American Economic Journal: Applied Economics 2020, 12(1): 61–85
https://doi.org/10.1257/app.20180306

A Precinct Too Far: Turnout and Voting Costs†

By Enrico Cantoni*

I study the effects of voting costs—specifically, distance to polling location—using geographic discontinuities. Opposite sides of boundaries between voting precincts are observationally identical, except for their assigned polling locations. This discontinuous assignment produces sharp changes in voters' travel distance to cast their ballots. In nine municipalities in Massachusetts and Minnesota, a 1 standard deviation (0.245 mile) increase in distance reduces ballots cast by 2 to 5 percent across four elections. During non-presidential elections, effects are three times larger in high-minority areas than in low-minority areas. Finally, I simulate the impact of various counterfactual assignments of voters to polling places. (*JEL* D72, J15, R41)

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—the one 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 rests 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

*Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126, Bologna, Italy (email: enrico.cantoni@unibo.it). Seema Jayachandran was coeditor for this article. I am very thankful to the George and Obie Shultz Fund at MIT Economics for its generous funding. I am deeply indebted to Ben Olken, Daron Acemoglu, Charles Stewart III, and Josh Angrist for their constant and patient supervision. I thank Eda Matchak, Sabino Piemonte, and election administrators in Massachusetts and Minnesota for their precious help collecting and understanding the voter data. I gratefully acknowledge the invaluable feedback of Isaiah Andrews, Peter Aronow, Alex Bartik, Giulia Brancaccio, David Colino, Tommaso Denti, Selene Ghisolfi, Donghee Jo, Daniel Pollman, Vincent Pons, Brendan Price, Cory Smith, Irena Stojkov, Marco Tabellini, and seminar participants at MIT, Yale, the University of Bologna, Analysis Group, Edgeworth Economics, and SAEe 2016. I have no relevant or material financial interests that relate to the research described in this paper. In addition, neither I nor members of my family hold any positions as officer, director, or board member of a relevant organization.
†Go to https://doi.org/10.1257/app.20180306 to visit the article page for additional materials and author disclosure statement(s) or to comment in the online discussion forum.

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, Kelly, and Witko 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, whether they equally impact voters across socioeconomic 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 crossing 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—especially in states that feature Election Day voter registration—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 me to focus on units located in extreme proximity (i.e., as little as 0.05 miles, or 80.5 meters) to precinct borders while preserving good statistical power. I then test robustness of the main results to an alternative specification that augments boundary fixed effects with local polynomials in latitude and longitude (as in, e.g., Dell 2010, Dell and Querubin 2018). The last specification builds on recent work by Keele, Titiunik, and Zubizarreta (2015). 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,[1] 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

[1] The Massachusetts municipalities are the cities of Boston, Cambridge, Fall River, Lowell, Newton, Quincy, Somerville, and the town of Brookline. The Minnesota municipality is Minneapolis.

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, voters in Minnesota experienced this requirement in 2012 and 2013 but saw it drop 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 (0.245 mile) increase in distance to the polling place reduces the number of ballots cast by parcel residents by 2–5 percent, which translates to a 1–3 percentage point decrease in turnout. Turnout effects are paralleled by sizable but imprecisely estimated impacts on voter registration, which underscore the importance of accounting for possible sample selection bias.

Since the Supreme Court *Shelby County v. Holder* ruling of 2013,[2] there is growing concern that changes to election practices may be used to disenfranchise Democratic-leaning poor and minority voters by disproportionately increasing their voting costs (e.g., through voter ID laws or restrictions in early voting; see Biggers and Hanmer 2015, Hicks et al. 2015, and Minnite 2013). Unlike voter ID laws and convenience voting restrictions, the assignment of voters to polling locations in my sample is the result of a transparent, nonpartisan process. Yet, I find that the effects of distance to the polling place carry potential partisan consequences, thus highlighting the importance of monitoring both state electoral practices—like voter ID laws—and local ones—like polling place assignment. Specifically, heterogeneous effects 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. 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 contrast emerges from interactions with census block group income and car availability. However, minority presence, income, and car availability are quite correlated,[3] so specifications that simultaneously control for all these interactions reveal the difficulty of assessing the exact contribution of each interaction variable.

The large impact on voter participation suggests that distance to the polling place might also affect voter behavior in non-electoral activities. For example, due to lack of interest in making an educated decision on Election Day, voters who are deterred from voting by distance to the polling place might decide to acquire less political information. Similarly, if voter participation and political contributions are complementary activities, distance to the polling place might reduce the

[2] In the case, the Court ruled Section 4(b) of the 1965 Voting Rights Act (VRA) unconstitutional, thereby rendering Section 5 inapplicable too. Jurisdictions covered by Section 5 of the VRA could change election practices only after the Attorney General determined that the proposed changes had no discriminatory purpose or effect.

[3] The block-level correlation between minority presence and income is −0.414, while the correlation between minority presence and the share of households without cars is 0.379. The correlation between income and the share of nonmotorized households is −0.473.

number of contributions to political organizations and candidates. I explore these possibilities using individual-level data on magazine/newspaper subscriptions and 2010–2012 contributions to political organizations and candidates. I find no effect of distance to the polling place on either parcel-level counts of magazine/newspaper subscribers or on the number of individuals who contributed to political organizations or candidates. This finding appears to cast doubt on theories of endogenous information, which predict that increasing costs to participate will also deter citizens from investing in the acquisition of political information. Caution is needed in interpreting these results, however, as several reasons—all potentially consistent with the complementarity of voter participation and information acquisition/political donations—could justify the zero effects.

Finally, I combine optimization tools from location science with my econometric estimates to simulate the turnout effects of counterfactual assignments of voters to polling places. Efficiently re-drawing precinct lines can increase voter turnout by 0.2–0.4 percentage points (i.e., between 0.5 and 0.7 percent), while a hypothetical policy that erased the impact of distance to the polling place would increase turnout by 1.6–4 percentage points (i.e., between 4.2 and 7.3 percent) and reduce participation gaps across low- and high-minority areas in non-presidential elections by 11.2–12.8 percent.

The next section discusses the institutional setting. I detail my empirical strategy in Section II and discuss summary statistics and tests of the identification assumption in Section III. Section IV presents main effects at the parcel- and census block-levels. Section V explores heterogeneous effects. Effects on magazine/newspaper subscriptions and campaign contributions are presented in Section VI. Section VII describes the efficient re-precincting algorithm. Section VIII concludes.

I. Institutional Background

A. *Precincts, Census Blocks, and Polling Places*

American voters are assigned to Election Day polling locations based on the precinct they live in.[4] Precincts represent the basic geographic unit for administering elections: they partition municipalities and constitute the building blocks of every geographic aggregation used for election purposes, including congressional and state legislative districts.

After each decennial federal census, precinct boundaries in Massachusetts and Minnesota are revised through a joint effort of municipal officials from the election and assessing departments.[5] These nonpartisan local officials have no direct, electoral incentive to manipulate precinct maps.[6] This contrasts with state legislative

[4] This is not the case in Colorado, Oregon, and Washington, which feature all-mail voting. Moreover, most states offer the possibility to vote before Election Day, through early in-person and/or absentee voting. Source: http://www.ncsl.org/research/elections-and-campaigns/absentee-and-early-voting.aspx (accessed January 9, 2019).

[5] The requirement to re-draw precinct lines after the decennial census does not apply to the city of Boston (1982 Mass. Acts ch. 605, section 3).

[6] Although vacancies for these positions are filled by mayors' or city managers' appointees, the tenure of appointed officials typically continues after the mayor or city manager leaves office.

and congressional districts boundaries, which, in both states, are drawn by the state legislature. Re-precincting is only one of the many tasks fulfilled by these officials (e.g., election offices administer the yearly municipal census, run elections; assessing offices maintain assessors' information). Also for this reason, anecdotally, these officials' objective is to minimize changes in precinct boundaries. In fact, making small revisions to existing boundaries—to account for changes in precinct population that occurred since the previous census[7]—is logistically simpler than re-drawing precincts from scratch. Moreover, changes in precinct borders automatically trigger the requirement to notify affected voters, which entails logistical and financial costs.

Prior to implementation, revised precinct maps are subject to multiple levels of approval. This approval process is intended as a check on the fairness of any revisions.[8] After re-precincting, precinct boundaries are then left unchanged until the next decennial census. Both states recommend (Minnesota), or outright prescribe (Massachusetts), that precincts be bounded by census block boundaries.

Census blocks are the smallest geographic unit used by the US Census Bureau for tabulation of decennial census data from all houses. In urban areas, census blocks typically coincide with city blocks (Chapter 11 of US Census Bureau 1994). 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 socioeconomic status (SES), including income and educational attainment, are nonetheless available by block group.[9]

Polling places can change over time for two reasons. First, a census block can be assigned to a new precinct as a result of post-censal re-precincting. Second, 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 a polling site).

B. *Voter List Maintenance and Election Day Registration*

Distance to the polling place may affect voter registration through two channels. First, both states in my sample purge voters because of electoral inactivity.[10] Between

[7] Precincts in Massachusetts municipalities must contain an equal share of the municipal population, ±5 percent, not to exceed 4,000 inhabitants. Similarly, Minnesota re-precincting guidelines suggest that precincts contain fewer than 2,000 registered voters (Minnesota Secretary of State 2011).

[8] First, they must be approved by the city council. Revised maps in Massachusetts must then be approved by an independent state commission (the Local Election Districts Review Commission, "LEDRC"), which verifies the maps' compliance with state and federal regulations. In Minnesota, a similar supervisory role is played directly by the Secretary of State. Boards of elections commissioners in Massachusetts municipalities provide an additional layer of supervision (MGL ch.51 §16A). Each board consists of two Democratic and two Republican members. While the board's main duty is to oversee the electoral process on Election Day, the board must also approve any change in polling locations, either caused by re-drawn precinct lines or by the reassignment of entire precincts to different polling stations.

[9] Census blocks are grouped into block groups, which are the smallest geographic aggregation at which the US Census Bureau releases income and schooling data. Block groups typically contain between 600 and 3,000 people.

[10] Technically, inactive voters in Massachusetts can be purged only after they fail to respond to a mail notification of removal.

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 percent of the total removals in the two states (US Election Assistance Commission 2015).

Second, distance to the polling place may dissuade eligible but unregistered voters from registering. 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. Ballots can be observed and aggregated at the parcel level independently of the residents' registration status. Thus, ballot counts and parcels constitute, respectively, an outcome and a sample that are robust to selection concerns. By contrast, voter-level turnout is only observed for registered voters, so analyses at this level would miss any effect of distance on voter registration.

The 2012–2016 electoral calendars in the two states are described in online Appendix A.

C. *Existing Evidence on Distance to the Polling Place and Turnout*

My finding that voting costs induced by "mechanisms of voting" disproportionately affect low-propensity voters is consistent with at least four studies from different geographic contexts. Gerber, Huber, and Hill (2013) documents that the introduction of all-mail elections in Oregon increased participation by 2–4 percentage points, an effect that was larger among less active registrants than for frequent voters. Hodler, Luechinger, and Stutzer (2015) shows that postal voting in Switzerland increased voter turnout, especially among low-educated voters. Fujiwara (2015) reports that electronic voting in Brazil fostered the enfranchisement of less educated citizens. Using an empirical design similar to mine, Kaplan and Yuan (2020) estimates the effect of early voting on turnout. The authors compare voter participation across county borders, exploiting a 2010 Ohio law that forced some counties to expand and others to contract early voting. They find that early voting disproportionately benefits women, Democrats, and working-age voters.

Earlier studies on the disenfranchising effect of distance to the polling place in the United States (Dyck and Gimpel 2005; Gimpel, Dyck, and Shaw 2004; Haspel and Knotts 2005) are mostly mute about endogeneity concerns, so resulting estimates could be partly driven by (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. Causal identification in their setting hinges on the comparability of (matched) voters in consolidated and unconsolidated precincts. My geographic designs replace this assumption with an arguably weaker one: that voters living in close proximity to each other, but on opposite sides of the same precinct border, are on average identical, except for the distance they need to travel to cast their ballots. Moreover, my empirical strategy allows for the possibility that distance to the polling place affects voter registration (e.g., dissuading voters who would register on Election Day in Minnesota).

II. 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. Moreover, particularly in states that feature Election Day registration (like Minnesota), some voters might refrain from registering altogether if they anticipate their polling place being too far away. This section discusses how I make progress on these two issues.

A. *Level of Analysis*

To address the potential endogeneity of voter registration, the main analysis is conducted at the parcel 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."[11]

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.

[11] Interestingly, the issue of endogenous voter registration in micro-level studies has, until recently, received limited scholarly attention; see Erikson (1981) and Timpone (1998) for two early exceptions, while a recent paper by Nyhan, Skovron, and Titiunik (2017) 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. Braconnier, Dormagen, and Pons (2017) examines the effects of voter registration drives during the 2012 French presidential election by randomly assigning entire buildings to canvassing treatments or to a control group, and then measuring the impact on registration counts at the address level.

B. *Boundary Fixed Effects*

Discontinuous changes in assignment to polling locations across adjacent precincts provide plausibly exogenous treatment variation. The idea is to compare parcels that are close to each other, but on opposite sides of a precinct border (coinciding, for example, with a street). Units on the two sides of the border share identical characteristics, on average. However, because of their assignment to different polling locations, residents of the two sides must travel different distances to cast their ballots, thus motivating a boundary discontinuity design à la Black (1999):

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

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.

Causal identification requires that, except for distance to the polling place, voters who live close to but on opposite sides of a boundary share similar determinants of voter turnout (e.g., identical socioeconomic background). By Tobler's (1970) first law of geography—"Everything is related to everything else, but near things are more related than distant things."—this is more likely to hold for units located right on opposite sides of a precinct border. Hence, all analyses are based on samples drawn from two narrow bands around precinct boundaries: 0.10 and 0.05 miles (160.9 and 80.5 meters, respectively).

Equation (1) also requires that distance to the polling place is the only determinant of voter participation that changes discontinuously across precinct boundaries. Although the mere overlap with other boundaries is not problematic,[12] I conservatively exclude precinct boundaries overlapping other institutional or geographic discontinuities. Online Appendices B and C detail, respectively, the data sources and the sample restrictions. Unless specified otherwise, standard errors are clustered by boundary throughout.

Figure 1 illustrates the identification strategy using two precinct boundaries from Cambridge, Massachusetts. The small polygons and the thick black lines represent parcels and precinct borders, respectively. Colored parcels are closer than 0.10 miles 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. Uncolored parcels are excluded from the sample for one or more of the following reasons: they are farther than 0.10 miles

[12] To see this point, consider a precinct border coinciding with the border between school assignment zones. In this case, 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.

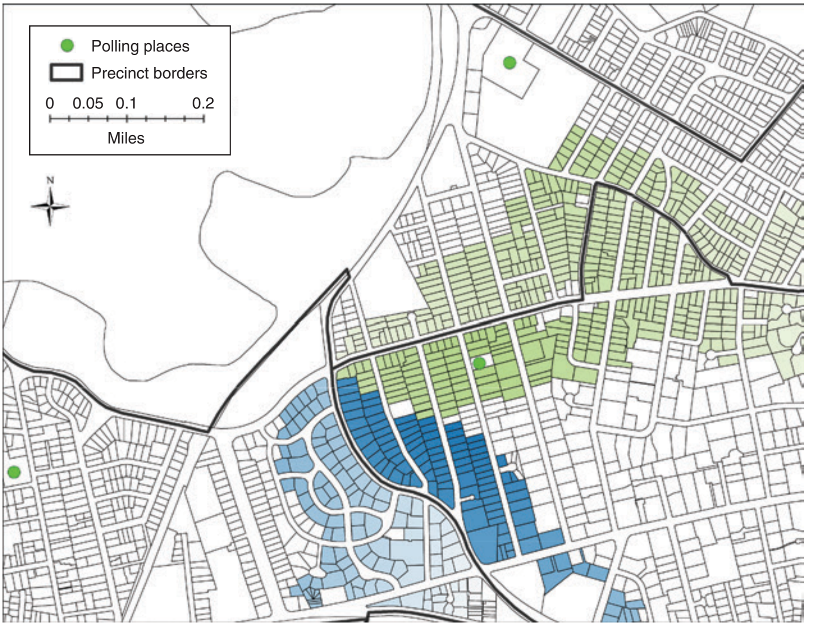


Figure 1. Example of Boundary Fixed Effects

Notes: This map illustrates the within-boundary identification strategy using two precinct boundaries from Cambridge, Massachusetts. The small polygons and the thick black lines represent parcels and precinct borders, respectively. Colored parcels are closer than 0.10 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, the blue boundary reveals a discontinuous change in distance to the polling place across the corresponding precinct border: parcels on the east side are closer to their polling location than parcels on the west side, so they display darker shades of blue.

to the nearest precinct boundary, their precinct boundary overlaps other discontinuities,[13] and/or they are nonresidential lots.

A visual analysis of the blue boundary in Figure 1 reveals a discontinuous change in distance to the polling place across the corresponding precinct border. The polling place of parcels on the west (east) side of the boundary is denoted by the green dot at the left (center) of the figure. Parcels on the east side are closer to their polling location than parcels on the west side, so they display darker shades of blue.

The green boundary reveals a second type of continuous, within-boundary-side variation. Specifically, moving eastward along the south side of the green boundary increases distance to the polling place relative to other southwest-side parcels, as reflected by lighter shades of green. The presence of this second, within-boundary-side variation contrasts with existing geographic discontinuity designs (e.g., Black 1999; Dell 2010; Dell and Querubin 2018; Keele, Titiunik, and Zubizarreta 2015; Keele and Titiunik 2015, 2016; Lavy 2010), which only feature sharp changes in treatment assignment across boundary sides. Within-boundary-side variation would be problematic if systematically correlated with other determinants of voter participation.

