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:

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

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

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:

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

where $\beta_i^{\text{0.1-0.2mi}}$, $\beta_i^{\text{0.2-0.3mi}}$, $\beta_i^{\text{0.3-0.5mi}}$, $\beta_i^{\text{0.5-0.75mi}}$, and $\beta_i^{\text{0.75+mi}}$ 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:

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

[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 reprecincting.

5. **Service costs:** the cost of assigning census block j to polling station i is equal to the population-weighted travel distance from block j to station i. That is: $c_{ij} = d_j \times dist(i, j)$, where $dist(i, j)$ denotes the j-to-i distance.

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

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

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

$$\begin{aligned} \text{minimize} \sum_{i \in I} \sum_{j \in J} c_{ij} x_{ij} & & & (5) \\ \text{subject to} \sum_{i \in I} x_{ij} = 1 & & j \in J & \quad (6) \\ \sum_{j \in J} d_J x_{ij} \leq q_i & & i \in I & \quad (7) \\ x_{ij} \in \{0, 1\} & & i \in I, j \in J. & \quad (8) \end{aligned}$$

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.

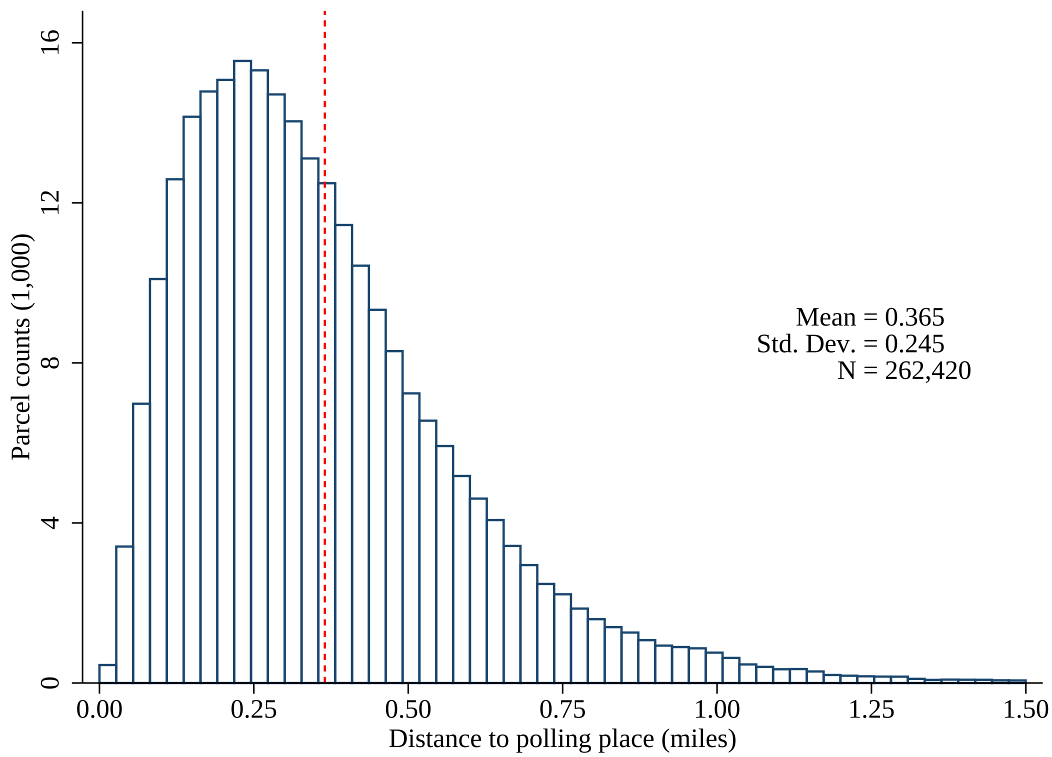
McNulty, John E., Conor M. Dowling, and Margaret H. Ariotti, "Driving Saints to Sin: How Increasing the Difficulty of Voting Dissuades Even the Most Motivated Voters," *Political Analysis*, 2009, *17* (4), 435–455.

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.

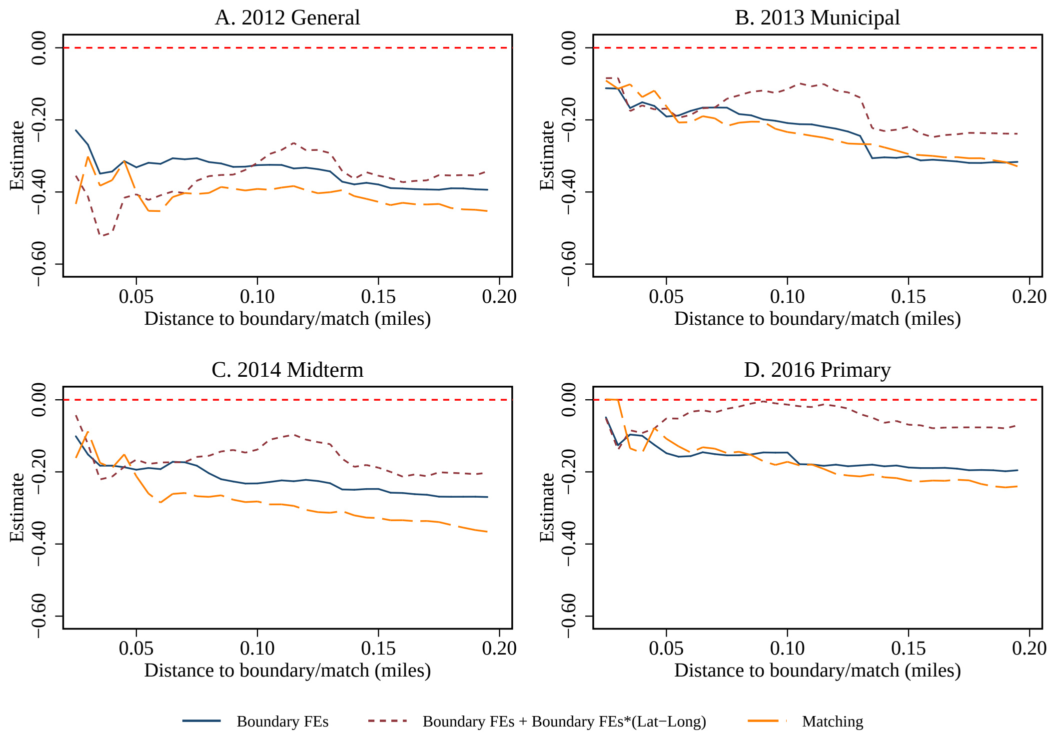
Zandbergen, Paul A., "A comparison of address point, parcel and street geocoding techniques," *Computers, Environment and Urban Systems*, 2008, *32* (3), 214–232.

