/o Cars</u>								
% w/o cars $\leq$ median	1.67	-.160 (.106)	.92	-.076 (.074)	1.29	-.031 (.049)	1.20	.067 (.043)
% w/o cars $>$ median	2.32	-.096 (.064)	1.08	-.125 (.091)	1.53	-.117 ~ (.066)	1.51	-.033 (.061)
F-test (within year)		.53		.34		1.61		2.72
p		.47		.56		.21		.10
F-test (across years)		1.42						
p		.23						
N	59,990	59,990	45,705	45,705	59,990	59,990	42,939	42,939

Notes: This table replicates estimates of heterogeneous effects from Table 6 using interacted Poisson specifications, that is, regressions that control for boundary-specific linear interactions with latitude and longitude.

\*\*  $p < 0.01$ , \*  $p < 0.05$ , ~  $p < 0.10$

Table A4: Heterogeneous Effects by Census Characteristics  
Matching Specifications

Election:	2012 Presidential		2013 Municipal		2014 Midterm		2016 Primary	
	Mean (1)	Effect (2)	Mean (3)	Effect (4)	Mean (5)	Effect (6)	Mean (7)	Effect (8)
<u>A. By % Minority</u>								
% minority $\leq$ median	2.06	-.218 * (.099)	1.07	-.264 * (.108)	1.53	-.058 (.069)	1.48	-.054 (.076)
% minority $>$ median	2.44	-.153 ** (.057)	1.02	-.237 * (.095)	1.57	-.261 ** (.067)	1.61	-.142 ~ (.075)
F-test (within year)		.38		.03		5.70		.84
p		.54		.85		.02		.36
N	133,644	133,644	96,084	96,084	133,644	133,644	99,082	99,082
<u>B. By Median HH Income</u>								
Income $\leq$ median	2.21	-.155 ** (.059)	.94	-.346 ** (.098)	1.41	-.285 ** (.077)	1.35	-.166 ~ (.094)
Income $>$ median	2.33	-.200 * (.084)	1.18	-.141 ~ (.082)	1.67	-.091 (.068)	1.70	-.065 (.067)
F-test (within year)		.23		2.80		3.90		.91
p		.63		.09		.05		.34
N	133,644	133,644	96,084	96,084	133,644	133,644	99,082	99,082
<u>C. By % Units w/o Cars</u>								
% w/o cars $\leq$ median	1.75	-.120 (.093)	.92	-.156 (.095)	1.32	-.052 (.068)	1.24	-.014 (.070)
% w/o cars $>$ median	2.66	-.234 ** (.064)	1.13	-.329 ** (.098)	1.72	-.294 ** (.077)	1.74	-.193 * (.089)
F-test (within year)		1.08		1.72		5.81		2.72
p		.30		.19		.02		.10
N	133,644	133,644	96,084	96,084	133,644	133,644	99,082	99,082

Notes: This table replicates estimates of heterogeneous effects from Table 6 using matching specifications.

\*\*  $p < 0.01$ , \*  $p < 0.05$ , ~  $p < 0.10$

Table A5: Heterogeneous Turnout Effects by Census Characteristics  
Boundary FE Specifications

Election:	2012 Presidential		2013 Municipal		2014 Midterm		2016 Primary	
	Mean	Effect	Mean	Effect	Mean	Effect	Mean	Effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<u>A. By % Minority</u>								
% minority $\leq$ median	.63	-.125 * (.053)	.36	-.037 (.029)	.49	-.042 ~ (.023)	.39	-.025 (.020)
% minority $>$ median	.51	-.097 ** (.024)	.24	-.091 ** (.019)	.34	-.106 ** (.019)	.31	-.073 ** (.022)
F-test (within year)		.24		2.83		4.78		2.99
p		.62		.09		.03		.08
F-test (across years)		2.92						
p		.02						
N	3,339	3,339	2,552	2,552	3,339	3,339	2,376	2,376
<u>B. By Median HH Income</u>								
Income $\leq$ median	.49	-.118 ** (.019)	.22	-.078 ** (.017)	.32	-.102 ** (.015)	.27	-.066 ** (.019)
Income $>$ median	.65	-.107 * (.049)	.41	-.042 (.028)	.50	-.043 ~ (.023)	.42	-.030 (.021)
F-test (within year)		.05		2.05		6.34		2.18
p		.82		.15		.01		.14
F-test (across years)		1.69						
p		.15						
N	3,339	3,339	2,552	2,552	3,339	3,339	2,376	2,376
<u>C. By % Units w/o Cars</u>								
% w/o cars $\leq$ median	.64	-.105 * (.046)	.36	-.044 (.027)	.50	-.050 * (.022)	.41	-.021 (.020)
% w/o cars $>$ median	.52	-.122 ** (.022)	.26	-.086 ** (.019)	.35	-.100 ** (.018)	.32	-.081 ** (.020)
F-test (within year)		.13		2.83		4.00		6.13
p		.72		.09		.05		.01
F-test (across years)		1.97						
p		.10						
N	3,339	3,339	2,552	2,552	3,339	3,339	2,376	2,376

Notes: This table reports estimates from boundary fixed effects OLS regressions that interact distance to the polling place with dummies for lower- and higher-than-median values of census block minority presence (Panel A), census block group median income (Panel B), and block group percentage of residential units without cars (Panel C). The null hypothesis of within-year F-tests is that the effect of distance to the polling place is the same across census blocks with higher-than-median and lower-than-median values of the interacting characteristic. The null hypothesis of across-years F-tests is that the effects are identical in every election.