[13] For example, parcels in the west part of the figure are excluded because their precinct boundary coincides with the Fresh Pond water reservoir.

In the green boundary example, this would be the case if moving eastward along the south side increased *both* distance to the polling place *and* socioeconomic status (which itself also affects voter participation).

In this context, within-boundary-side variation is unlikely to threaten identification for at least three reasons. First, precinct boundaries encompass extremely small geographic areas, leaving arguably limited scope for changes in determinants of voter turnout other than distance to the polling place. Second, any such changes would be problematic only if systematically correlated with distance to the polling place; for example, *within each side of each boundary*, parcels that are relatively farther to polling locations should be systematically wealthier than closer ones. Third, within-boundary correlations between determinants of voter participation and distance to the polling place would tend to show up in balancing exercises, of which I find no evidence. The next section and online Appendix D describe alternative, less parsimonious specifications that rely exclusively on the discontinuous variation across boundary sides.

C. *Matching*

Following Keele, Titiunik, and Zubizarreta (2015), I use distance-based nearest-neighbor matching (with replacement) as an alternative method to identify the impact of distance to the polling place. The idea is to compare *pairs* of neighboring parcels or blocks that span across precinct borders, instead of comparing all units on one side of the border with all units on the opposite side, as implicitly done by within-boundary specifications. By proximity, units within pairs should share identical characteristics, on average. However, because they live in different precincts, their residents are assigned to vote at different polling locations, thereby leaving plausibly exogenous variation in 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 online Appendix C. 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:

$$y_{ip} = \delta_p + \beta dist_i + \varepsilon_i, \tag{2}$$

where y_{ip} denotes that parcels are repeated for all pairs they are part of, and δ_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 0.10 and 0.05 miles 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 0.10 miles 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 2. Example of Matching

Notes: This map shows the sample of matched parcels in the same geographic area of Figure 1. Colored parcels are within 0.10 miles 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 2 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, Lester, and Reich 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.[14] 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, Gelbach, and Miller 2011).

The matching procedure and equation (2) apply without changes to the census block sample. However, because of their size, few census blocks are within 0.05 miles of their matches, so block-level estimates based on that distance are very imprecise.[15]

[14] Dube, Lester, and Reich (2010) compares contiguous pairs of counties located in different US 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.

[15] For example, Minneapolis has only four pairs of matched census blocks within 0.05 miles of each other.

Table 1—Parcel Summary Statistics and Covariate Balance

	Sample mean	Specification				
		OLS	Boundary FEs		Matched pair FEs	
Distance to boundary/match	Any (1)	Any (2)	<0.10 mi (3)	<0.05 mi (4)	<0.10 mi (5)	<0.05 mi (6)
Lot size (sq ft)	6,440	4,399	333	30	−64	86
		(509)	(207)	(196)	(229)	(215)
Building area (sq ft)	4,295	−658	−163	−218	−53	−11
		(117)	(140)	(151)	(207)	(201)
Residential area (sq ft)	2,938	−749	−128	−27	−9	−55
		(155)	(177)	(218)	(287)	(290)
Value of buildings ($1K)	306	−148	−17	1	−25	−24
		(30)	(18)	(25)	(37)	(41)
Value of land ($1K)	175	−13	9	5	−1	4
		(18)	(7)	(8)	(7)	(8)
Units	1.74	−0.94	−0.12	−0.09	−0.07	−0.05
		(0.12)	(0.13)	(0.15)	(0.22)	(0.21)
Stories	2.06	−0.59	0.02	0.01	0.05	0.03
		(0.05)	(0.03)	(0.03)	(0.04)	(0.04)
Rooms	9.79	−2.22	−0.56	−0.36	−0.66	−0.64
		(0.28)	(0.33)	(0.33)	(0.55)	(0.56)
Owner occupied (fraction)	0.71	0.16	−0.01	−0.004	0.02	0.06
		(0.03)	(0.02)	(0.027)	(0.03)	(0.04)
F-test		38.56	1.30	0.78	0.86	0.79
p-value		0.00	0.23	0.63	0.56	0.62
Clusters		1,323	382	387	431	419
Observations	262,420	262,420	59,805	35,918	133,202	56,968

Notes: This table describes the parcel samples. Column 1 provides summary statistics. Each cell in columns 2 through 6 reports estimates from a separate regression on distance to the polling place, measured in miles. Estimates reported in column 2 are from bivariate regressions of parcel characteristics (in rows) on distance to the polling place. Columns 3 and 4 control for boundary fixed effects. The specification in columns 5 and 6 controls for matched-pair fixed effects. Standard errors are clustered by precinct boundary (separately by precinct boundary and precinct in columns 5 and 6) and reported in parentheses.

III. Summary Statistics and Balancing Exercises

Under my identification assumption, parcels within boundaries/matched pairs should share similar characteristics, on average. Consequently, using parcel characteristics as placebo outcomes in specifications (1) and (2) should yield small and insignificant estimates. Table 1 reports test results (columns 2 through 6) 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 3–6 are subject to the additional parcel and precinct boundary restrictions detailed in online Appendix C.

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 is for residential purposes. The mean value of

Table 2—Census Block Summary Statistics and Covariate Balance

		Specification				
	Sample mean	OLS	Boundary FEs		Matched pair FEs	
Dist. to boundary/match	Any (1)	Any (2)	<0.10 mi (3)	<0.05 mi (4)	<0.10 mi (5)	<0.05 mi (6)
Adult population	82.1	−41.8	−7.6	−5.7	−16.0	0.1
		(5.1)	(7.3)	(7.8)	(11.6)	(19.5)
Non-Hispanic whites	52.0	−20.7	−4.3	−2.2	−3.4	1.1
		(3.6)	(4.9)	(4.8)	(5.5)	(14.0)
Non-Hispanic blacks	10.8	−7.9	−1.4	−2.1	−7.7	−6.9
		(1.3)	(1.9)	(2.6)	(4.3)	(5.7)
Hispanics, all races	8.3	−7.6	−0.9	−1.5	−3.4	0.7
		(1.0)	(1.4)	(1.8)	(2.4)	(3.7)
Nonwhites/Hispanics (fraction)	0.32	−0.10	0.02	0.03	0.01	0.09
		(0.02)	(0.02)	(0.03)	(0.02)	(0.09)
Median HH income ($1K)	71.0	13.8	1.3	3.6	3.5	8.4
		(4.6)	(3.2)	(3.3)	(2.9)	(7.5)
Units without cars (fraction)	0.19	−0.15	0.002	−0.01	−0.01	−0.02
		(0.02)	(0.016)	(0.02)	(0.02)	(0.02)
HS non-completers (fraction)	0.12	−0.01	0.01	0.01	0.01	0.004
		(0.01)	(0.01)	(0.01)	(0.01)	(0.022)
F-test		15.21	0.49	0.54	1.11	0.85
p-value		0.00	0.86	0.83	0.35	0.56
Clusters		1,289	353	288	372	99
Observations	15,037	15,037	3,333	1,694	4,108	534

Note: This table replicates the covariate balance tests from Table 1 on census block samples.

buildings and land are $306,000 and $175,000, respectively, while 71 percent of the parcels in Boston and Minneapolis are owner occupied.[16]

Column 2 reports coefficients from bivariate regressions of parcel attributes on distance to the polling place measured in miles. Out of nine estimates, eight are significant at the 1 percent level, thus confirming that, unconditionally, parcels that are farther to polling places are different than closer ones.

By contrast, regressions that control for boundary (columns 3 and 4) or matched-pair fixed effects (columns 5 and 6) 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 and only two are significant at the 10 percent level.

Table 2 reports summary statistics (column 1) and balancing exercises (columns 2 through 6) for census block samples. The average census block has 82.1 residents aged 18 or older, of which 52 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

[16] 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.

(\$53,046) and state levels (respectively, \$66,866 and \$59,836 for Massachusetts and Minnesota). Nineteen percent of occupied residential units in the average census block have no cars, and 12 percent of residents 25 or older never completed high school.

Like for parcel attributes, bivariate regressions of census block characteristics on distance to the 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 or matched-pair fixed effects. F-tests of joint significance similarly support the conditional un-correlation of distance to the polling place.

IV. 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.

A. *Effects on Parcel-Level Ballots Cast*

Table 3 reports effects on parcel-level counts of ballots cast. Each panel corresponds to a different election, and each cell reports estimates from a separate regression. All regressions control for parcel and census block covariates.[17] Columns 1 and 2 of Table 3 report estimates from boundary fixed effects specifications. Columns 3 and 4 report matching estimates.

As shown in panel A, 2012 parcel-level impact estimates range from -0.325 to -0.401. The treatment standard deviation is 0.245 miles, so bounds on estimated effects per standard deviation are -0.08 and -0.098. 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 4–5 percent. All estimates are significant at the 1 percent level.

Results from other elections reveal an overall similar picture. Ranges of impact estimates are $[-0.234, -0.162]$, $[-0.282, -0.194]$, and $[-0.172, -0.108]$ in, respectively, 2013 (panel B), 2014 (panel C), and 2016 (panel D). Matching and

[17] 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, 0–1, 1–2, 2–3, taller than 3; right bounds are included), and number of rooms (categories: 0, 0–4, 4–7, 7–11, more than 11; right bounds are included). Census block and block group covariates are: total population aged 18 or older, percentage of nonwhite 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.

Table 3—Effects on Parcel-Level Votes Cast

	Specification			
	Boundary FEs		Matched pair FEs	
Dist. to boundary/match	<0.10 mi (1)	<0.05 mi (2)	<0.10 mi (3)	<0.05 mi (4)
Panel A. Votes cast in 2012				
Distance to polling place	−0.325	−0.332	−0.391	−0.401
	(0.090)	(0.109)	(0.117)	(0.127)
Mean dep. var.	2.04	2.13	2.27	2.24
Observations	59,805	35,918	133,202	56,968
Panel B. Votes cast in 2013				
Distance to polling place	−0.209	−0.191	−0.234	−0.162
	(0.041)	(0.051)	(0.070)	(0.073)
Mean dep. var.	1.01	1.05	1.05	1.08
Observations	45,519	27,148	95,642	39,732
Panel C. Votes cast in 2014				
Distance to polling place	−0.232	−0.194	−0.282	−0.212
	(0.055)	(0.069)	(0.084)	(0.091)
Mean dep. var.	1.43	1.47	1.55	1.53
Observations	59,805	35,918	133,202	56,968
Panel D. Votes cast in 2016				
Distance to polling place	−0.146	−0.148	−0.172	−0.108
	(0.056)	(0.069)	(0.077)	(0.075)
Mean dep. var.	1.40	1.44	1.55	1.51
Observations	42,754	26,906	98,640	45,166

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. All regressions control for parcel and census block covariates.

within-boundary estimates display comparable magnitudes, while the latter dominate in terms of precision.

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, online Appendix Figure A2 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 3 and online Appendix Figure A3 plot coefficients from, respectively, within-boundary and matching regressions along with 95 percent confidence intervals. Virtually all estimates support a statistically significant effect of distance to the polling place on parcel-level voter participation. Estimates from regressions that add boundary-specific

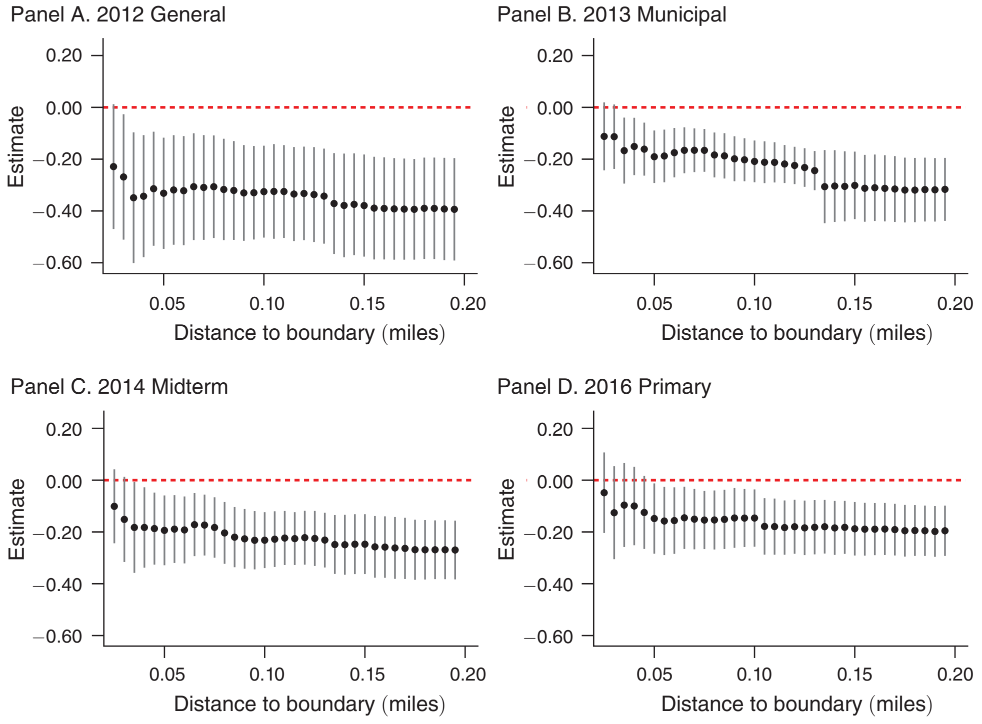


Figure 3. Within-Boundary Parcel-Level Estimates across Distances to Boundary

Notes: These figures plot estimated parcel-level treatment effects and 95 percent confidence intervals based on boundary fixed effects specifications across different distances to the nearest precinct border. Each pair of estimate and confidence interval comes from a separate regression.

interactions with latitude and longitude interactions—"interacted specifications" henceforth—(online Appendix Figure A4) are noisier and slightly smaller in magnitude than corresponding within-boundary estimates. Yet, confidence intervals of the three specifications largely overlap one another, suggesting that the underlying estimates are statistically similar to each other.

B. *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 test directly.

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

Table 4—Effects on Census Block Turnout and Registration

	Specification			
	Boundary FEs		Matched pair FEs	
	<0.10 mi	<0.05 mi	<0.10 mi	<0.05 mi
Dist. to boundary/match	(1)	(2)	(3)	(4)
Panel A. Census block turnout in 2012				
Distance to polling place	−0.113	−0.118	−0.086	−0.145
	(0.030)	(0.038)	(0.040)	(0.087)
Mean dep. var.	0.57	0.55	0.53	0.53
Observations	3,333	1,694	4,108	534
Panel B. Census block turnout in 2013				
Distance to polling place	−0.062	−0.064	−0.031	−0.028
	(0.021)	(0.030)	(0.029)	(0.055)
Mean dep. var.	0.30	0.28	0.28	0.30
Observations	2,546	1,222	2,916	334
Panel C. Census block turnout in 2014				
Distance to polling place	−0.072	−0.054	−0.047	−0.079
	(0.017)	(0.026)	(0.029)	(0.069)
Mean dep. var.	0.41	0.38	0.37	0.36
Observations	3,333	1,694	4,108	534
Panel D. Census block turnout in 2016				
Distance to polling place	−0.045	−0.045	−0.017	−0.099
	(0.016)	(0.025)	(0.022)	(0.061)
Mean dep. var.	0.35	0.34	0.33	0.33
Observations	2,370	1,404	3,312	526
Panel E. Census block registration in 2014				
Distance to polling place	−0.049	−0.022	−0.045	−0.032
	(0.020)	(0.030)	(0.038)	(0.119)
Mean dep. var.	0.82	0.81	0.79	0.78
Observations	3,333	1,694	4,108	534

Notes: This table reports estimates from regressions of census block turnout (panels A–D) and 2014 census block registration (panel E) on distance to the polling place. Panels represent outcomes defined by different elections. Each column represents a different combination of sample and specification. 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. All regressions control for parcel and census block covariates.

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. Depending on the combination of sample and specification, a 1 mile increase in distance to the polling place reduces turnout by 8.6–14.5 percentage points in the 2012 presidential (panel A) and 2.8–6.4 percentage points in the 2013 municipal elections (panel B). Ranges of estimated effects in the 2014 midterm and 2016 primary elections are, respectively, $[-7.9, -4.7]$ and $[-9.9, -1.7]$ percentage points (panels C and D). Relative to average outcomes, census block estimates are broadly consistent with corresponding figures from parcel regressions.[18] This suggests measurement error is not of critical concern at the block level.

Online Appendix Figure A5 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. Online Appendix Figures A6, A7, and A8 focus on within-boundary, matching, and interacted effects and plot point estimates along with 95 percent confidence intervals. The statistical significance of the estimates is supported by most combinations of specification and cutoff of distance to boundary/match.

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, a channel that seems particularly plausible in Minnesota, where voters have the option of registering on Election Day. To explore this possibility, 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 4, panel E, reports effects on voter registration from boundary fixed effects and matching specifications.[19]

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 2.2 (column 2) to 4.9 (column 1) percentage points, on average, which compares to an outcome mean of approximately 82 percent. Though large, registration effects should be taken with caution because only one out of four estimates reaches conventional levels of significance.

To investigate if registration effects in Minneapolis are driven by discouraged Election Day registrants, panel B of online Appendix Table A3 reports impact

[18] For example, take the mid-ranges of 2012 block- (i.e., $(-0.086 - 0.145)/2 = -0.116$) and parcel-level estimates (i.e., $(-0.325 - 0.401)/2 = -0.363$). These two figures compare to, respectively, an average census block turnout of 0.57 (i.e., $-0.116/0.57 = -0.203$) and 2.04 ballots cast per parcel (i.e., $-0.363/2.04 = -0.178$).