Figure A1: Distance to the Polling Place



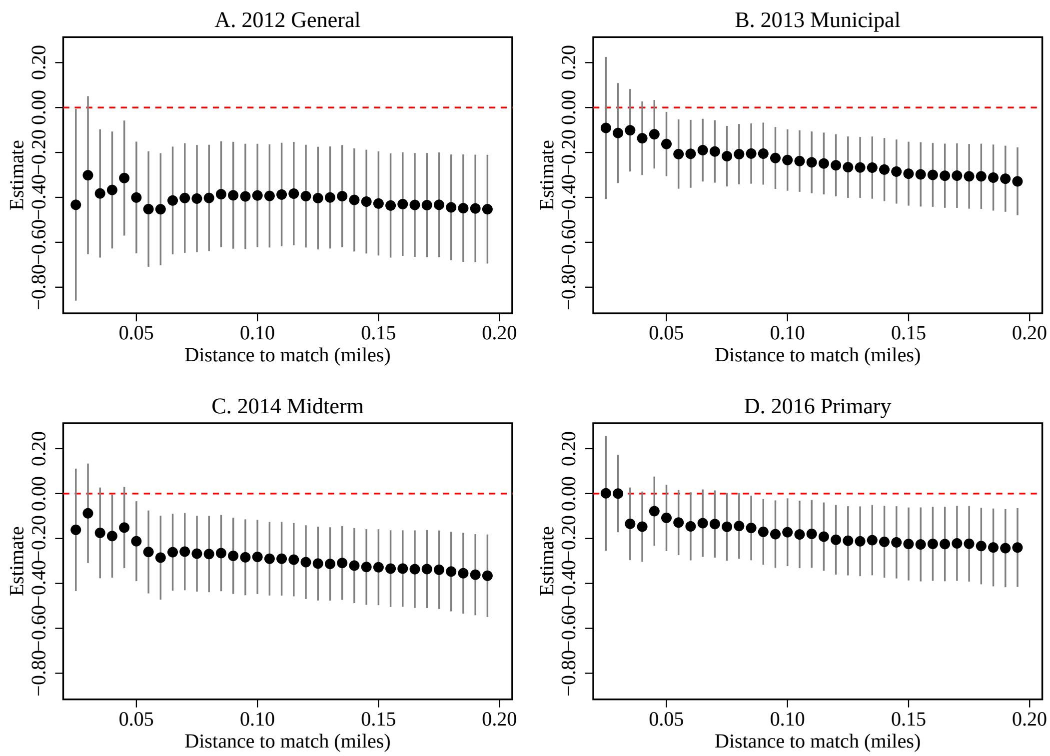
Notes: This histogram plots the distribution of distance to the polling place in the full parcel sample.