Table A6: Heterogeneous Turnout Effects by Census Characteristics  
Interacted Specifications

Election:	2012 Presidential		2013 Municipal		2014 Midterm		2016 Primary	
	Mean (1)	Effect (2)	Mean (3)	Effect (4)	Mean (5)	Effect (6)	Mean (7)	Effect (8)
<u>A. By % Minority</u>								
% minority $\leq$ median	.63	-.125 * (.055)	.36	-.036 (.029)	.49	-.045 * (.021)	.39	-.021 (.018)
% minority $>$ median	.51	-.094 ** (.023)	.24	-.089 ** (.019)	.34	-.103 ** (.019)	.31	-.068 ** (.021)
F-test (within year)		.26		2.66		4.26		3.11
p		.61		.10		.04		.08
F-test (across years)		2.93						
p		.02						
N	3,339	3,339	2,552	2,552	3,339	3,339	2,376	2,376
<u>B. By Median HH Income</u>								
Income $\leq$ median	.49	-.119 ** (.020)	.22	-.076 ** (.017)	.32	-.100 ** (.015)	.27	-.062 ** (.019)
Income $>$ median	.65	-.103 * (.051)	.41	-.042 (.029)	.50	-.044 * (.021)	.42	-.024 (.019)
F-test (within year)		.11		1.79		6.45		2.89
p		.74		.18		.01		.09
F-test (across years)		1.69						
p		.15						
N	3,339	3,339	2,552	2,552	3,339	3,339	2,376	2,376
<u>C. By % Units w/o Cars</u>								
% w/o cars $\leq$ median	.64	-.105 * (.049)	.36	-.044 (.027)	.50	-.053 ** (.021)	.41	-.019 (.018)
% w/o cars $>$ median	.52	-.118 ** (.023)	.26	-.083 ** (.018)	.35	-.095 ** (.018)	.32	-.074 ** (.020)
F-test (within year)		.07		2.11		2.56		4.96
p		.79		.15		.11		.03
F-test (across years)		1.63						
p		.16						
N	3,339	3,339	2,552	2,552	3,339	3,339	2,376	2,376

Notes: This table replicates estimates of heterogeneous effects from Table A4 using interacted specifications, that is, regressions that control for municipality-specific linear interactions with latitude and longitude.

\*\*  $p < 0.01$ , \*  $p < 0.05$ , ~  $p < 0.10$

Table A7: Heterogeneous Turnout Effects by Census Characteristics  
Matching Specifications

Election:	2012 Presidential		2013 Municipal		2014 Midterm		2016 Primary	
	Mean	Effect	Mean	Effect	Mean	Effect	Mean	Effect
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
<u>A. By % Minority</u>								
% minority $\leq$ median	.59	-.114 (.075)	.33	.023 (.035)	.44	.003 (.035)	.38	.018 (.030)
% minority $>$ median	.49	-.057 (.041)	.23	-.087 ** (.028)	.31	-.099 ** (.037)	.29	-.052 ~ (.032)
F-test (within year)		.41		6.66		4.10		2.41
p		.52		.01		.04		.12
N	4,130	4,130	2,938	2,938	4,130	4,130	3,334	3,334
<u>B. By Median HH Income</u>								
Income $\leq$ median	.47	-.032 (.035)	.21	-.047 ~ (.024)	.29	-.056 * (.028)	.25	-.013 (.025)
Income $>$ median	.60	-.135 ~ (.069)	.39	-.007 (.051)	.45	-.040 (.046)	.40	-.020 (.034)
F-test (within year)		2.00		.52		.09		.03
p		.16		.47		.76		.86
N	4,130	4,130	2,938	2,938	4,130	4,130	3,334	3,334
<u>C. By % Units w/o Cars</u>								
% w/o cars $\leq$ median	.60	-.087 (.067)	.33	.009 (.035)	.46	-.000 (.040)	.41	.018 (.030)
% w/o cars $>$ median	.50	-.086 * (.035)	.25	-.077 ** (.029)	.33	-.101 ** (.028)	.30	-.059 * (.026)
F-test (within year)		.00		4.10		4.50		3.90
p		.99		.04		.03		.05
N	4,130	4,130	2,938	2,938	4,130	4,130	3,334	3,334

Notes: This table replicates estimates of heterogeneous effects from Table A4 using matching specifications.

\*\*  $p < 0.01$ , \*  $p < 0.05$ , ~  $p < 0.10$