[19] Estimates from interacted specifications, available upon request, are virtually identical to within-boundary estimates.

estimates on parcel-level counts of Election Day registrants. Estimated effects range from -0.069 to -0.122, and they all reach (at least marginal) statistical significance. Estimates on the total number of registered voters, reported in panel A, are approximately four times as large, but they are imprecisely estimated. This suggests that voting costs can affect voter registration both by dissuading would-be Election Day registrants and through other forces, like voter purging. However, limited statistical power makes it impossible to assess the exact relative contribution of these channels.

In the online Appendix, I present three additional exercises related to the main estimates. First, I run placebo regressions that simultaneously control for own distance *and* distance to the polling location of units across the precinct boundary. Corroborating a causal interpretation of the estimates, distance to the polling location across the precinct boundary does not affect voter participation. Second, I construct RD-like plots of residualized outcomes and covariates using pairs of matched parcels and blocks. The running variable is distance to the matched unit (the negative of distance to the matched unit) for the unit that, within a pair, is relatively farther (closer) to its respective polling location. Third, I explore nonlinear treatment effects by replacing a single treatment with five dummies corresponding to different non-overlapping ranges of distance.

V. 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).[20] 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:

$$E\big[y_i|\mathbf{X}_i\big] = \exp\bigg(\delta_{b(i)} + \sum_{s\in\{L,H\}} \mathbf{1}\big(SES_{c(i)} = s\big) \times \big(\theta^s + \beta^s dist_i\big) + \mathbf{X}'_{ic(i)}\eta\bigg), \tag{3}$$

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; $\mathbf{X}_{ic(i)}$ 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.[21] Table 5 reports estimates based on interactions

[20] Sample medians of the fraction minority, median household income, and car availability are 0.186, \$63,021, and 0.13, respectively. Percentage minority is defined as the proportion of nonwhite or Hispanic residents in the census block.

[21] Equation (3) is estimated by Poisson conditional maximum likelihood using $\sum_{i=1}^{N(b)} y_{ib}$ as a sufficient statistic for the boundary fixed effects $\delta_{b(i)}$. The consistency of the maximum likelihood estimator for β 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).

Table 5—Heterogeneous Effects by Census Characteristics

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)
Panel A. By percent minority								
Percent minority $\leq$ median	1.88	−0.159	1.04	−0.110	1.42	−0.064	1.33	−0.050
		(0.079)		(0.049)		(0.044)		(0.048)
Percent minority $>$ median	2.17	−0.165	0.98	−0.333	1.43	−0.269	1.45	−0.193
		(0.047)		(0.066)		(0.048)		(0.055)
F-test (within year)		0.01		8.34		13.0		5.47
p-value		0.93		0.00		0.00		0.02
F-test (across years)		3.89						
p-value		0.00						
Observations	59,805	59,805	45,519	45,519	59,805	59,805	42,754	42,754
Panel B. By median HH income								
Income $\leq$ median	1.99	−0.145	0.88	−0.268	1.29	−0.207	1.21	−0.179
		(0.045)		(0.059)		(0.055)		(0.067)
Income $>$ median	2.08	−0.177	1.17	−0.182	1.55	−0.134	1.54	−0.089
		(0.075)		(0.053)		(0.041)		(0.042)
F-test (within year)		0.19		1.74		1.68		1.95
p-value		0.67		0.19		0.19		0.16
F-test (across years)		1.05						
p-value		0.38						
Observations	59,805	59,805	45,519	45,519	59,805	59,805	42,754	42,754
Panel C. By percent units without cars								
Percent without cars $\leq$ median	1.67	−0.160	0.91	−0.136	1.29	−0.093	1.20	−0.018
		(0.077)		(0.054)		(0.043)		(0.042)
Percent without cars $>$ median	2.33	−0.163	1.08	−0.313	1.54	−0.251	1.52	−0.241
		(0.047)		(0.063)		(0.052)		(0.057)
F-test (within year)		0.00		4.86		6.50		13.82
p-value		0.97		0.03		0.01		0.00
F-test (across years)		5.02						
p-value		0.00						
Observations	59,805	59,805	45,519	45,519	59,805	59,805	42,754	42,754

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.

with minority (panel A), income (panel B), and car availability (panel C). Online Appendix Table A4 presents OLS equivalents of these interacted specifications. Results from parcel-level matching specifications are reported in online Appendix Table A5, while OLS estimates of heterogeneous block-level effects are shown in online Appendix Tables A6 and A7.

Except for the 2012 presidential election, panel A reveals significant heterogeneity by minority presence. Effects on high-minority parcels are -16.5, -33.3, -26.9, and -19.3 log points in the 2012, 2013, 2014, and 2016 elections, respectively, while corresponding effects on low-minority parcels are -15.9, -11, -6.4, and -5 log points. 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 5, 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 5, 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 higher-salience 2012 election, while almost the entire effect in 2016 concentrates among parcels with lower-than-median car availability.

Heterogeneous effects by minority presence, income, and car availability are similar to one another at least partly due to the high correlation between these variables. Consequently, it may be difficult to assess the relative contribution of each interacting characteristic. Indeed, as reported in online Appendix Tables A8 and A9, regressions that control simultaneously for the three interactions yield imprecise and mostly insignificant estimates. That is, I lack statistical power to quantify the extent to which heterogeneous effects are due to differential sensitivities to voting costs by white versus minority, rich versus poor, or motorized versus nonmotorized voters.

Online Appendices H and I report estimated effects by partisan affiliation and state, respectively. State-specific impact estimates offer no evidence of an immediate reduction of the (relative) impact of distance to the polling place in Minnesota after the state introduced no-excuse absentee voting in 2014.

VI. Effects on Magazine and Newspaper Subscriptions and FEC Contributions

Recent research (e.g., Shineman 2018) has argued that information acquisition is endogenous to participation; that is, mobilizing voters to participate also motivates them to become more politically informed. My setting allows for an indirect test of this proposition. Because of its detrimental impact on turnout, distance to the polling place may reduce voters' demand of political information, which could take the form of fewer subscriptions to magazines or newspapers. A within-boundary estimate of the effect of distance to the polling place on the parcel-level count of magazine or newspaper subscribers is reported in column 1 of Table 6.

Distance to the polling place appears to have no effect on the number of newspaper and magazine subscribers, on average. The point estimate is an insignificant and precisely estimated -0.028 subscribers for a 1 mile increase in distance to the polling place, which compares to a sample mean of 0.541 subscribers per parcel.

Table 6—Effect on Magazine/Newspaper Subscriptions and Campaign Contributions

	Magazine, newspaper subscribers (1)	All contributors (2)	Contributors to Republican candidates (3)	Contributors to Democratic candidates (4)
Distance to polling place	−0.028	0.002	−0.003	0.004
	(0.027)	(0.028)	(0.003)	(0.016)
Mean dep. var.	0.541	0.133	0.008	0.089
Observations	59,805	59,805	59,805	59,805

Notes: This table reports within-boundary estimates of the effect of distance to the polling place on parcel-level counts of magazine/newspaper subscribers (column 1), contributors to any FEC-registered candidate or political organization (column 2), contributors to Republican candidates (column 3), and contributors to Democratic candidates (column 4).

Crucially, however, the zero effect on the number of magazine and newspaper subscribers does not disprove theories of endogenous information, as several reasons consistent with endogenous information could underlie the lack of effects. For example, the outcome variable, which aggregates newspaper and magazine subscriptions without distinction of outlet types, may be a poor proxy for the amount of political information acquired by voters. Alternatively, voter participation effects may fall disproportionately on voters who do not gather political information from printed medias. Either way, these limitations of the outcome variable caution against using the results reported in Table 6 to draw stark conclusions on theories of endogenous information.

Table 6 also documents a zero effect of distance to the polling place on the number of contributors to candidates and political organizations during the 2010–2012 election cycle. I find no effect on the number of people who made any FEC-recorded contribution to candidates or PACs (column 2), on the number of contributors to Republican candidates (column 3), or on the number of contributors to Democratic candidates (column 4). Though, again, the lack of effects could arise from the asymmetry between the typical set of voters who make political contributions (i.e., high-propensity-to-vote individuals) and the low-propensity voters who are deterred by distance to the polling place.

VII. Efficient Re-drawing 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 contains an equal number of residents. 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 re-precincting. Motivated by this observation, I examine if efficiently re-drawing existing precinct lines can reduce distance to the polling place, on average. Here, I summarize the problem setup and the findings; online Appendix J describes the technical details.

I start with a set of conservative but realistic assumptions about the re-precincting problem faced by election administrators. First, I fix the set of available polling sites to the actual 2012 polling stations. Second, I assume that each station cannot serve more than the number of voters it served in the 2012 election. Third, I assume that an administrator's objective is to minimize the sum of population-weighted distances between census blocks and polling stations.[22]

Efficiently re-drawing precinct lines reduces distance to the polling place by 0.035 miles, on average (from its 2012 mean of 0.36 miles). This seemingly small improvement would increase turnout by 0.4 percentage points in 2012 and between 0.2 and 0.3 percentage points in the other elections. However, this higher turnout would not narrow participation gaps across high- and low-SES areas. Finally, I examine a hypothetical policy that erased the impact of distance to the polling place. Under this policy, turnout would be considerably higher and, in non-presidential elections, less unequal across low- and high-SES areas. For example, participation gaps between low- and high-minority areas would decrease by 1.7 percentage points in 2013 (starting from 13.3 percentage points), 2 percentage points in 2014 (from 17.9 percentage points), and 1.5 percentage points in 2016 (from 12.7 percentage points). In other words, the disproportionate impact of distance to the polling place in high-minority areas contributes to between 11 and 13 percent of the participation gap between low- and high-minority areas during non-presidential elections.

VIII. Conclusion

In nine 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 2–5 percent 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 re-draw 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 0.03 miles. This would be enough to raise average turnout by 0.2–0.4 percentage points, but not enough to

[22] Durán et al. (2018) uses a set of similar assumptions to simulate re-precincting in the 2013 Argentinian midterm elections.

narrow the turnout gap that separates voters across different socioeconomic strata. By contrast, a hypothetical benchmark policy that eliminated distance to the polling place would increase average turnout by 1.6–4 percentage points and narrow the turnout gap between low- and high-minority areas in non-presidential elections by as much as 11–13 percent. 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 a goal for future research on the determinants of voter participation.

REFERENCES

Avery, James M. 2015. "Does Who Votes Matter? Income Bias in Voter Turnout and Economic Inequality in the American States from 1980 to 2010." *Political Behavior* 37 (4): 955–76.

Avery, James M., and Mark Peffley. 2005. "Voter Registration Requirements, Voter Turnout, and Welfare Eligibility Policy: Class Bias Matters." *State Politics and Policy Quarterly* 5 (1): 47–67.

Biggers, Daniel R., and Michael J. Hanmer. 2015. "Who Makes Voting Convenient? Explaining the Adoption of Early and No-Excuse Absentee Voting in the American States." *State Politics and Research Quarterly* 15 (2): 192–210.

Black, Sandra E. 1999. "Do Better Schools Matter? Parental Valuation of Elementary Education." *Quarterly Journal of Economics* 114 (2): 577–99.

Braconnier, Céline, Jean-Yves Dormagen, and Vincent Pons. 2017. "Voter Registration Costs and Disenfranchisement: Experimental Evidence from France." *American Political Science Review* 111 (3): 584–604.

Brady, Henry E., and John E. McNulty. 2011. "Turning Out to Vote: The Costs of Finding and Getting to the Polling Place." *American Political Science Review* 105 (1): 115–34.

Cameron, A. Colin, Jonah B. Gelbach, and Douglas L. Miller. 2011. "Robust Inference with Multiway Clustering." *Journal of Business and Economic Statistics* 29 (2): 238–49.

Cameron, A. Colin, and Pravin K. Trivedi. 2013. *Regression Analysis of Count Data*. 2nd ed. Cambridge, UK: Cambridge University Press.

Dell, Melissa. 2010. "The Persistent Effects of Peru's Mining Mita." *Econometrica* 78 (6): 1863–1903.

Dell, Melissa, and Pablo Querubin. 2018. "Nation Building through Foreign Intervention: Evidence from Discontinuities in Military Strategies." *Quarterly Journal of Economics* 133 (2): 701–64.

Dube, Arindrajit, T. William Lester, and Michael Reich. 2010. "Minimum Wage Effects across State Borders: Estimates Using Contiguous Counties." *Review of Economics and Statistics* 92 (4): 945–64.

Durán, Guillermo, Mauro Giormenti, Mario Guajardo, Pablo M. Pinto, Pablo A. Rey, and Nicolás E. Stier-Moses. 2018. "Improving Access to Voting with Optimized Matchings." *Electoral Studies* 51: 38–48.

Dyck, Joshua J., and James G. Gimpel. 2005. "Distance, Turnout, and the Convenience of Voting." *Social Science Quarterly* 86 (3): 531–48.

Erikson, Robert S. 1981. "Why Do People Vote? Because They Are Registered." *American Politics Research* 9 (3): 259–76.

Franko, William W., Nathan J. Kelly, and Christopher Witko. 2016. "Class Bias in Voter Turnout, Representation, and Income Inequality." *Perspectives on Politics* 14 (2): 351–68.

Fujiwara, Thomas. 2015. "Voting Technology, Political Responsiveness, and Infant Health: Evidence from Brazil." *Econometrica* 83 (2): 423–64.

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

Gimpel, James G., Joshua J. Dyck, and Daron R. Shaw. 2004. "Registrants, Voters, and Turnout Variability across Neighborhoods." *Political Behavior* 26 (4): 343–75.

Hajnal, Zoltan L. 2009. *America's Uneven Democracy: Race, Turnout, and Representation in City Politics*. 1st ed. Cambridge, UK: Cambridge University Press.

Hajnal, Zoltan, and Jessica Trounstine. 2005. "Where Turnout Matters: The Consequences of Uneven Turnout in City Politics." *Journal of Politics* 67 (2): 515–35.

Haspel, Moshe, and H. Gibbs Knotts. 2005. "Location, Location, Location: Precinct Placement and the Costs of Voting." *Journal of Politics* 67 (2): 560–73.

Hicks, William D., Seth C. McKee, Mitchell D. Sellers, and Daniel A. Smith. 2015. "A Principle or a Strategy? Voter Identification Laws and Partisan Competition in the American States." *Political Research Quarterly* 68 (1): 18–33.

Hill, Kim Quaile, and Jan E. Leighley. 1992. "The Policy Consequences of Class Bias in State Electorates." *American Journal of Political Science* 36 (2): 351–65.

Hodler, Roland, Simon Luechinger, and Alois Stutzer. 2015. "The Effects of Voting Costs on the Democratic Process and Public Finances." *American Economic Journal: Economic Policy* 7 (1): 141–71.

Kaplan, Ethan, and Haishan Yuan. 2020. "Early Voting Laws, Voter Turnout, and Partisan Vote Composition: Evidence from Ohio." *American Economic Journal: Applied Economics* 12 (1): 32–60.

Keele, Luke J., and Rocío Titiunik. 2015. "Geographic Boundaries as Regression Discontinuities." *Political Analysis* 23 (1): 127–55.

Keele, Luke, and Rocío Titiunik. 2016. "Natural Experiments Based on Geography." *Political Science Research and Methods* 4 (1): 65–95.

Keele, Luke, Rocío Titiunik, and José R. Zubizarreta. 2015. "Enhancing a Geographic Regression Discontinuity Design through Matching to Estimate the Effect of Ballot Initiatives on Voter Turnout." *Journal of the Royal Statistical Society. Series A: Statistics in Society* 178 (1): 223–39.

Lavy, Victor. 2010. "Effects of Free Choice among Public Schools." *Review of Economic Studies* 77 (3): 1164–91.

Lijphart, Arend. 1997. "Unequal Participation: Democracy's Unresolved Dilemma Presidential Address, American Political Science Association, 1996." *American Political Science Review* 91 (1): 1–14.

McDonald, Michael P., and Samuel L. Popkin. 2001. "The Myth of the Vanishing Voter." *American Political Science Review* 95 (4): 963–74.

Minnesota Secretary of State. 2011. *2011 Minnesota Redistricting Guide*. St. Paul: Office of the Minnesota Secretary of State. https://www.sos.state.mn.us/media/1102/2011-redistricting-guide.pdf.

Minnite, Lorraine C. 2013. "Voter Identification Laws: The Controversy over Voter Fraud." In *Law and Election Politics: The Rules of the Game*, 2nd ed., edited by Matthew J. Streb, 88–133. New York: Routledge.

Nickerson, David W. 2015. "Do Voter Registration Drives Increase Participation? For Whom and When?" *Journal of Politics* 77 (1): 88–101.

Nyhan, Brendan, Christopher Skovron, and Rocío Titiunik. 2017. "Differential Registration Bias in Voter File Data: A Sensitivity Analysis Approach." *American Journal of Political Science* 61 (3): 744–60.

Pintor, Rafael López, and Maria Gratschew. 2002. *Voter Turnout Since 1945: A Global Report*. Stockholm: International Institute for Democracy and Electoral Assistance.

Shineman, Victoria Anne. 2018. "If You Mobilize Them, They Will Become Informed: Experimental Evidence that Information Acquisition Is Endogenous to Costs and Incentives to Participate." *British Journal of Political Science* 48 (1): 189–211.

Taylor, Steven L., Matthew S. Shugart, Arend Lijphart, and Bernard Grofman. 2014. *A Different Democracy: American Government in a Thirty-One-Country Perspective*. New Haven: Yale University Press.