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



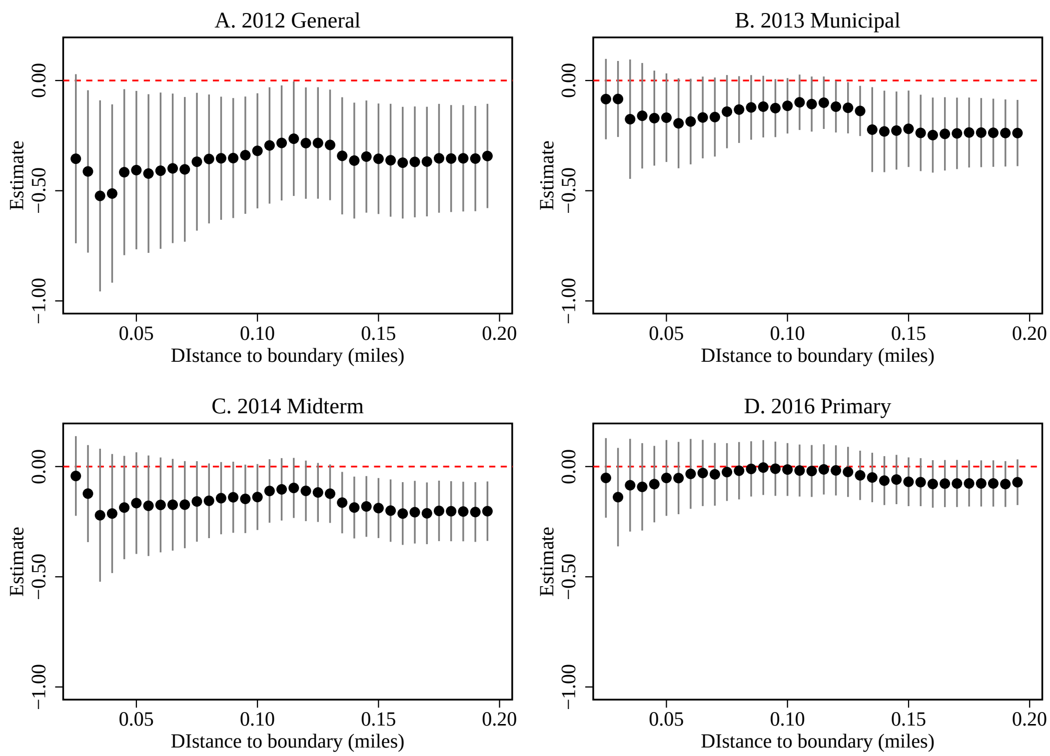
Notes: These figures plot estimated parcel-level treatment effects based on boundary fixed effects, boundary effects with lat-long interactions, and matching specifications across different bandwidths (i.e., distance to the nearest precinct border or distance to the matched unit). Different panels correspond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

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



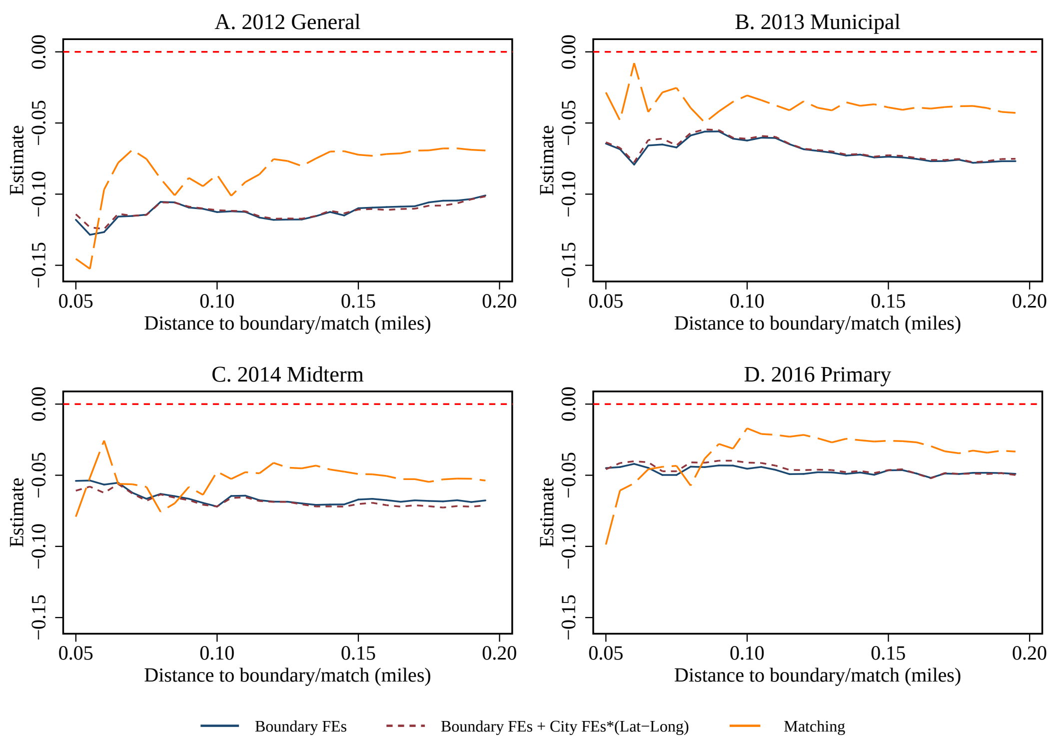
Notes: These figures plot estimated parcel-level treatment effects and 95-percent confidence intervals based on matching specifications across different distances to the matched unit. Each pair of estimate and confidence interval comes from a separate regression.