Timpone, Richard J. 1998. "Structure, Behavior, and Voter Turnout in the United States." *American Political Science Review* 92 (1): 145–58.

Tobler, W.R. 1970. "A Computer Movie Simulating Urban Growth in the Detroit Region." *Economic Geography* 46 (S1): 234–40.

US Census Bureau. 1994. "Geographic Areas References Manual." www2.census.gov/geo/pdfs/reference/GARM/Ch11GARM.pdf (accessed October 10, 2019).

US Election Assistance Commission. 2015. *The EAC 2014 Election Administration and Voting Survey Comprehensive Report*. Silver Spring, MD: US Election Assistance Commission. https://www.eac.gov/assets/1/1/2014_EAC_EAVS_Comprehensive_Report_508_Compliant.pdf.

A Precinct Too Far: Turnout and Voting Costs – Online Appendix

Enrico Cantoni*

February 11, 2019

A 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, and held presidential primaries on March 1, 2016; in addition, the most populous cities in the two states (Boston and Minneapolis) 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.[1] Thus, the only legitimate way most voters had to cast their ballots was by traveling to their assigned polling places on Election Day.[2] Unlike the voters in Massachusetts, where an ex-

*Department of Economics, University of Bologna, Piazza Scaravilli 2, 40126, Bologna (BO), Italy (e-mail: enrico.cantoni@unibo.it).

[1]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).

[2]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 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.

cuse was required throughout the sample period, the registered voters in Minnesota no longer need an excuse to vote absentee from June 2014 forward.

B 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 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[3], 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 primary 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® census block shapefiles. I then use census block identifiers to retrieve: population counts and racial makeup by census block, median household income, the proportion of occupied residential units without a

[3]Because Minnesota does not record a voter's party affiliation, this variable is not available for Minneapolis.

car, and the fraction of high-school noncompleters by block groups.[4]

Data on political contributions are from Bonica (2013), which collects every contribution registered in the Federal Election Commission (FEC) public records and made by individuals or organizations to local, state, and federal elections from 1979 to 2012. I restrict attention to contributions made by individuals during the 2010–2012 election cycle. Each record contains a contributor's ID, along with the latitude and longitude of the contributor's address.[5] Because of geocoding approximation, address coordinates often correspond to points in front of (i.e., on the street), rather than inside, the parcels containing the addresses. For this reason, I use ArcGIS to assign each geocoded contribution to its closest parcel polygon. I then construct three outcomes: the parcel-level count of all individuals who made any FEC-recorded contribution; the count of contributors to Republican candidates to local, state, or federal offices; and the count of contributors to Democratic candidates.[6]

Data on newspaper and magazine subscriptions were purchased from InfoUSA. InfoUSA uses a variety of sources, including actual subscription records from an undisclosed number of magazines and newspapers, to estimate the probability that individuals are currently subscribed to at least one magazine or newspaper. Each record contains the geocoded latitude and longitude of a likely subscriber's address. Similarly to FEC contributions, I match subscribers' address points to the nearest parcel polygons. Then, I use the total number of likely subscribers living in each parcel as outcome variable. The data were obtained in April 2015 and are updated as of that date.

C 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[7] could amass

[4]Block-level total and adult population by race and ethnicity come from, respectively, Tables P9 and 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.

[5]See the dataset codebook for details on the geocoding procedure. https://sdr.stanford.edu/uploads/tm/608/bd/7390/tm608bd7390/content/dime_codebook_v1.pdf Accessed: October 9, 2016.

[6]Results are substantively unchanged when outcomes are defined as the corresponding dollar amounts donated by parcel residents.

[7]The address locator is the dataset containing address attributes and geographic coordinates (typically, latitude and longitude) that serves as a crosswalk between addresses and geographic coordinates.

groups of geocoded addresses 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."[8] I start by standardizing voter addresses following the conventions used by MassGIS and Hennepin GIS for their address point shapefiles.[9] 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 percent 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.[10]

Analysis samples satisfy several restrictions. First, samples of parcels are limited to residential lots whose area does not exceed 70,000 square feet.[11] 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 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 129

[8]For a review of the superior precision of address-point matching relative to alternative geocoding techniques, see Zandbergen (2008).

[9]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.

[10]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.

[11]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.

blocks where the number of cast ballots in one or more elections exceeds the 2010 VAP.[12] 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 A1 plots the distribution of distance to the polling place in the parcel sample. The average residential parcel has a distance of 0.365 mile to its polling place, with a standard deviation of 0.245. Because the sample consists of densely populated urban areas, the overwhelming majority of parcels are assigned to polling locations that are less than 0.5 mile away.[13]

D 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 Dell (2010); Dell and Querubin (2018) and Gelman and Imbens (2018), I augment regression (1) with boundary-specific linear polynomials in latitude and longitude:

$$y_i = \delta_{b(i)} + \gamma^{lat}_{b(i)} latitude_i + \gamma^{long}_{b(i)} longitude_i + \beta dist_i + \varepsilon_i, \tag{1}$$

where $\gamma^{lat}_{b(i)}$ and $\gamma^{long}_{b(i)}$ denote the boundary-specific coefficients on parcel i's latitude and longitude, respectively. These boundary-specific interactions are the RD polynomial, which controls for relevant factors (besides the treatment) that vary smoothly across precinct boundaries. I refer to equation 1 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.

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 1. In my context, all boundaries are shorter than 1 mile, and the large sample size allows to restrict attention to parcels located within 0.05 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

[12] These are typically census blocks that contain large residential buildings constructed after 2010 (i.e., the year the decennial census was published).

[13] A regression of distance to the polling place on boundary dummies yields a residual standard deviation of 0.17 mile. Adding boundary-specific linear polynomials in latitude and longitude reduces the residual standard deviation to 0.12 mile. Similarly, the residual standard deviation in the full matching sample is approximately 0.15 mile.

to the polling place. By contrast, Dell (2010); Dell and Querubin (2018); 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.

Moreover, equation 1 requires explicitly estimating two controls – one for latitude, one for longitude – for each precinct boundary. As estimation samples include about four hundred precinct boundaries, the total number of controls in interacted specifications is large, thus reducing statistical power. At the same time, the number of lat-long controls grows with the number of boundaries (and hence with sample size), thus potentially complicating statistical inference (e.g., Cattaneo et al., 2018). For these reasons, I limit the use of interacted specifications to robustness checks. Corroborating the limited role that the RD polynomial plays in my design, balancing tests (available upon request) and main results from interacted specifications are substantively in line with within-boundary estimates.

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 1 interacts latitude and longitude with city (instead of boundary) fixed effects.

E Placebo Regressions

In this appendix, I run placebo regressions to assess whether, even conditioning on boundary or matched-pair fixed effects, unobservable voter characteristics spuriously drive my impact estimates. Because balance checks in Tables 1 and 2 show that distance to the polling place is conditionally uncorrelated with parcel and block characteristics, omitted variable bias seems unlikely. Yet, maybe voters living close to the institutional buildings typically used as polling locations (e.g., schools, city halls) have higher sense of civic duty – and are thus more likely to vote – than those who live farther away, even if both sets of voters have the same education and income, on average. For example, teachers and public employees, who may have higher-than-average levels of civic engagement, may be more likely to live in proximity to schools or public buildings. To rule out this possibility, Table A2 reports estimates from regressions of the following forms:

$$y_i = \delta_{b(i)} + \beta dist_i + \gamma distOtherStation_i + \mathbf{X}'_{ic(i)}\eta + \varepsilon_i \quad (2)$$

$$y_{ip} = \delta_p + \beta dist_i + \gamma distOtherStation_i + \mathbf{X}'_{ic(i)}\eta + \varepsilon_i, \quad (3)$$

where equations 2 and 3 refer to boundary fixed effects and matching specifications, respectively. The two regressions simply augment the corresponding original specifications with distance to the polling station of units on the opposite side of the precinct boundary: $distOtherStation_i$.

To corroborate a causal interpretation of the main results, estimates of β should be virtually unaffected by the inclusion of $distOtherStation_i$, while estimates of γ should be small and insignificant. By contrast, if voters living close to schools and municipal buildings have relatively higher civic duty – and thus higher propensities to vote, independently of whether they are actually assigned to vote at that specific polling location –, estimates of γ should be negative and significant.

Two observations are required to correctly interpret the results. First, controlling for distance to own polling place is crucial. If I simply replaced $dist_i$ with $distOtherStation_i$, I would obtain *positive* and significant estimates. The reason is that the two measures are highly negatively correlated: within boundaries, moving away from one polling location means moving closer to the opposite polling place, on average. Second, because of the high correlation between $dist_i$ and $distOtherStation_i$, controlling for both variables sharply reduces the treatment variation available to estimate effects (see bottom of Table A2). This is particularly true for matching specifications, which exploit *within-pair* treatment variation.

Reassuringly, controlling for $distOtherStation_i$ leaves within-boundary estimates of β (columns 1–4) virtually identical to the main estimates reported in Tables 3 and 4. At the same time, the estimated effect of distance to the other polling place in the boundary is always small and insignificant. Matching specifications (columns 5–8) are less revealing, as including $distOtherStation_i$ renders the estimated β's insignificant while the estimated γ's span large confidence intervals. But this is unsurprising in light of the minuscule variation that, conditioning on matched-pair fixed effects, remains to simultaneously estimate the effects of $dist_i$ and $distOtherStation_i$. Overall, I find no evidence that my estimates are spuriously driven by unobservable correlates of living close to schools or polling places (independently of the actual assignment to vote at those sites).

F RD-Like Plots

Here, I present one-dimensional RD-like plots. Defining a one-dimensional running variable for within-boundary specifications is complicated. A possible candidate is distance to the boundary, assigning negative (positive) values to units that, within each boundary, fall on the side that is relatively closer (farther) to its respective polling station. However, maps are two-dimensional and whichever side is closer depends on the specific point used to compute distances to the two polling locations. Moreover, choosing an arbitrary point on the border (e.g., the midpoint of each border between voting precincts) may be misleading, as parcels and census blocks in the boundary may not concentrate around that point. Finally, even assuming there are sensible, non-arbitrary ways to define a running variable, it is not obvious how the resulting graphs would map to the within-boundary specifications presented in the paper.

These issues are largely absent in matching specifications: within each matched pair, there is always one unit that is relatively closer to its polling location, and one unit that is relatively farther. A natural running variable is thus distance to the matched unit (the negative of distance to the matched unit) for the unit that, within a pair, is relatively farther (closer) to its polling location.

Using this running variable, Figures A9 and A10 show that, within pairs, units that are relatively closer to polling places (left side of each plot) have markedly higher voter participation than units that are relatively farther (right side of each plot). To visualize the same variation captured by regression (2), the graphs plot residualized outcomes after partialling out matched-pair fixed effects. The solid red lines denote linear fits of residualized outcomes on the running variable, estimated separately on each side of a ± 0.15-mile neighborhood around the discontinuity that separates closer (left) vs. farther (right) units. Point clouds represent sample means of plotted variables by (equally spaced) bins of the running variable, where the number of bins is based on Calonico et al. (2015)'s IMSE-optimal estimator.

Figures A11 and A12 plot residualized covariates. Except for distance to the polling place (panel A of the two figures), there are no systematic differences in covariates across the two sides of the discontinuity. Any differences are small in magnitude and consistent with the conditional exogeneity of distance to the polling place documented in the balancing exercises (Tables 1 and 2).

G Non-Linear Effects

In this appendix, I report estimates from regressions that replace distance to the polling place with indicators for non-overlapping ranges of distance. Using samples of units within .10 mile to the nearest precinct border/match, I estimate, respectively, within-boundary and matching specifications of the following forms:

$$y_i = \delta_{b(i)} + \beta_i^{0.1\text{-}0.2\text{mi}} + \beta_i^{0.2\text{-}0.3\text{mi}} + \beta_i^{0.3\text{-}0.5\text{mi}} + \beta_i^{0.5\text{-}0.75\text{mi}} + \beta_i^{0.75\text{+mi}} + \mathbf{X}'_{ic(i)}\eta + \varepsilon_i,$$
$$y_i = \delta_p + \beta_i^{0.1\text{-}0.2\text{mi}} + \beta_i^{0.2\text{-}0.3\text{mi}} + \beta_i^{0.3\text{-}0.5\text{mi}} + \beta_i^{0.5\text{-}0.75\text{mi}} + \beta_i^{0.75\text{+mi}} + \mathbf{X}'_{ic(i)}\eta + \varepsilon_i$$

where $\beta_i^{0.1\text{-}0.2\text{mi}}$, $\beta_i^{0.2\text{-}0.3\text{mi}}$, $\beta_i^{0.3\text{-}0.5\text{mi}}$, $\beta_i^{0.5\text{-}0.75\text{mi}}$, and $\beta_i^{0.75\text{+mi}}$ denote fixed effects for whether parcel or block i is within $0.1-0.2$, $0.2-0.3$, $0.3-0.5$, $0.5-0.75$, or

0.75+ mile to its polling station. The omitted category is being within 0–0.1 mile to one's polling place. Figures A13 and A14 report estimates from within-boundary and matching specifications, respectively. In each figure, Panel A reports the estimated β_i's and 95-percent confidence intervals from four boundary parcel-level regressions (i.e., one regression per election); panel B reports analogous estimates from block-level regressions.

In Figure A13, the estimated effects appear to grow linearly with distance to the polling place. The only possible exception is the seemingly "exponential" drop in participation going from 0.3-0.5mi to 0.5-0.75mi, which is particularly visible in census block regressions. This drop is perhaps explained by a combination of two factors. First, the maximum distance voters in my sample are willing to walk to cast a ballot may be in the 0.3-to-0.5 mile range.[14] Second, there may be a fixed cost associated with driving to the polls (e.g., the time necessary to find parking). If so, distances beyond walkability may induce a fraction of voters to drive instead of walking; at the same time, these distances may induce voters with large driving fixed costs to abstain entirely. Albeit noisier, patterns of matching estimates in Figure A14 are substantively in line with corresponding within-boundary estimates.

H 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 A10 reports the results.

In every election, proportional effects on votes cast by Democrats and unaffiliated/third-

[14]Incidentally, 1/4 and 1/2 mile are the two standard measures of "walkability" used in the United States Green Building Council, LEED 2009 guidelines; for example, see: https://www.usgbc.org/credits/lt32 Accessed: October 3, 2018.

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 percent. 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 10.9 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 5.3 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-percent level in 2014 and at the 10-percent level in 2016, while a joint test of equal proportional effects across the four elections is marginally significant.

I Effects by State

Does absentee voting alleviate the negative turnout effect of distance to the polling place?[15] 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:

$$E[y_i|\mathbf{X_i}] = \exp\left(\delta_{b(i)} + \beta^{MA} dist_i + \beta^{MN} dist_i + \mathbf{X}'_{ic(i)}\eta\right), \tag{4}$$

[15] 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.

where β^{MA} and β^{MN} denote state-specific proportional effects.[16] DD estimates, reported in Table A11, are then computed as $(\beta^{MN} - \beta^{MA})_{'14} - (\beta^{MN} - \beta^{MA})_{BaselineYr}$, 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.

Estimated proportional effects in the Massachusetts municipalities are remarkably stable across elections (respectively, -17.7, -18.1, and -15.1 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., $-0.215 + 0.151 = -0.064$) is more pronounced than the corresponding 2012 gap (i.e., $-0.112 + 0.177 = 0.065$), but less so than the 2013 difference (i.e., $-0.353 + 0.181 = -0.172$). 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.

J Efficient Redrawing of Precinct Boundaries: Technical Appendix

I formalize the reprecincting problem faced by election administrators as a generalized assignment problem (GAP, Fernández and Landete, 2015; Kundakcioglu and Alizamir, 2008). In each city, a finite set of census blocks, $J = \{1, \ldots, j, \ldots, n\}$, must be optimally allocated to a finite, predetermined set of polling places, $I = \{1, \ldots, i, \ldots, m\}$. 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

[16] Boundary fixed effects are defined within city, so they already incorporate the states main effects.

the problem realistic and consistent with the regulations discussed in Section 1.1, 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 the two terms indistinctly. 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 re-precincting.

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:

$$\text{minimize} \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} \tag{5}$$

$$\text{subject to} \sum_{i \in I} x_{ij} = 1 \qquad j \in J \tag{6}$$

$$\sum_{j \in J} d_J x_{ij} \leq q_i \qquad i \in I \tag{7}$$

$$x_{ij} \in \{0, 1\} \qquad i \in I, j \in J. \tag{8}$$

Constraints 6 and 8 guarantee that each census block is entirely assigned to exactly one precinct, while constraint 7 ensures that precinct capacities are not exceeded. These assumptions are 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, 2008 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.[17] Column 1 of Table A12 reports the average census block-to-polling-place distance (in miles). Column 2 shows the average difference between 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 into 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).[18]

[17]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.

[18]The simulated turnout effects of the two policies are computed using census block point estimates from Table A6 times the average distances shown in columns 1 and 2 of Table A12. Results are unchanged when I exclude Boston, which, as mentioned in Section 1.1, is exempted from the decennial requirement to redraw precinct lines.

References

Bonica, Adam, "Database on Ideology, Money in Politics, and Elections," 2013.

Calonico, Sebastian, Matias D. Cattaneo, and Rocío Titiunik, "Optimal Data-Driven Regression Discontinuity Plots," *Journal of the American Statistical Association*, 2015, *110* (512), 1753–1769.

Cattaneo, Matias D., Michael Jansson, and Whitney K. Newey, "Inference in Linear Regression Models with Many Covariates and Heteroscedasticity," *Journal of the American Statistical Association*, 2018, *0* (0), 1–12.

Dell, Melissa, "The Persistent Effects of Peru's Mining Mita," *Econometrica*, 2010, *78* (6), 1863–1903.