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



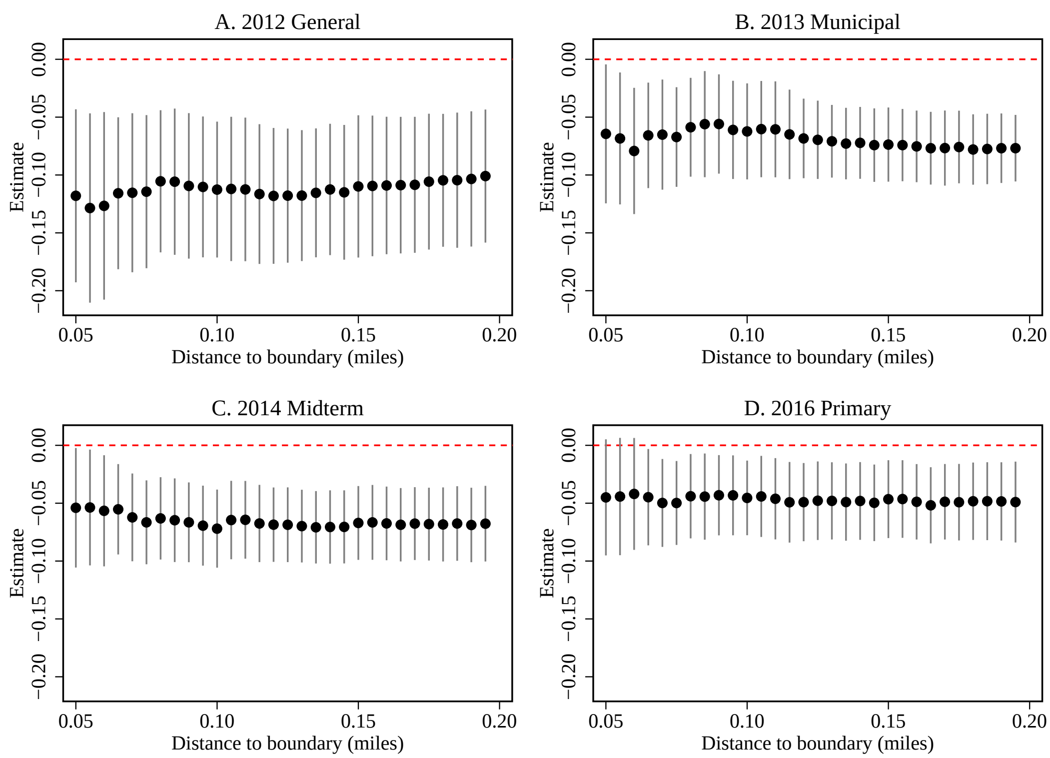
Notes: These figures plot estimated parcel-level treatment effects and 95-percent confidence intervals based on boundary fixed effects specifications with lat-long interactions across different distances to the nearest precinct border. Each pair of estimate and confidence interval comes from a separate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

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



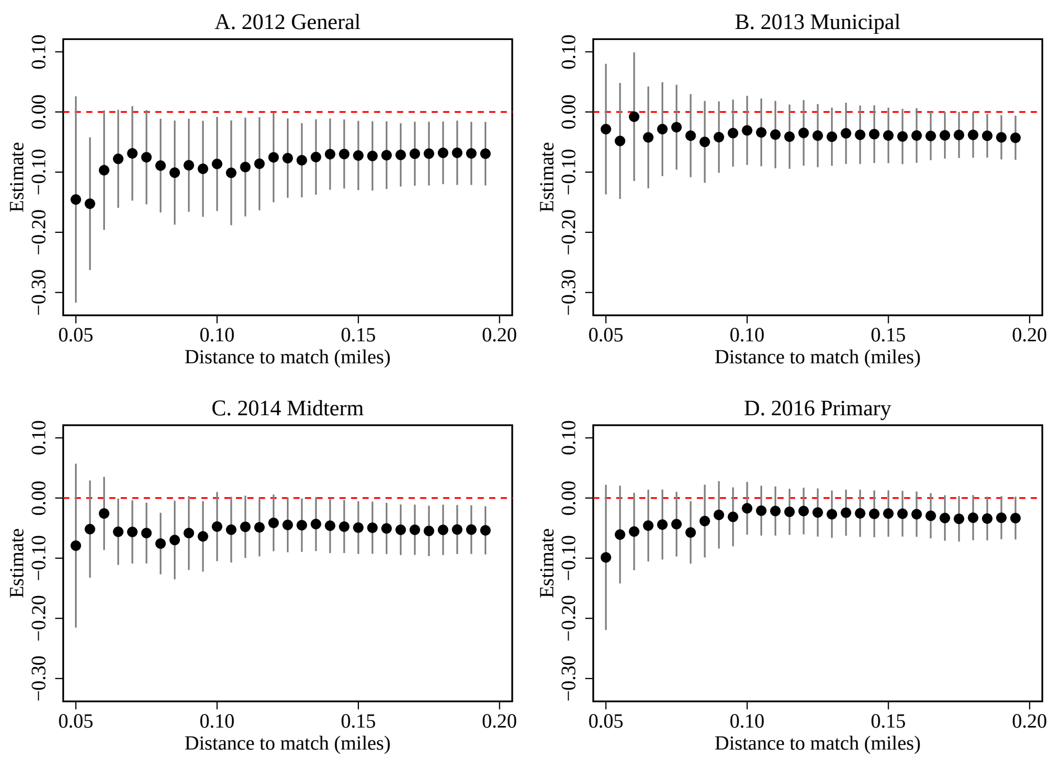
Notes: These figures plot estimated block treatment effects based on boundary fixed effects, boundary fixed effects with lat-long interactions, and matching specifications across different bandwidths (i.e., distance to the nearest precinct border or distance to the matched unit). Different panels correspond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