_ **and Pablo Querubin**, "Nation Building Through Foreign Intervention: Evidence from Discontinuities in Military Strategies," *The Quarterly Journal of Economics*, 2018, *133* (2), 701–764.

Fernández, Elena and Mercedes Landete, "Location Science," in Gilbert Laporte, Stefan Nickel, and Francisco Saldanha da Gama, eds., *Location Science*, 2015 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.

Gelman, Andrew and Guido Imbens, "Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs," *Journal of Business & Economic Statistics*, 2018, *0* (0), 1–10.

Gronke, Paul, Eva Galanes-Rosenbaum, and Peter A. Miller, "Early Voting and Turnout," *PS: Political Science & Politics*, 2007, *40* (4), 639–645.

Karp, Jeffrey A. and Susan A. Banducci, "Absentee Voting, Mobilization, and Participation," *American Politics Research*, 2001, *29* (2), 183–195.

Kundakcioglu, Erhun O. and Saed Alizamir, "Generalized Assignment Problem," in Christodoulos A. Floudas and Panos M. Pardalos, eds., *Encyclopedia of Optimization*, 2nd ed., Springer US, 2008, pp. 1153–1162.

Larocca, Roger and John S. Klemanski, "U.S. State Election Reform and Turnout in Presidential Elections," *State Politics & Policy Quarterly*, 2011, *11* (1), 76–101.

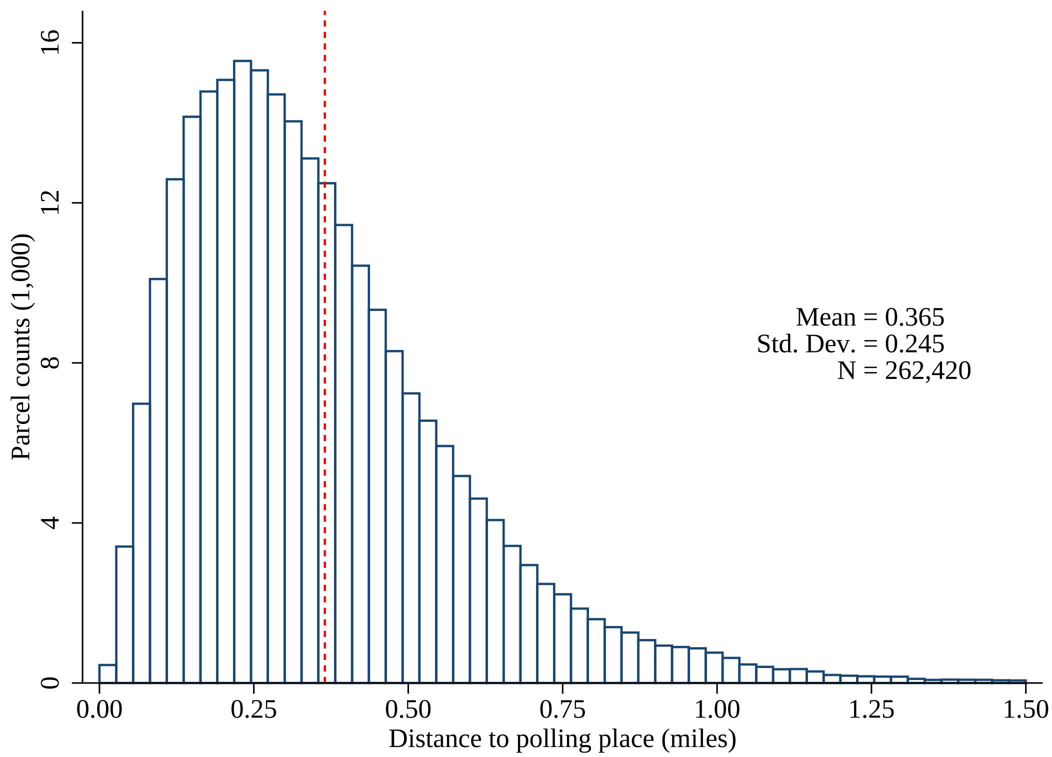
McNulty, John E., Conor M. Dowling, and Margaret H. Ariotti, "Driving Saints to Sin: How Increasing the Difficulty of Voting Dissuades Even the Most Motivated Voters," *Political Analysis*, 2009, *17* (4), 435–455.

Meredith, Marc and Zac Endter, "Aging into Absentee Voting: Evidence from Texas," 2015.

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

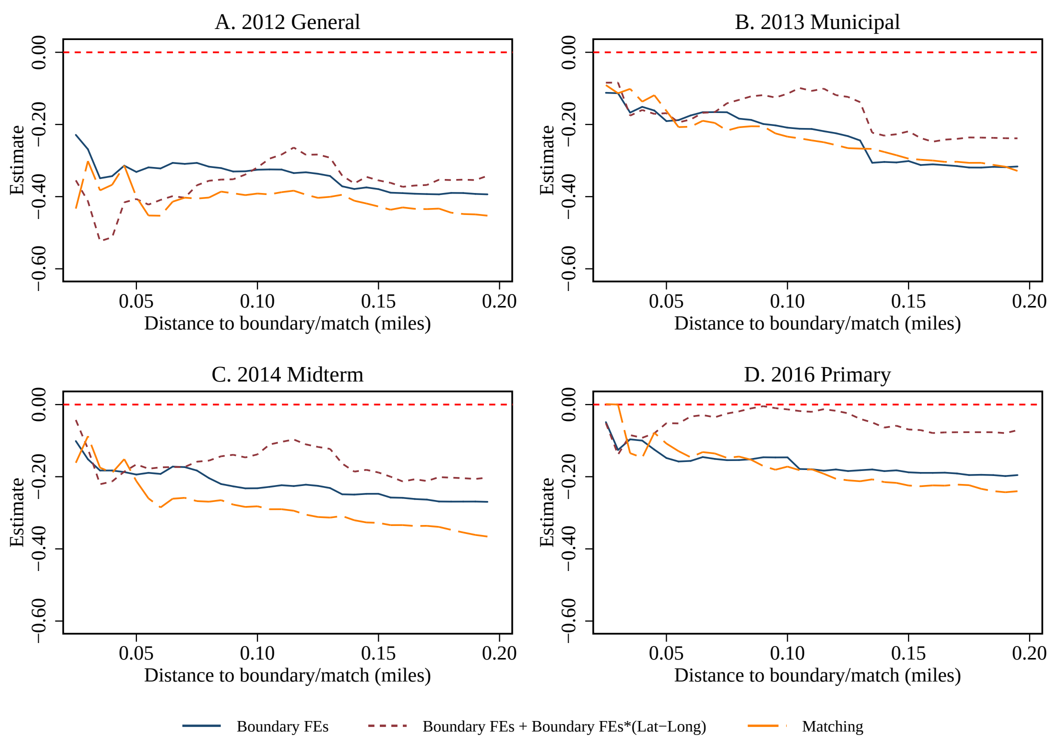
Zandbergen, Paul A., "A comparison of address point, parcel and street geocoding techniques," *Computers, Environment and Urban Systems*, 2008, *32* (3), 214–232.

Figure A1: Distance to the Polling Place



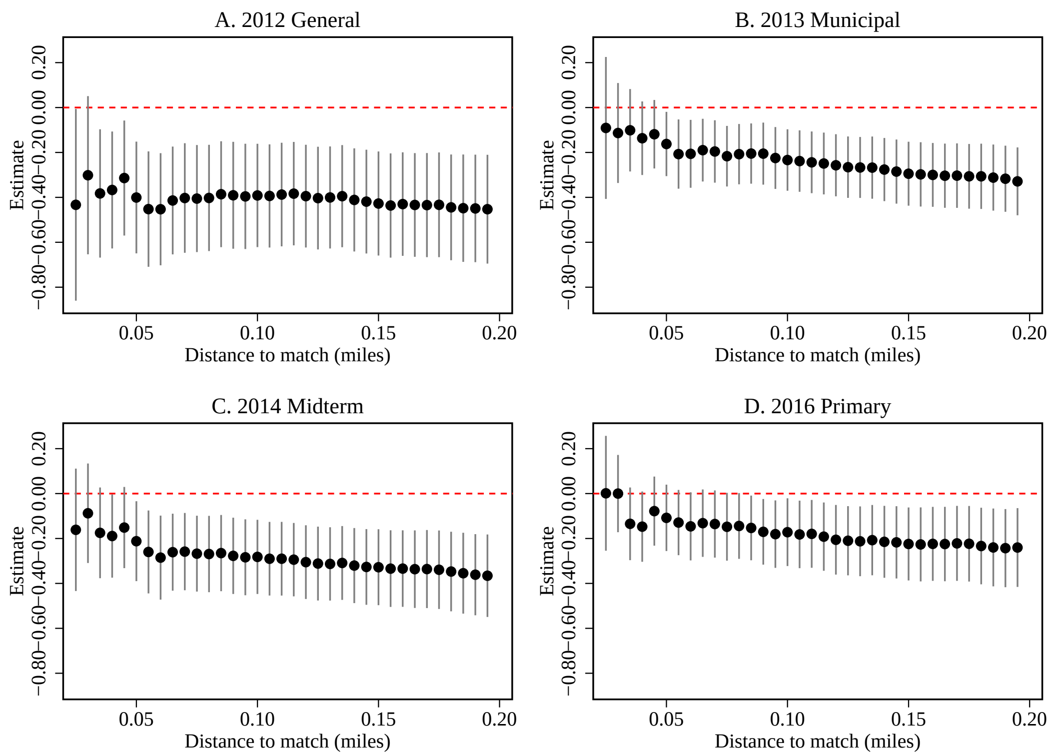
Notes: This histogram plots the distribution of distance to the polling place in the full parcel sample.

Figure A2: Sensitivity of Parcel Estimates to Distance to Boundary/Match



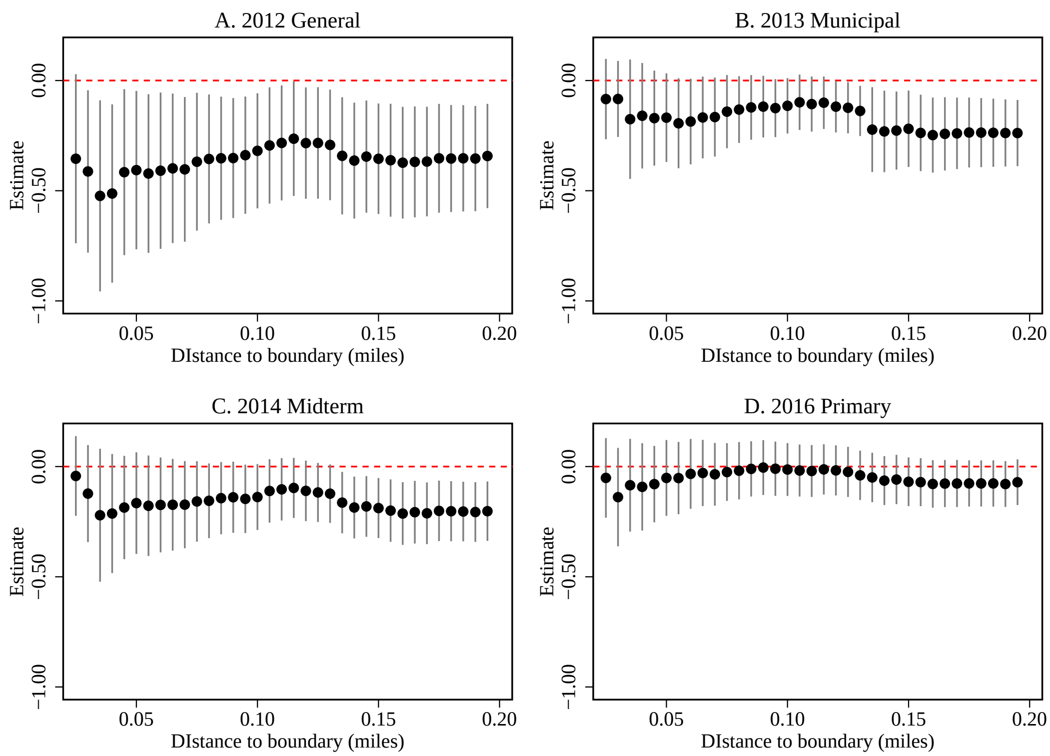
Notes: These figures plot estimated parcel-level treatment effects based on boundary fixed effects, boundary effects with lat-long interactions, and matching specifications across different bandwidths (i.e., distance to the nearest precinct border or distance to the matched unit). Different panels correspond to different elections.

Figure A3: Matching Parcel-Level Estimates Across Distances to Match



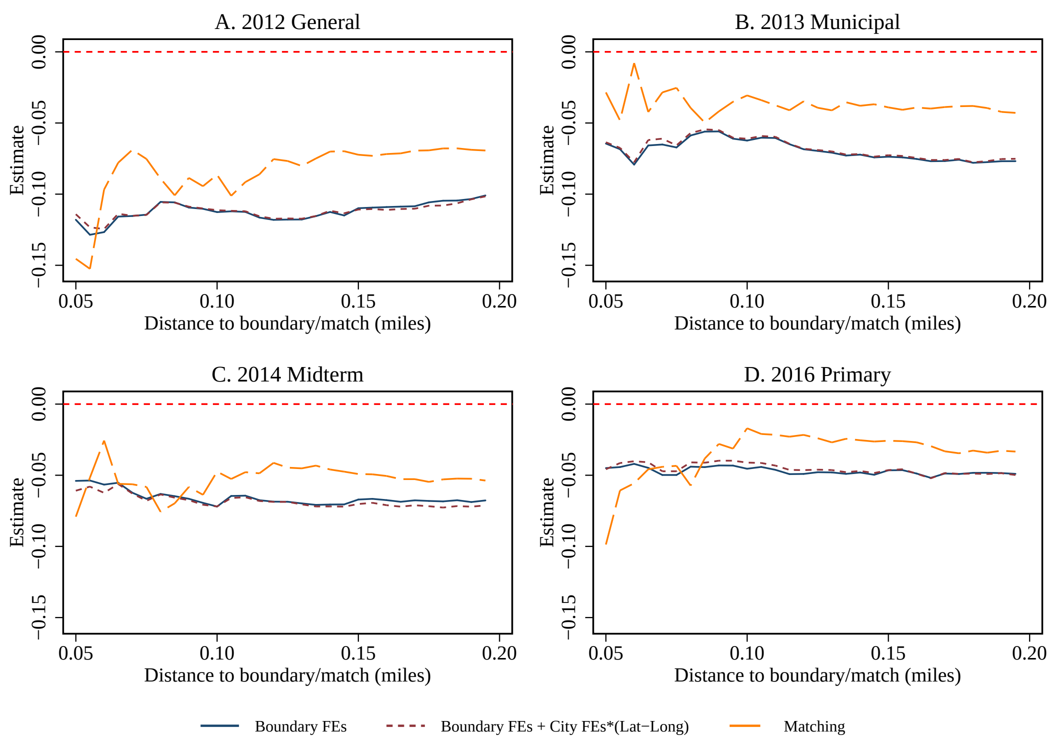
Notes: These figures plot estimated parcel-level treatment effects and 95-percent confidence intervals based on matching specifications across different distances to the matched unit. Each pair of estimate and confidence interval comes from a separate regression.

Figure A4: Within-Boundary Parcel-Level Estimates with Lat-Long Interactions Across Distances to Boundary



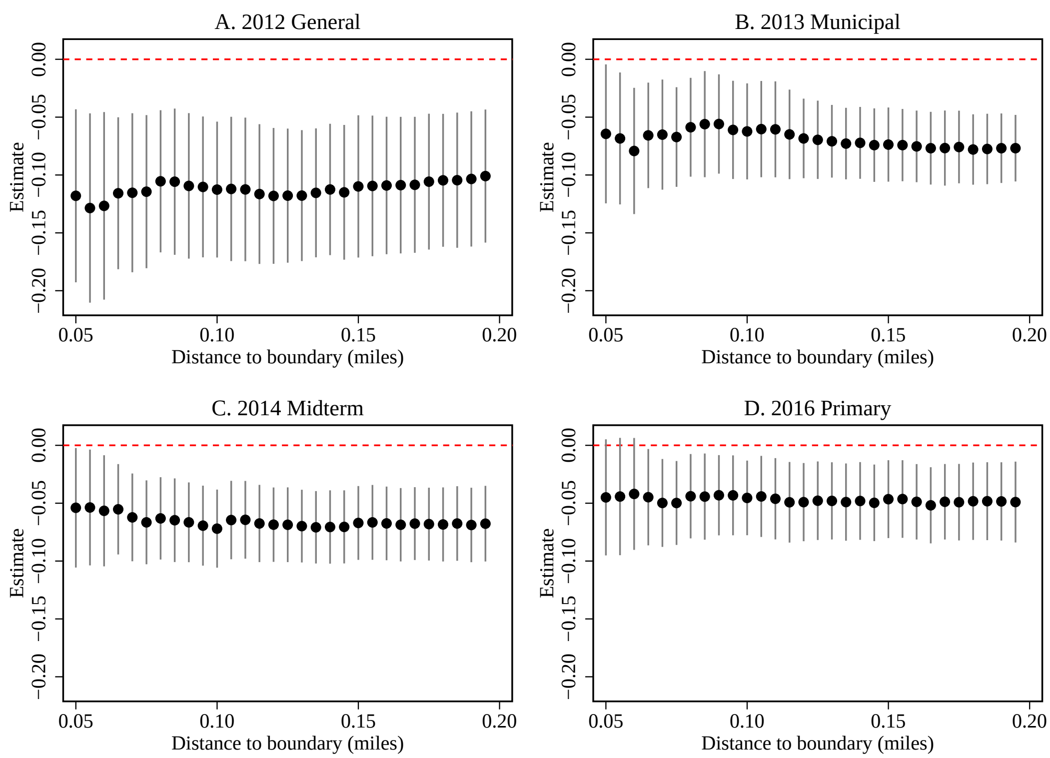
Notes: These figures plot estimated parcel-level treatment effects and 95-percent confidence intervals based on boundary fixed effects specifications with lat-long interactions across different distances to the nearest precinct border. Each pair of estimate and confidence interval comes from a separate regression.