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



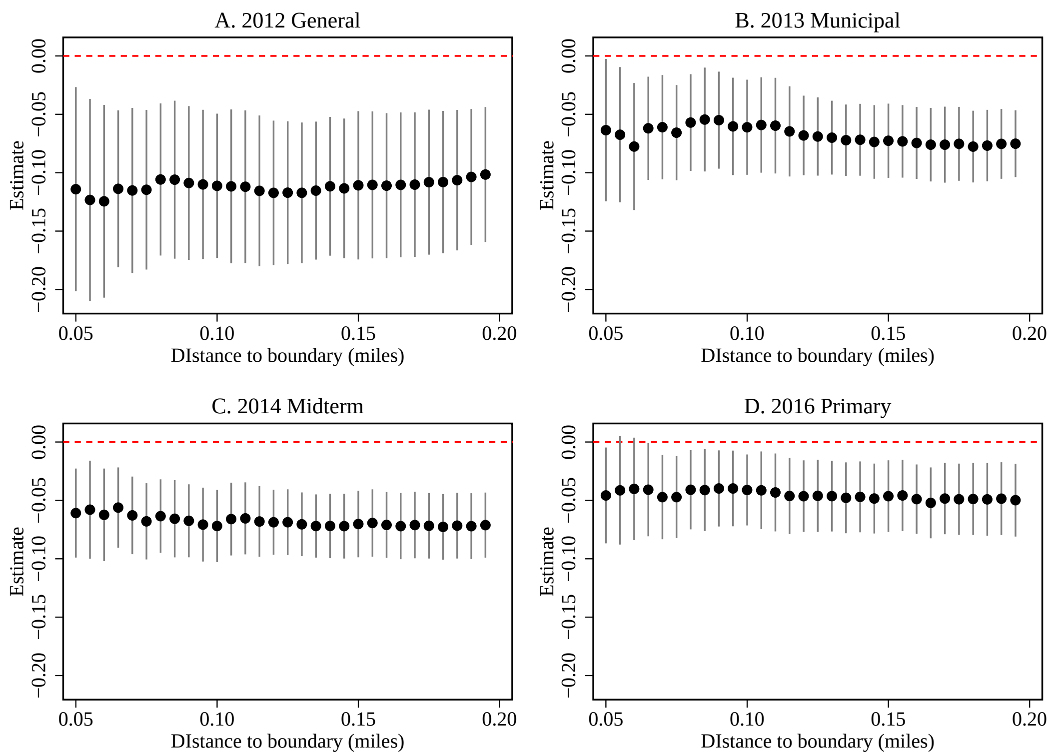
Notes: These figures plot estimated block-level treatment effects and 95-percent confidence intervals based on boundary fixed effects specifications across different distances to the nearest precinct border. Each pair of estimate and confidence interval comes from a separate regression.

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



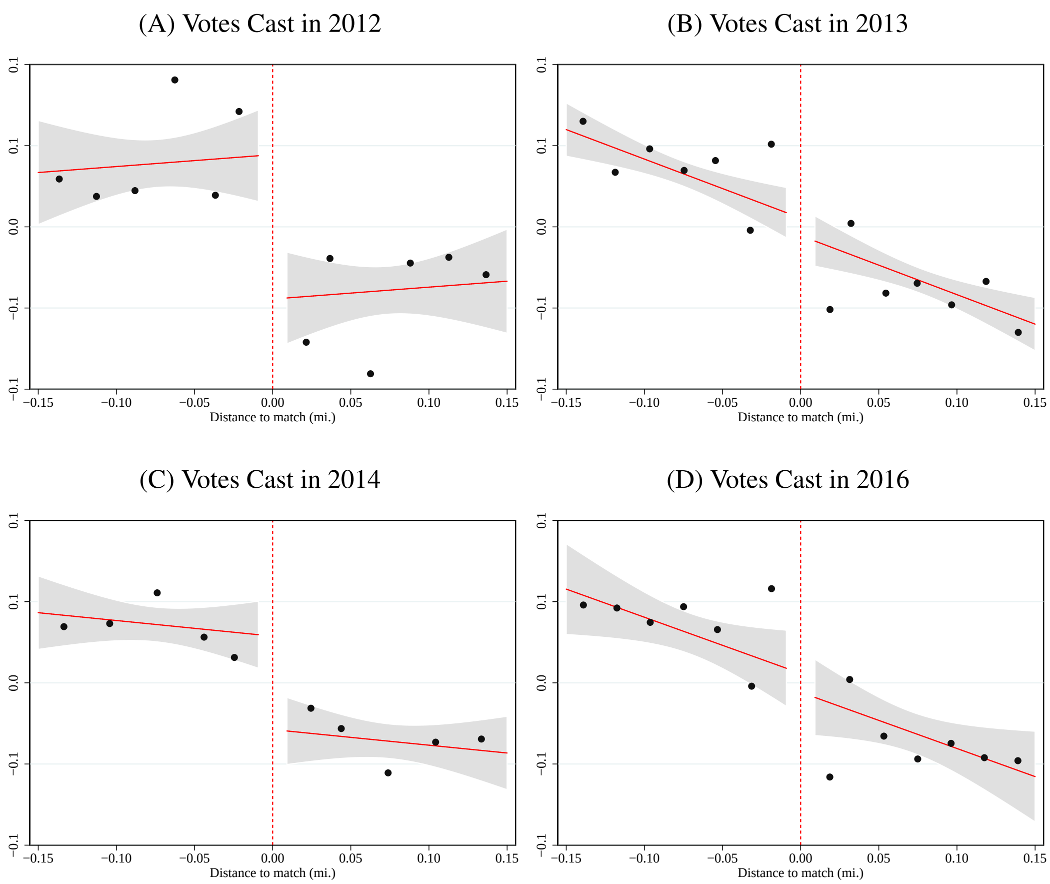
Notes: These figures plot estimated block-level treatment effects and 95-percent confidence intervals based on matching specifications across different distances to matched unit. Each pair of estimate and confidence interval comes from a separate regression.

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



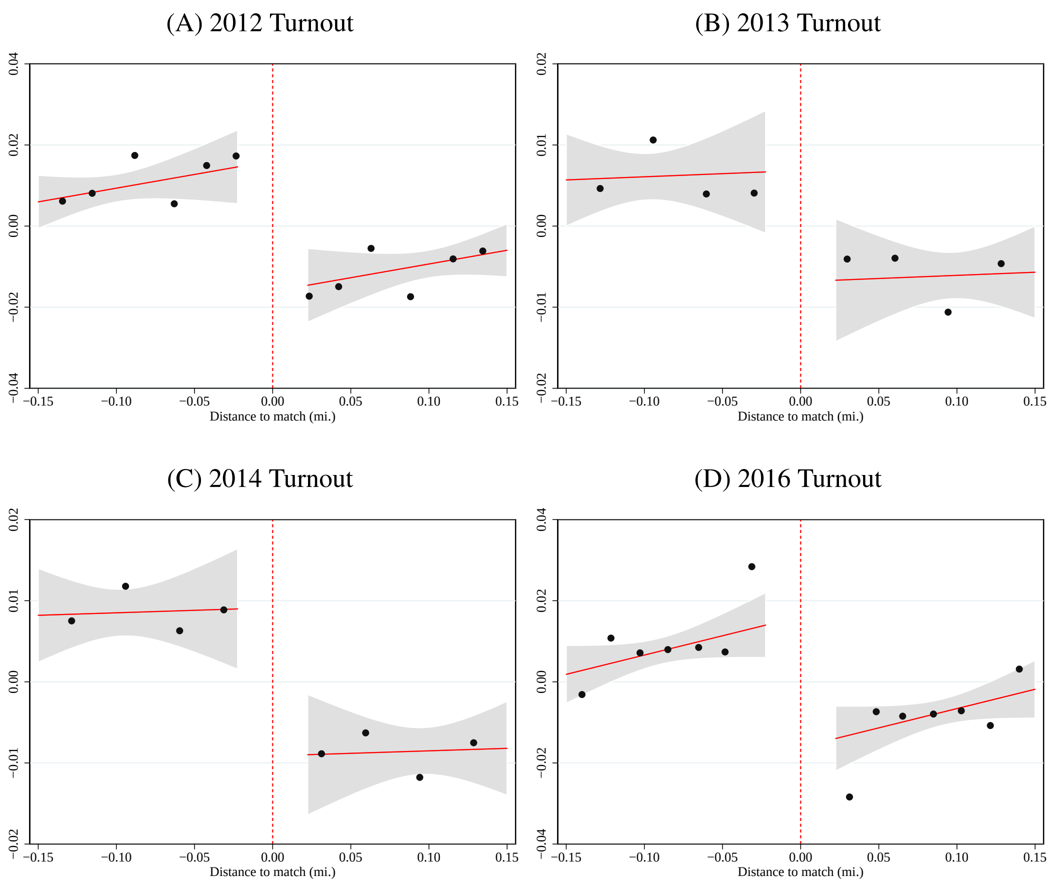
Notes: These figures plot estimated block-level treatment effects and 95-percent confidence intervals based on boundary fixed effects specifications with lat-long interactions across different distances to the nearest precinct border. Each pair of estimate and confidence interval comes from a separate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

Figure A9: Residualized Parcel Outcomes Against Distance to Match



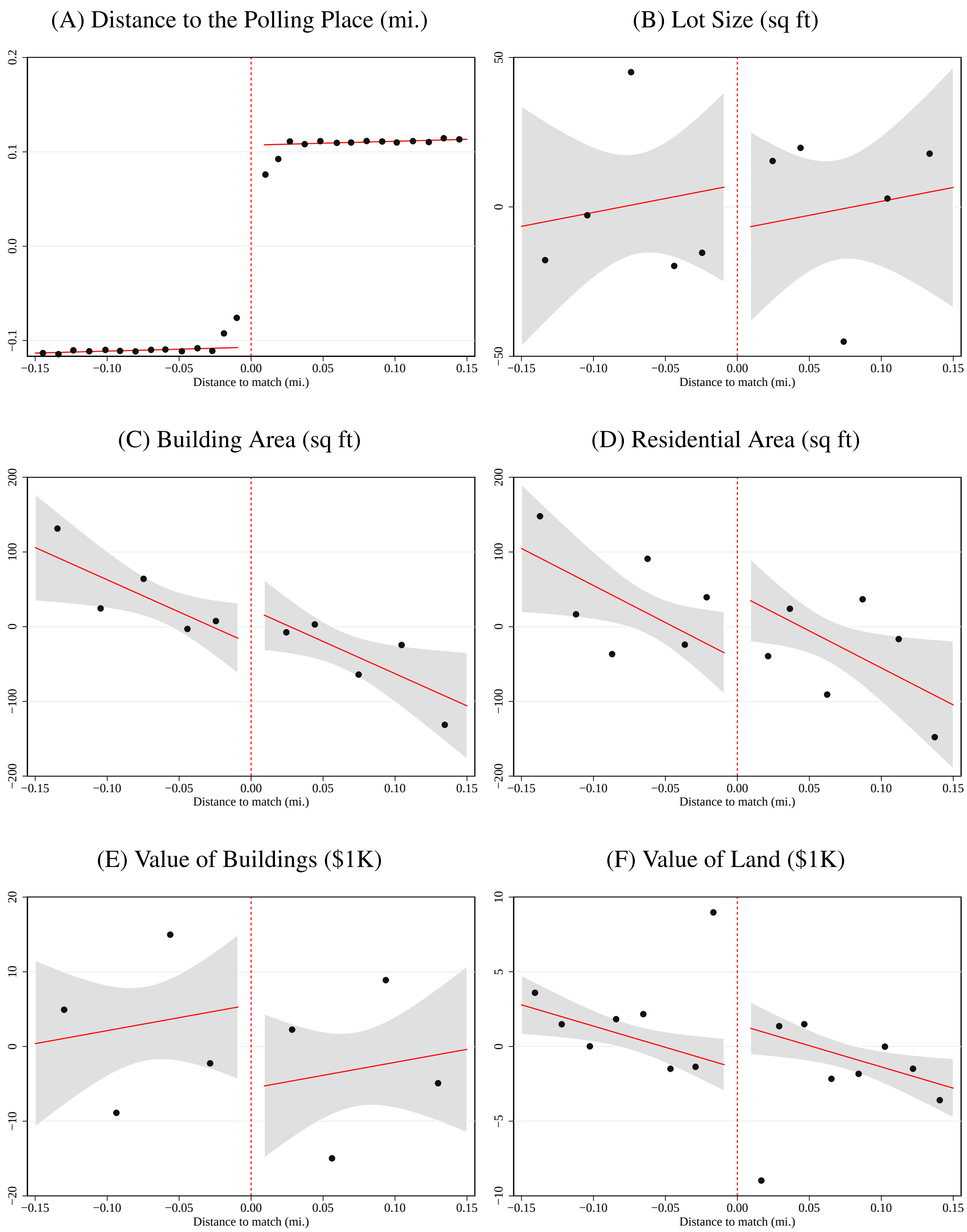
Notes: Using samples of matched parcels, these figures plot votes cast as a function of distance to the matched unit (or the negative thereof). Within each matched pair, the unit that is relatively closer to its polling place is assigned a negative value of distance to the match; the unit that is relatively 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.