Figure A5: Sensitivity of Census Block Estimates to Distance to Boundary/Match



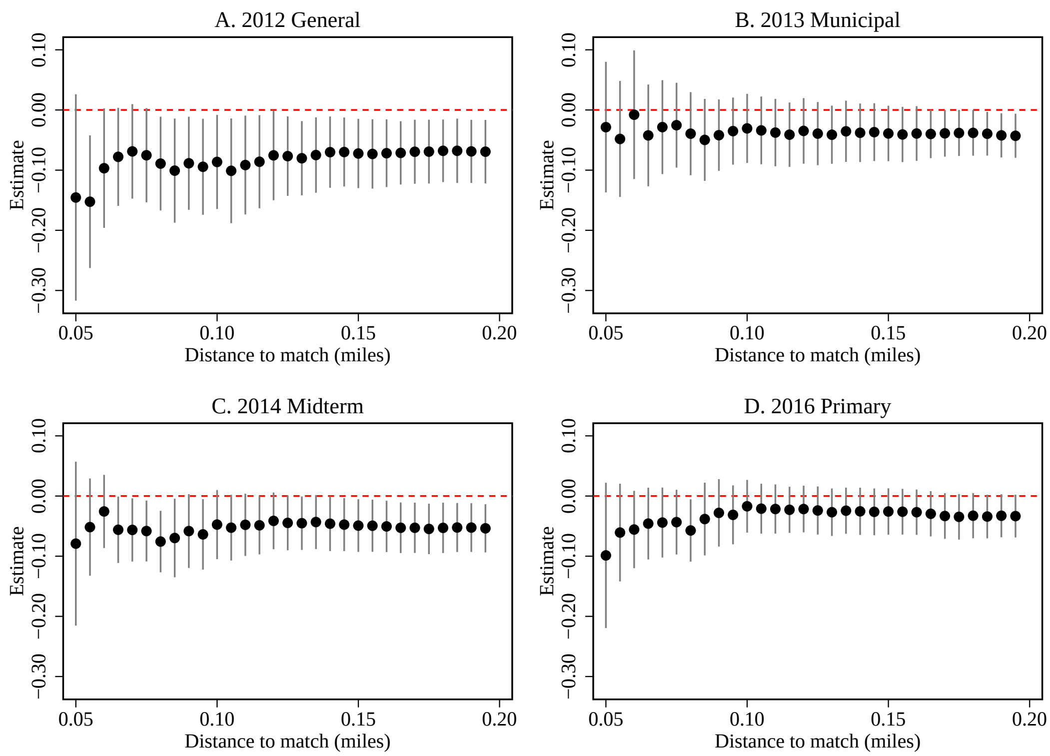
Notes: These figures plot estimated block treatment effects based on boundary fixed effects, boundary fixed effects with lat-long interactions, and matching specifications across different bandwidths (i.e., distance to the nearest precinct border or distance to the matched unit). Different panels correspond to different elections.

Figure A6: Within-Boundary Block-Level Estimates Across Distances to Boundary



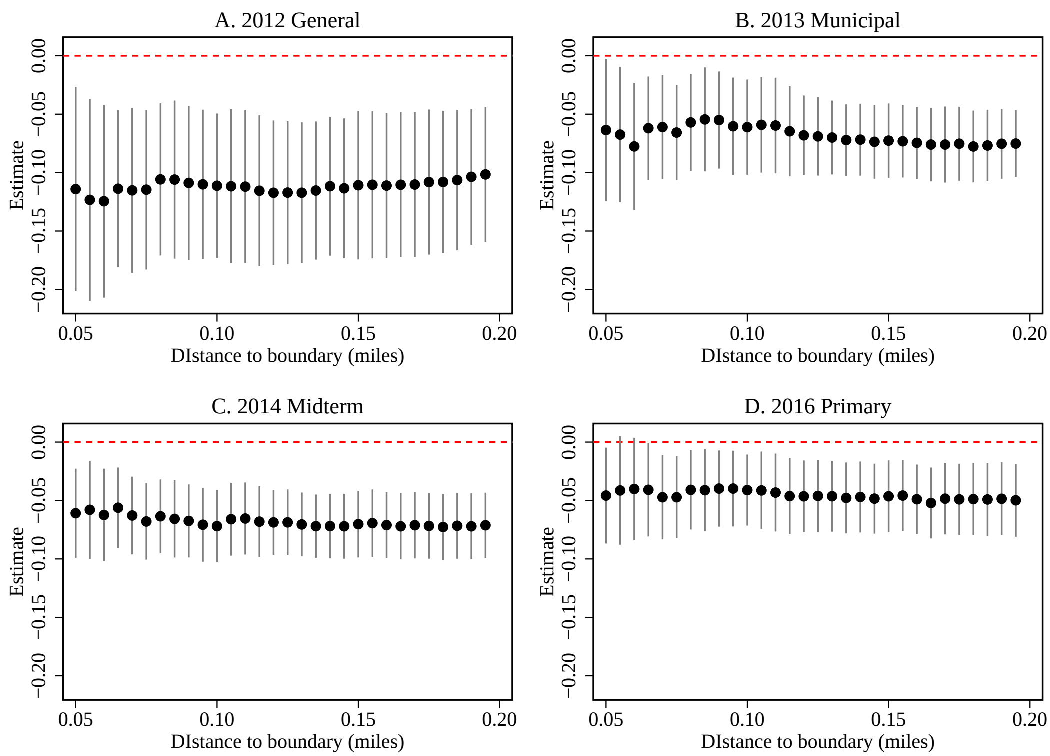
Notes: These figures plot estimated block-level treatment effects and 95-percent confidence intervals based on boundary fixed effects specifications across different distances to the nearest precinct border. Each pair of estimate and confidence interval comes from a separate regression.

Figure A7: Matching Block-Level Estimates Across Distances to Match



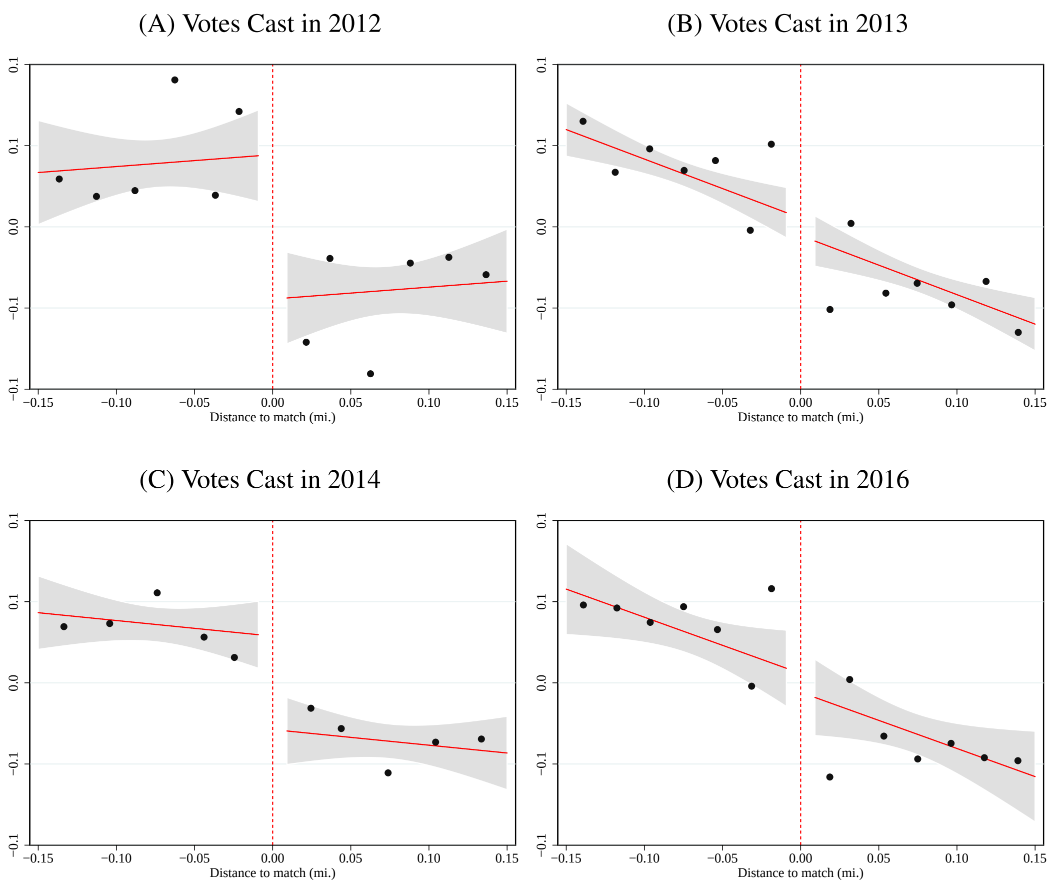
Notes: These figures plot estimated block-level treatment effects and 95-percent confidence intervals based on matching specifications across different distances to matched unit. Each pair of estimate and confidence interval comes from a separate regression.

Figure A8: Within-Boundary Block-Level Estimates with Lat-Long Interactions Across Distances to Boundary



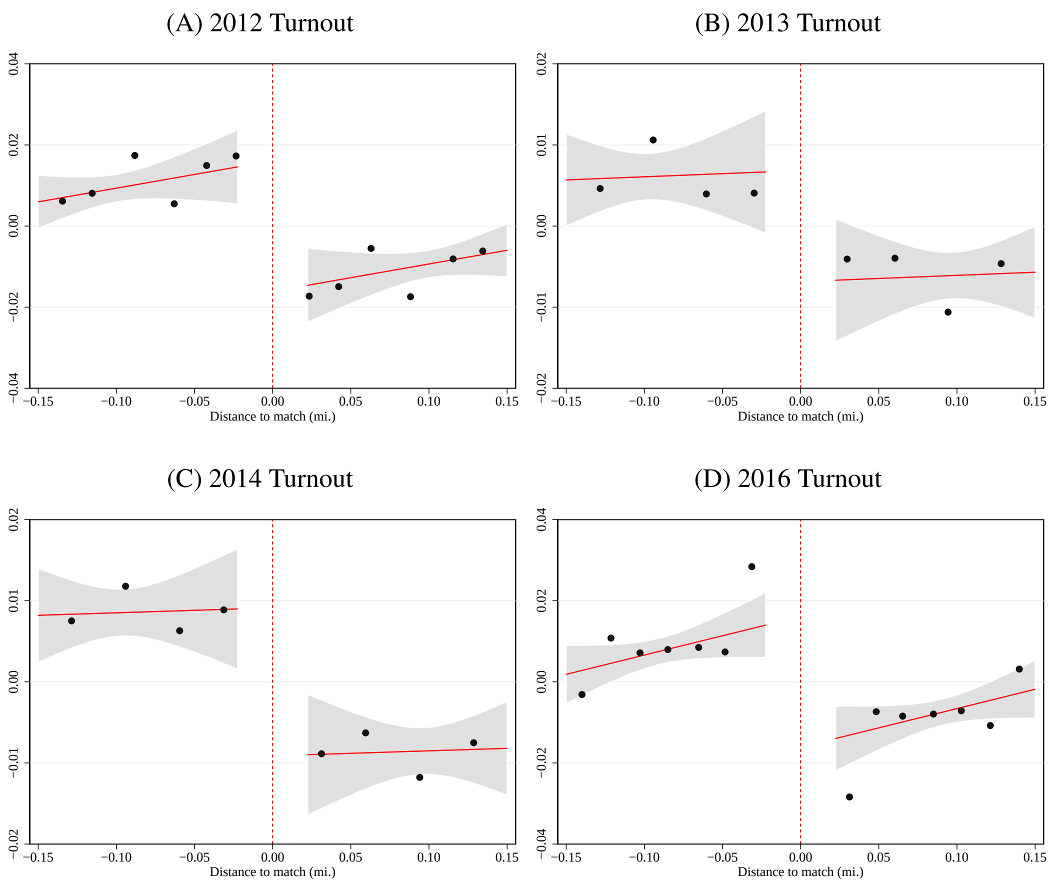
Notes: These figures plot estimated block-level treatment effects and 95-percent confidence intervals based on boundary fixed effects specifications with lat-long interactions across different distances to the nearest precinct border. Each pair of estimate and confidence interval comes from a separate regression.

Figure A9: Residualized Parcel Outcomes Against Distance to Match



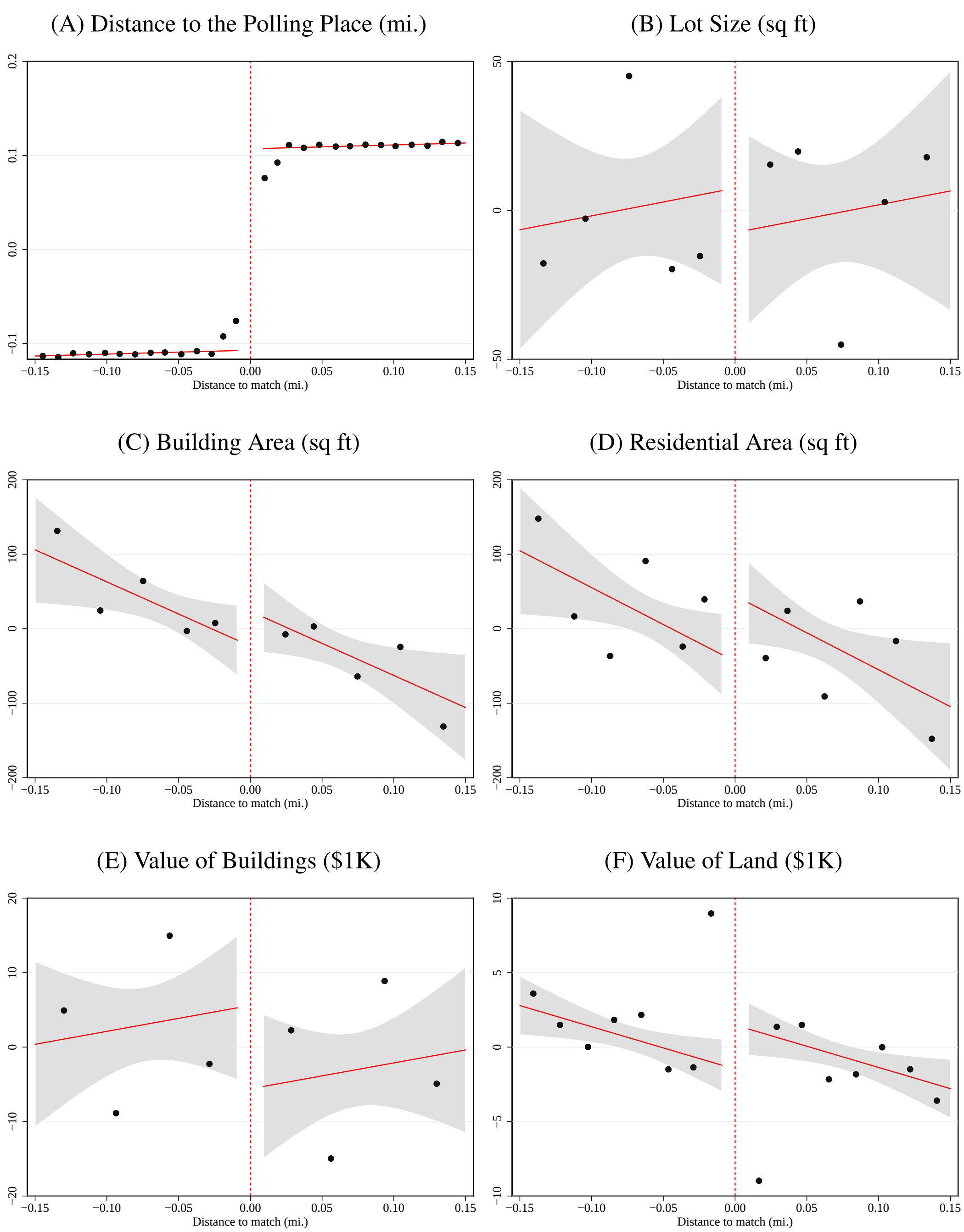
Notes: Using samples of matched parcels, these figures plot votes cast as a function of distance to the matched unit (or the negative thereof). Within each matched pair, the unit that is relatively closer to its polling place is assigned a negative value of distance to the match; the unit that is relatively farther to its polling place is assigned a positive value of distance to the match. Plotted variables are residualized after partialling out matched-pair fixed effects. Solid red lines are linear fits estimated separately on the two sides of the discontinuity. Shaded areas denote 95-percent confidence intervals. Point clouds are outcome means by (equally spaced) bins of the running variable, where the number of bins is based on Calonico et al. (2015)'s IMSE-optimal estimator.

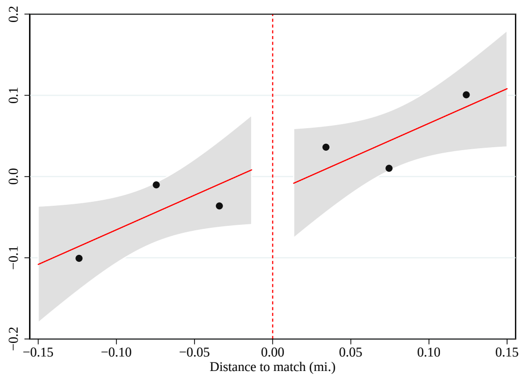
Figure A10: Residualized Census Block Outcomes Against Distance to Match



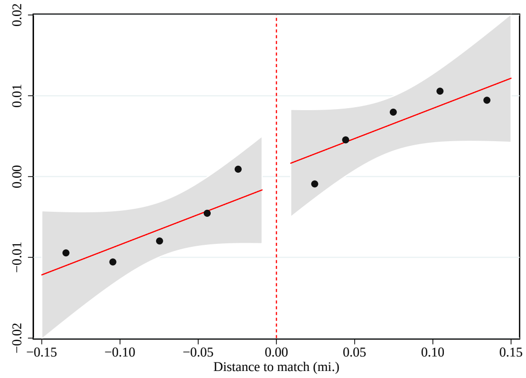
Notes: These figures are constructed in the same way as Figure A9 and plot residualized census block turnout.

Figure A11: Residualized Parcel Covariates Against Distance to Matched Unit

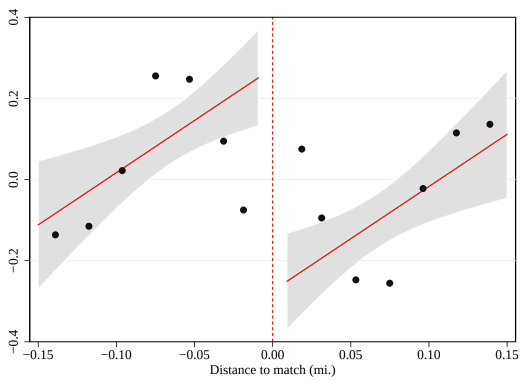




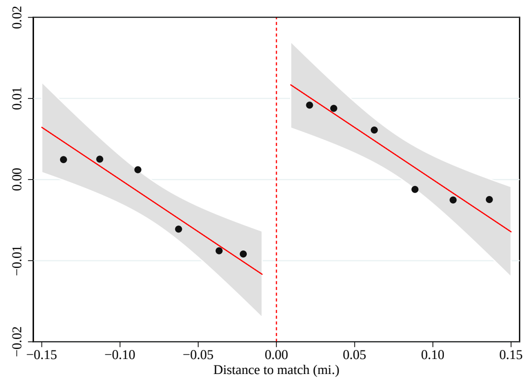
(G) Units



(H) Stories



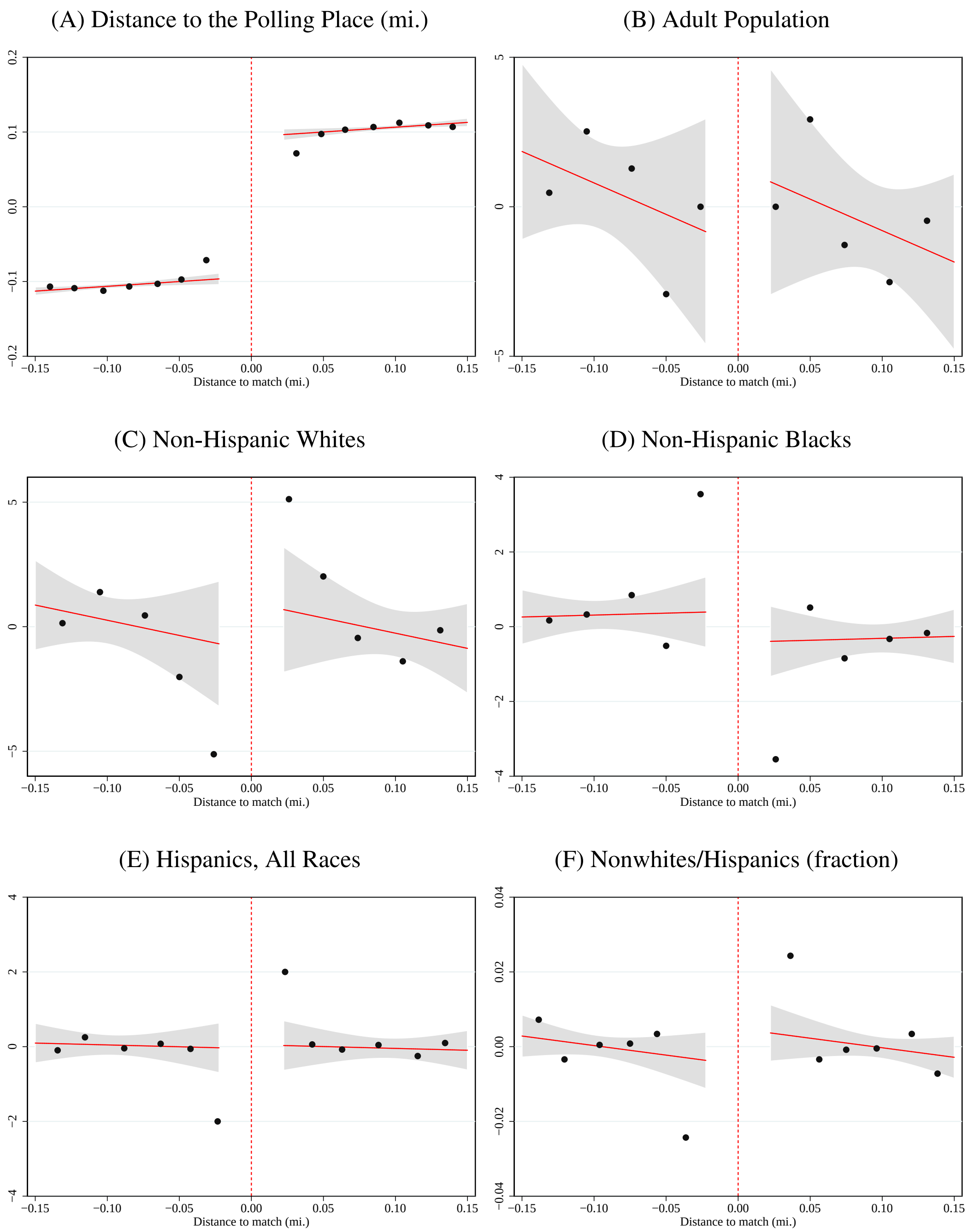
(I) Rooms

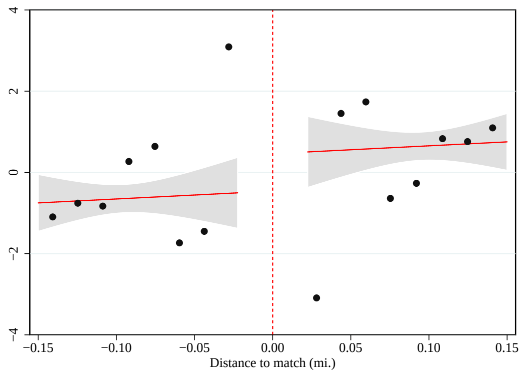


(J) Owner Occupied (fraction)

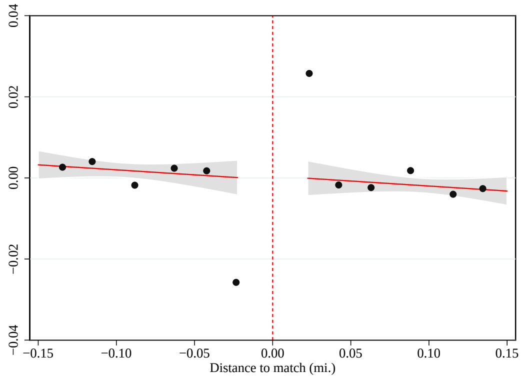
Notes: These figures are constructed in the same way as Figure A9 and plot residualized parcel covariates.

Figure A12: Residualized Census Block Covariates Against Distance to Matched Unit

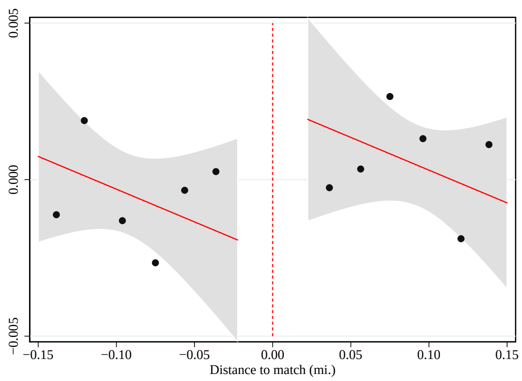




(G) Median HH Income ($1K)



(H) Units w/o Cars (fraction)

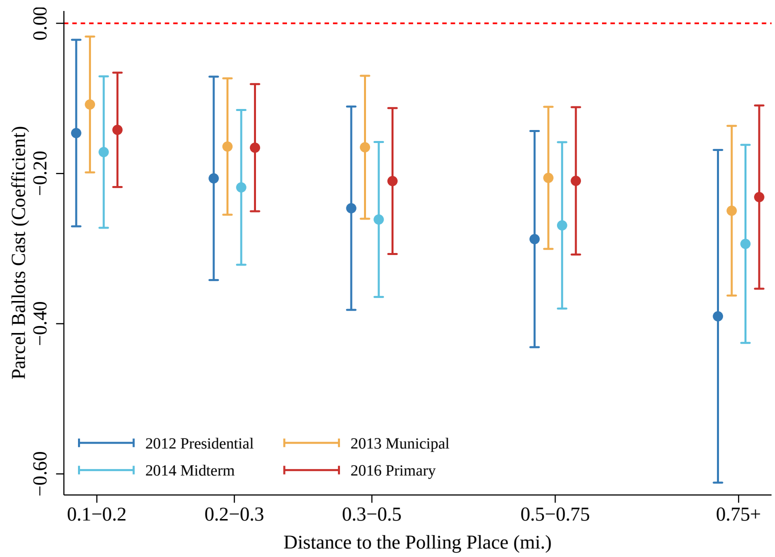


(I) High-School Noncompleters (fraction)

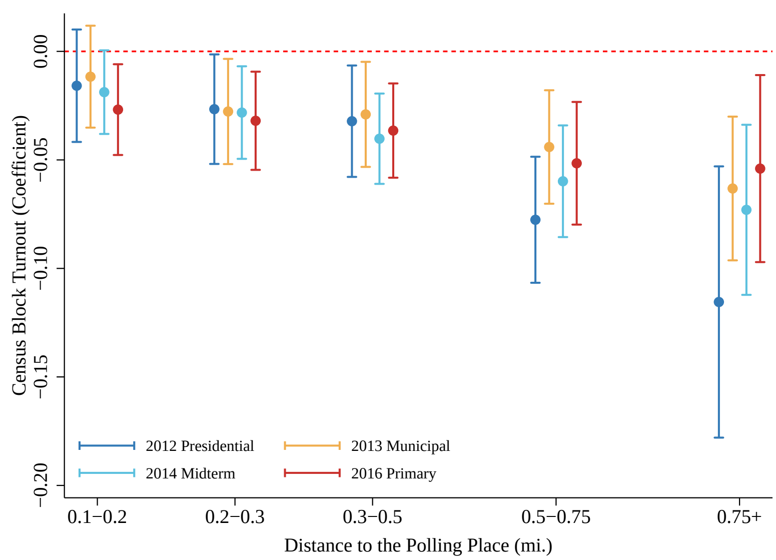
Notes: These figures are constructed in the same way as Figure A9 and plot residualized census block covariates.

Figure A13: Non-Linear Treatment Effects – Within-Boundary Estimates

(A) Non-Linear Effects on Parcel Votes Cast



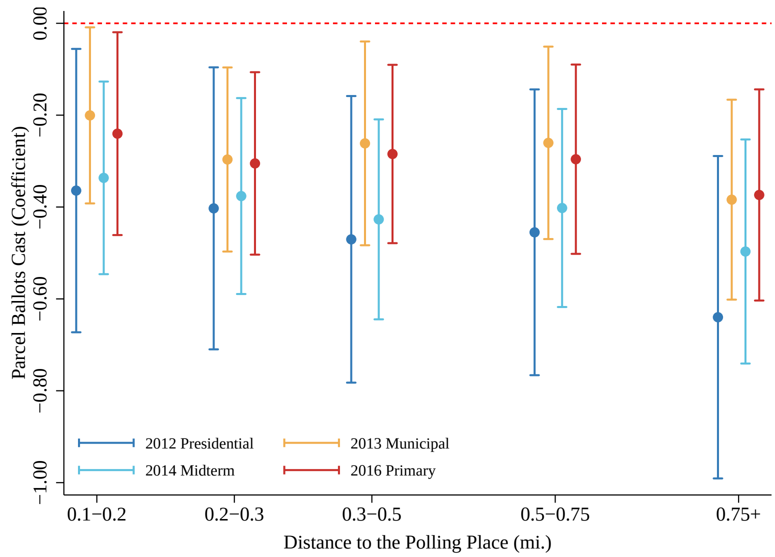
(B) Non-Linear Effects on Census Block Turnout



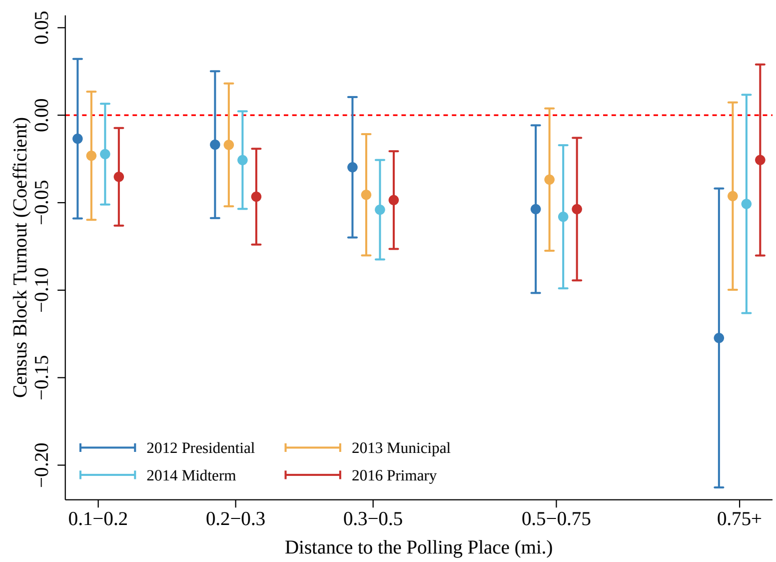
Notes: These figures plot estimates and 95-percent confidence intervals of non-linear treatment effects on the number of ballots cast by parcel residents (Panel A) and block-level voter turnout (Panel B). Each panel reports estimates from four distinct regressions, one for each election. All regressions are boundary fixed effects specifications run on 0.10-mile-to-boundary samples that control for five mutually exclusive dummies corresponding to different ranges of distance to the polling place. The omitted category is distance to the polling place between 0 and 0.1 mile.

Figure A14: Non-Linear Treatment Effects – Matching Estimates

(A) Non-Linear Effects on Parcel Votes Cast



(B) Non-Linear Effects on Census Block Turnout



Notes: These figures plot estimates and 95-percent confidence intervals of non-linear treatment effects on the number of ballots cast by parcel residents (Panel A) and block-level voter turnout (Panel B). Each panel reports estimates from four distinct regressions, one for each election. All regressions are matching specifications run on 0.10-mile-to-match samples that control for five mutually exclusive dummies corresponding to different ranges of distance to the polling place. The omitted category is distance to the polling place between 0 and 0.1 mile.

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

	Discontinuity Sample		Placebo Sample	
Dist. to Bound.:	<0.10 mi	<0.05 mi	<0.10 mi	<0.05 mi
	(1)	(2)	(3)	(4)
	Panel A. Boundary FEs			
R^2	0.54	0.56	0.70	0.73
	Panel B. Boundary FEs + City FEs×(Lat-Long)			
R^2	0.55	0.57	0.76	0.78
	Panel C. Boundary FEs + City FEs×(Lat-Long)^2			
R^2	0.56	0.58	0.80	0.81
	Panel D. Boundary FEs + City FEs×(Lat-Long)^3			
R^2	0.57	0.59	0.80	0.82
	Panel E. Boundary FEs + Boundary FEs×(Lat-Long)			
R^2	0.79	0.79	0.96	0.97
N	59,805	35,918	33,442	20,631

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: Placebo Effects of Distance to Other Polling Place in Boundary/Match

	Specification:							
	Boundary FEs				Matched Pair FEs			
Election:	2012 Presid.	2013 Munic.	2014 Midt.	2016 Primary	2012 Presid.	2013 Munic.	2014 Midt.	2016 Primary
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Panel A. Parcel Votes Cast							
Distance to own polling place	-0.355	-0.268	-0.285	-0.269	-0.997	0.130	-0.591	-1.094
	(0.109)	(0.059)	(0.083)	(0.101)	(0.736)	(0.581)	(0.526)	(0.567)
Distance to other polling	-0.043	-0.089	-0.078	-0.157	-0.601	0.363	-0.307	-0.921
place in boundary/match	(0.110)	(0.059)	(0.082)	(0.093)	(0.736)	(0.591)	(0.530)	(0.568)
Mean dep. var.	2.04	1.01	1.43	1.40	2.27	1.05	1.55	1.55
Residual std. dev. of own distance:								
before controlling for other distance	0.162	0.165	0.162	0.164	0.149	0.147	0.149	0.153
after controlling for other distance	0.121	0.124	0.121	0.108	0.023	0.023	0.023	0.024
N	59,805	45,519	59,805	42,754	133,202	95,642	133,202	98,640
	Panel B. Census Block Turnout							
Distance to own poling place	-0.113	-0.070	-0.075	-0.043	0.046	0.047	-0.058	-0.052
	(0.029)	(0.021)	(0.023)	(0.023)	(0.122)	(0.103)	(0.109)	(0.099)
Distance to other polling	0.004	-0.009	0.002	0.008	0.134	0.080	-0.010	-0.035
place in boundary/match	(0.032)	(0.019)	(0.022)	(0.021)	(0.118)	(0.096)	(0.103)	(0.097)
Mean dep. var.	0.57	0.30	0.41	0.35	0.53	0.28	0.37	0.33
Residual std. dev. of own distance:								
before controlling for other distance	0.167	0.172	0.167	0.170	0.145	0.150	0.145	0.150
after controlling for other distance	0.130	0.128	0.130	0.123	0.029	0.027	0.029	0.031
N	3,333	2,546	3,333	2,370	4,108	2,916	4,108	3,312

Notes: This table reports placebo estimates from regressions that simultaneously control for distance to own polling place and distance to the other polling place in a boundary (columns 1–4) or distance to the polling place of a parcel's/block's matched unit (columns 5–8). Each panel reports two standard deviations of distance to own polling place; namely, the residual standard deviation after controlling for all covariates included in the regression but distance to the other polling place in the boundary/match ("before controlling for other distance"), and the residual standard deviation controlling for all covariates including distance to the other polling place in the boundary/match ("after controlling for other distance").