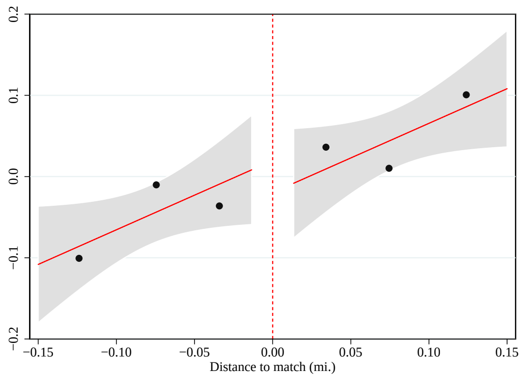
Figure A10: Residualized Census Block Outcomes Against Distance to Match



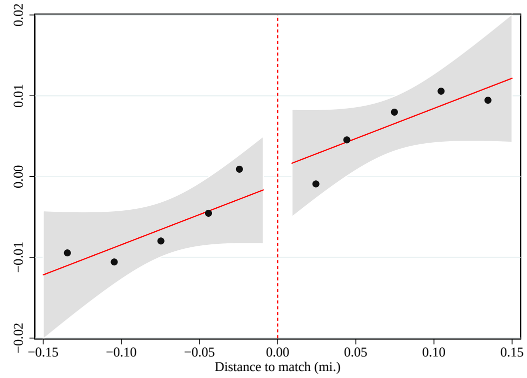
Notes: These figures are constructed in the same way as Figure A9 and plot residualized census block turnout.