Table A3: Effects on 2014 MN Parcel Counts of Registered Voters

	Specification:			
	Boundary FEs		Matched Pair FEs	
Dist. to Bdry/Match:	<0.10 mi	<0.05 mi	<0.10 mi	<0.05 mi
	(1)	(2)	(3)	(4)
	Panel A. 2014 Registrants - MN			
Distance to polling place	-0.382	-0.264	-0.479	-0.478
	(0.181)	(0.243)	(0.267)	(0.273)
Mean dep. var.	2.16	2.21	2.36	2.21
N	17,051	9,012	34,562	11,802
	Panel B. 2014 Election-Day Registrants - MN			
Distance to polling place	-0.069	-0.074	-0.110	-0.122
	(0.032)	(0.039)	(0.042)	(0.055)
Mean dep. var.	0.16	0.17	0.19	0.18
N	17,051	9,012	34,562	11,802

Notes: This table reports estimates from regressions of parcel-level counts of registered voters in the 2014 Minnesota sample. The outcomes in panels A and B are, respectively, counts of all registered voters and voters who registered on Election Day.

Table A4: Heterogeneous Effects by Census Characteristics
OLS 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)
				Panel A. By % Minority				
% minority ≤ median	1.88	-0.267	1.04	-0.110	1.42	-0.064	1.33	-0.050
		(0.124)		(0.049)		(0.044)		(0.048)
% minority > median	2.17	-0.384	0.98	-0.333	1.43	-0.269	1.45	-0.193
		(0.100)		(0.066)		(0.048)		(0.055)
F-test (within year)		0.77		8.34		13.03		5.47
p		0.38		0.00		0.00		0.02
F-test (across years)		3.52						
p		0.01						
N	59,805	59,805	45,519	45,519	59,805	59,805	42,754	42,754
				Panel B. By Median HH Income				
Income ≤ median	1.99	-0.296	0.88	-0.268	1.29	-0.207	1.21	-0.179
		(0.090)		(0.059)		(0.055)		(0.067)
Income > median	2.08	-0.353	1.17	-0.182	1.55	-0.134	1.54	-0.089
		(0.127)		(0.053)		(0.041)		(0.042)
F-test (within year)		0.19		1.74		1.68		1.95
p		0.66		0.19		0.19		0.16
F-test (across years)		0.59						
p		0.67						
N	59,805	59,805	45,519	45,519	59,805	59,805	42,754	42,754
				Panel C. By % Units w/o Cars				
% w/o cars ≤ median	1.67	-0.289	0.91	-0.136	1.29	-0.093	1.20	-0.018
		(0.118)		(0.054)		(0.043)		(0.042)
% w/o cars > median	2.33	-0.375	1.08	-0.313	1.54	-0.251	1.52	-0.241
		(0.109)		(0.063)		(0.052)		(0.057)
F-test (within year)		0.38		4.86		6.50		13.82
p		0.54		0.03		0.01		0.00
F-test (across years)		3.93						
p		0.00						
N	59,805	59,805	45,519	45,519	59,805	59,805	42,754	42,754

Notes: This table replicates estimates of heterogeneous effects from Table 5 using boundary fixed effects OLS specifications.

Table A5: Heterogeneous 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)
				Panel A. By % Minority				
% minority ≤ median	2.06	-0.218	1.08	-0.266	1.53	-0.058	1.48	-0.052
		(0.099)		(0.108)		(0.069)		(0.075)
% minority > median	2.44	-0.153	1.02	-0.236	1.57	-0.261	1.61	-0.143
		(0.057)		(0.094)		(0.067)		(0.076)
F-test (within year)		0.38		0.04		5.7		.91
p		0.54		0.84		0.02		0.34
N	133,202	133,202	95,642	95,642	133,202	133,202	98,640	98,640
				Panel B. By Median HH Income				
Income ≤ median	2.22	-0.156	0.94	-0.345	1.42	-0.280	1.37	-0.167
		(0.059)		(0.098)		(0.077)		(0.094)
Income > median	2.32	-0.200	1.18	-0.142	1.67	-0.094	1.70	-0.064
		(0.084)		(0.081)		(0.068)		(0.067)
F-test (within year)		0.21		2.75		3.58		0.94
p		0.64		0.10		0.06		0.33
N	133,202	133,202	95,642	95,642	133,202	133,202	98,640	98,640
				Panel C. By % Units w/o Cars				
% w/o cars ≤ median	1.75	-0.119	0.92	-0.160	1.32	-0.054	1.24	-0.011
		(0.093)		(0.096)		(0.068)		(0.070)
% w/o cars > median	2.66	-0.235	1.14	-0.325	1.73	-0.292	1.74	-0.195
		(0.064)		(0.098)		(0.078)		(0.090)
F-test (within year)		1.10		1.55		5.59		2.81
p		0.29		0.21		0.02		0.09
N	133,202	133,202	95,642	95,642	133,202	133,202	98,640	98,640

Notes: This table replicates estimates of heterogeneous effects from Table 5 using matching specifications.

Table A6: Block-Level 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)
				Panel A. By % Minority				
% minority ≤ median	0.63	-0.126	0.36	-0.036	0.49	-0.043	0.39	-0.026
		(0.053)		(0.029)		(0.023)		(0.020)
% minority > median	0.52	-0.097	0.24	-0.092	0.34	-0.106	0.31	-0.073
		(0.023)		(0.019)		(0.019)		(0.022)
F-test (within year)		0.25		2.90		4.7		2.95
p		0.62		0.09		0.03		0.09
F-test (across years)		2.93						
p		0.02						
N	3,333	3,333	2,546	2,546	3,333	3,333	2,370	2,370
				Panel B. By Median HH Income				
Income ≤ median	0.49	-0.118	0.22	-0.078	0.32	-0.103	0.27	-0.066
		(0.019)		(0.017)		(0.015)		(0.019)
Income > median	0.65	-0.107	0.41	-0.041	0.50	-0.043	0.42	-0.029
		(0.049)		(0.028)		(0.023)		(0.021)
F-test (within year)		0.06		2.16		6.56		2.31
p		0.80		0.14		0.01		0.13
F-test (across years)		1.74						
p		0.14						
N	3,333	3,333	2,546	2,546	3,333	3,333	2,370	2,370
				Panel C. By % Units w/o Cars				
% w/o cars ≤ median	0.64	-0.107	0.36	-0.044	0.50	-0.051	0.41	-0.023
		(0.046)		(0.027)		(0.022)		(0.020)
% w/o cars > median	0.52	-0.119	0.26	-0.085	0.35	-0.099	0.32	-0.078
		(0.022)		(0.019)		(0.018)		(0.020)
F-test (within year)		0.07		2.63		3.55		5.11
p		0.79		0.10		0.06		0.02
F-test (across years)		1.79						
p		0.13						
N	3,333	3,333	2,546	2,546	3,333	3,333	2,370	2,370

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 A7: Block-Level 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)
				Panel A. By % Minority				
% minority ≤ median	0.59	-0.115	0.33	0.023	0.44	0.003	0.38	0.017
		(0.075)		(0.035)		(0.036)		(0.030)
% minority > median	0.49	-0.055	0.23	-0.086	0.32	-0.098	0.29	-0.052
		(0.041)		(0.028)		(0.037)		(0.032)
F-test (within year)		0.45		6.44		4.0		2.34
p		0.50		0.01		0.05		0.13
N	4,108	4,108	2,916	2,916	4,108	4,108	3,312	3,312
				Panel B. By Median HH Income				
Income ≤ median	0.47	-0.031	0.21	-0.046	0.30	-0.055	0.25	-0.012
		(0.035)		(0.024)		(0.028)		(0.025)
Income > median	0.60	-0.136	0.39	-0.007	0.45	-0.041	0.40	-0.021
		(0.070)		(0.052)		(0.047)		(0.034)
F-test (within year)		2.12		0.49		0.07		0.04
p		0.15		0.48		0.79		0.84
N	4,108	4,108	2,916	2,916	4,108	4,108	3,312	3,312
				Panel C. By % Units w/o Cars				
% w/o cars ≤ median	0.60	-0.085	0.33	0.011	0.46	-0.000	0.41	0.020
		(0.067)		(0.035)		(0.041)		(0.030)
% w/o cars > median	0.50	-0.088	0.25	-0.079	0.33	-0.101	0.30	-0.061
		(0.036)		(0.029)		(0.029)		(0.027)
F-test (within year)		0.00		4.26		4.27		4.09
p		0.97		0.04		0.04		0.04
N	4,108	4,108	2,916	2,916	4,108	4,108	3,312	3,312

Notes: This table replicates estimates of heterogeneous effects from Table A6 using matching specifications.

Table A8: Parcel-Level Heterogeneous Turnout Effects by Census Characteristics Controlling for All Interactions Simultaneously

Election:	2012 Presidential (1)	2013 Municipal (2)	2014 Midterm (3)	2016 Primary (4)
		Panel A. Boundary FEs		
Distance to polling place	-0.131	-0.079	-0.029	0.003
	(0.082)	(0.069)	(0.066)	(0.071)
Distance ×	-0.011	-0.186	-0.175	-0.092
1(% minority > median)	(0.067)	(0.090)	(0.053)	(0.058)
Distance ×	-0.038	0.004	-0.000	0.025
1(Income > median)	(0.065)	(0.066)	(0.056)	(0.063)
Distance ×	-0.012	-0.106	-0.104	-0.191
1(% w/o cars > median)	(0.069)	(0.093)	(0.064)	(0.066)
N	59,805	45,519	59,805	42,754
		Panel B. Matched Pair FEs		
Distance to polling place	-0.108	-0.348	-0.078	-0.027
	(0.112)	(0.162)	(0.113)	(0.117)
Distance ×	0.107	0.186	-0.121	-0.016
1(% minority > median)	(0.102)	(0.183)	(0.089)	(0.099)
Distance ×	-0.076	0.190	0.093	0.036
1(Income > median)	(0.088)	(0.147)	(0.109)	(0.112)
Distance ×	-0.180	-0.176	-0.165	-0.176
1(% w/o cars > median)	(0.102)	(0.171)	(0.107)	(0.124)
N	127,342	72,686	118,052	85,108

Notes: This table reports estimates from parcel-level Poisson boundary fixed effects (Panel A) and matching (Panel B) specifications that simultaneously control for interactions between distance to the polling place with dummies for higher-than-median values of census block minority presence, census block group median income, and block group percentage of residential units without cars.

Table A9: Block-Level Heterogeneous Turnout Effects by Census Characteristics Controlling for All Interactions Simultaneously

Election:	2012 Presidential (1)	2013 Municipal (2)	2014 Midterm (3)	2016 Primary (4)
	Panel A. Boundary FEs			
Distance to polling place	-0.127	-0.044	-0.065	-0.025
	(0.044)	(0.029)	(0.025)	(0.023)
Distance ×	0.035	-0.051	-0.050	-0.036
1(% minority > median)	(0.052)	(0.030)	(0.030)	(0.028)
Distance ×	0.011	0.022	0.042	0.017
1(Income > median)	(0.035)	(0.024)	(0.025)	(0.022)
Distance ×	-0.016	-0.010	-0.011	-0.037
1(% w/o cars > median)	(0.035)	(0.025)	(0.026)	(0.026)
N	3,333	2,546	3,333	2,370
	Panel B. Matched Pair FEs			
Distance to polling place	-0.021	0.039	0.042	0.061
	(0.071)	(0.031)	(0.039)	(0.032)
Distance ×	0.061	-0.092	-0.076	-0.047
1(% minority > median)	(0.089)	(0.045)	(0.068)	(0.061)
Distance ×	-0.115	-0.013	-0.034	-0.043
1(Income > median)	(0.058)	(0.042)	(0.047)	(0.038)
Distance ×	-0.072	-0.041	-0.072	-0.069
1(% w/o cars > median)	(0.071)	(0.043)	(0.059)	(0.053)
N	4,108	2,916	4,108	3,312

Notes: This table reports estimates from block-level OLS boundary fixed effects (Panel A) and matching (Panel B) specifications that simultaneously control for interactions between distance to the polling place with dummies for higher-than-median values of census block minority presence, census block group median income, and block group percentage of residential units without cars.

Table A10: Heterogeneous Effects by Voter Party Identification

Election:	2012 Presidential		2013 Municipal		2014 Midterm		2016 Primary	
	Mean	Effect	Mean	Effect	Mean	Effect	Mean	Effect
Party Affiliation/Primary:	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Republican	0.12	-0.224	0.06	-0.167	0.08	0.110	0.22	0.022
		(0.121)		(0.149)		(0.125)		(0.085)
Democratic	1.25	-0.161	0.73	-0.162	0.87	-0.183	1.17	-0.158
		(0.055)		(0.051)		(0.048)		(0.051)
Unaffiliated	0.84	-0.193	0.41	-0.157	0.56	-0.132		
		(0.075)		(0.073)		(0.056)		
F-test (within year)		0.28		0.00		3.23		3.17
p		0.75		1.00		0.04		0.07
F-test (across years)		1.94						
p		0.06						
N	42,754	42,754	28,474	28,474	42,754	42,754	42,754	42,754

Notes: Each cell reports estimates from a separate Poisson, boundary fixed effects 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.

Table A11: Heterogeneous Effects by State

Election:	2012 Presidential		2013 Municipal		2014 Midterm	
	Mean	Effect	Mean	Effect	Mean	Effect
	(1)	(2)	(3)	(4)	(5)	(6)
Massachusetts - β^{MA}	2.22	-0.177	1.20	-0.181	1.51	-0.151
		(0.060)		(0.047)		(0.044)
Minnesota - β^{MN}	1.58	-0.112	0.68	-0.353	1.23	-0.215
		(0.085)		(0.100)		(0.075)
F-test (within year)		0.39		2.52		0.55
p		0.53		0.11		0.46
$(\beta^{MN} - \beta^{MA})_{14} - (\beta^{MN} - \beta^{MA})_{12}$		-0.130				
		(0.086)				
$(\beta^{MN} - \beta^{MA})_{14} - (\beta^{MN} - \beta^{MA})_{13}$		0.108				
		(0.084)				
N	59,805	59,805	45,519	45,519	59,805	59,805

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.

Table A12: Simulated Turnout with Reprecincting and 0-Distance Scenarios

	Actual Dist.	ΔDist. Actual-Simul.	2012 Turnout (%)			2013 Turnout (%)			2014 Turnout (%)			2016 Turnout (%)		
			Actual	Simulated		Actual	Simulated		Actual	Simulated		Actual	Simulated	
				Reprec.	0-Dist.		Reprec.	0-Dist.		Reprec.	0-Dist.		Reprec.	0-Dist.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
						Panel A. Average Census Block								
All census blocks	0.360	0.035	61.2	61.6	65.2	30.0	30.2	32.2	45.0	45.3	47.6	38.4	38.6	40.0
						Panel B. By % Minority								
% minority ≤ median	0.382	0.043	70.1	70.6	74.9	37.6	37.7	39.0	55.0	55.1	56.6	45.7	45.8	46.7
% minority > median	0.342	0.029	54.1	54.4	57.4	24.2	24.5	27.4	37.1	37.4	40.7	32.9	33.2	35.4
Turnout gap (High-Low SES)			16.0	16.3	17.5	13.3	13.2	11.6	17.9	17.7	15.9	12.7	12.6	11.2
						Panel C. By Median HH Income								
Income ≤ median	0.342	0.032	50.2	50.6	54.2	22.2	22.5	24.9	33.9	34.2	37.4	27.9	28.1	30.1
Income > median	0.376	0.039	71.7	72.1	75.7	41.1	41.2	42.6	55.7	55.9	57.3	46.9	47.0	48.0
Turnout gap (High-Low SES)			21.5	21.5	21.4	18.9	18.8	17.7	21.8	21.7	19.9	19.1	19.0	17.9
						Panel D. By % Housing Units w/o Cars								
% w/o cars ≤ median	0.417	0.043	69.5	70.0	74.0	35.2	35.4	37.1	54.6	54.8	56.7	45.0	45.1	45.9
% w/o cars > median	0.309	0.029	53.9	54.3	57.6	25.7	26.0	28.4	36.7	37.0	39.8	33.9	34.1	36.3
Turnout gap (High-Low SES)			15.6	15.7	16.4	9.5	9.4	8.7	17.8	17.8	16.9	11.1	11.0	9.7

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.