Figure A11: Residualized Parcel Covariates Against Distance to Matched Unit

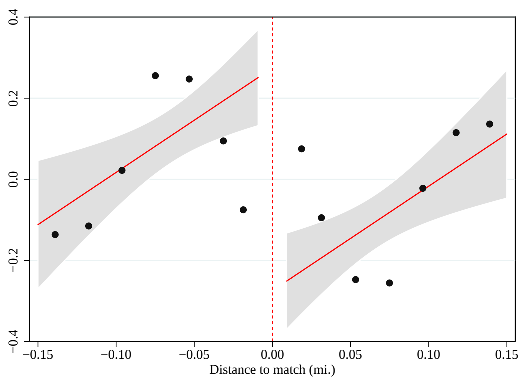




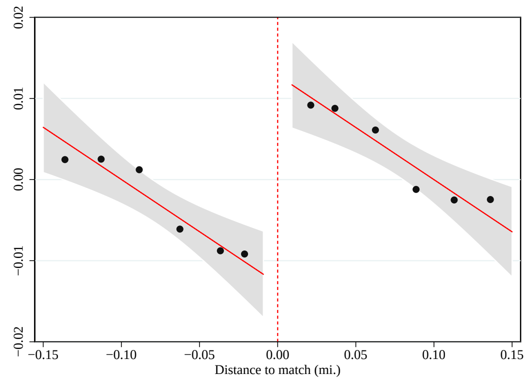
(G) Units



(H) Stories



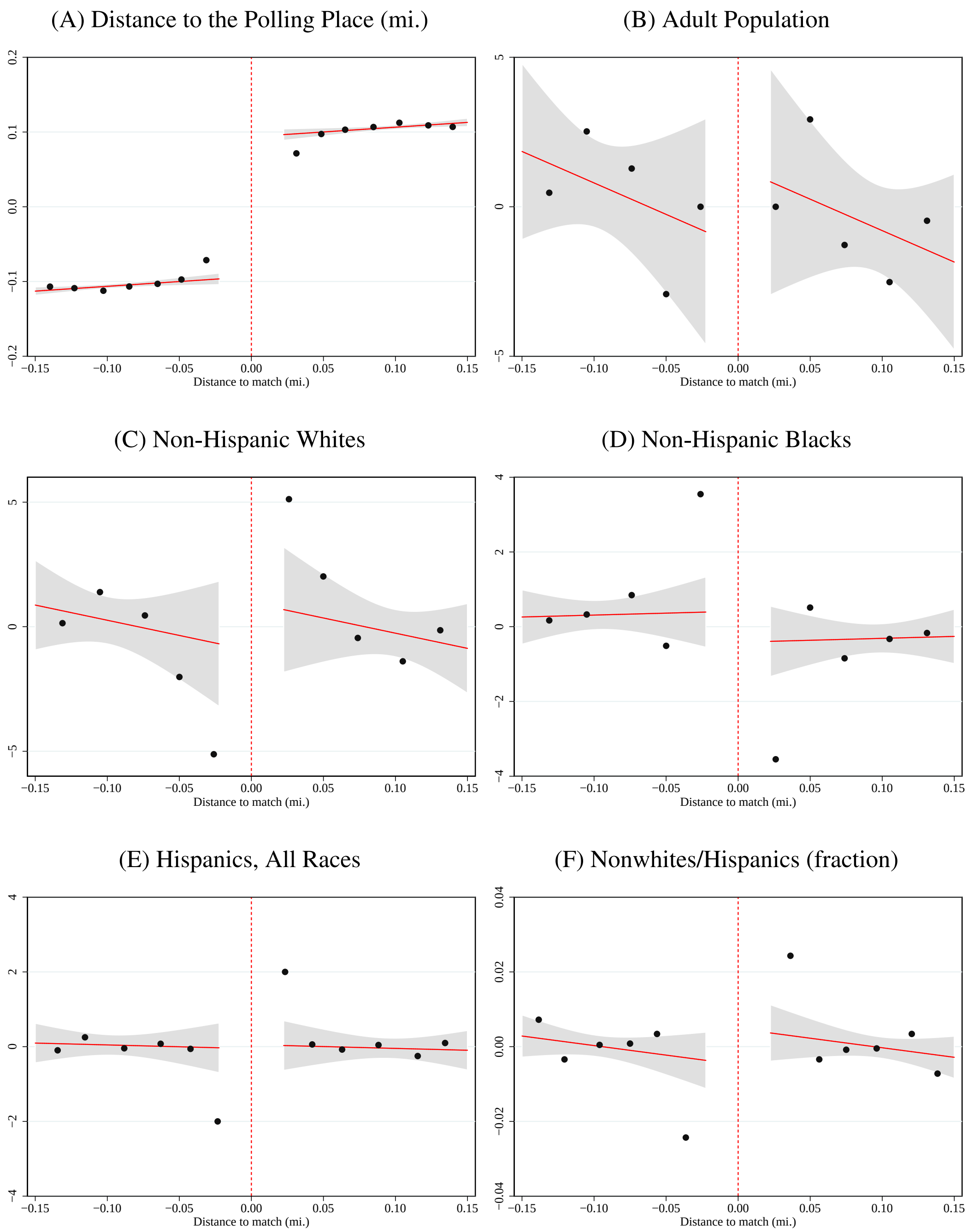
(I) Rooms

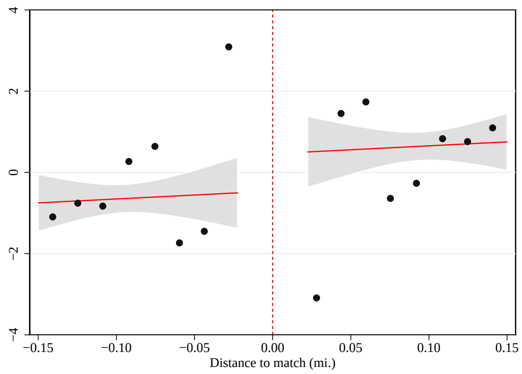


(J) Owner Occupied (fraction)

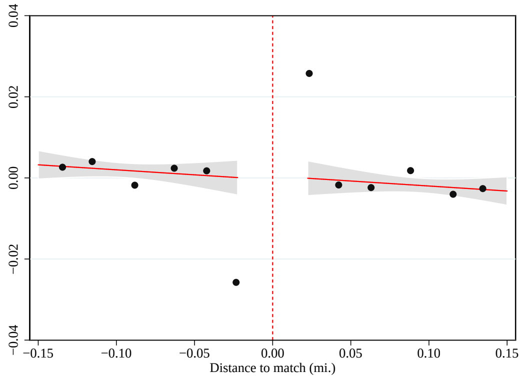
Notes: These figures are constructed in the same way as Figure A9 and plot residualized parcel covariates.

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

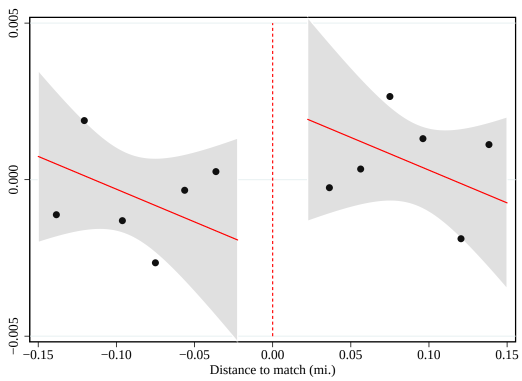




(G) Median HH Income ($1K)



(H) Units w/o Cars (fraction)

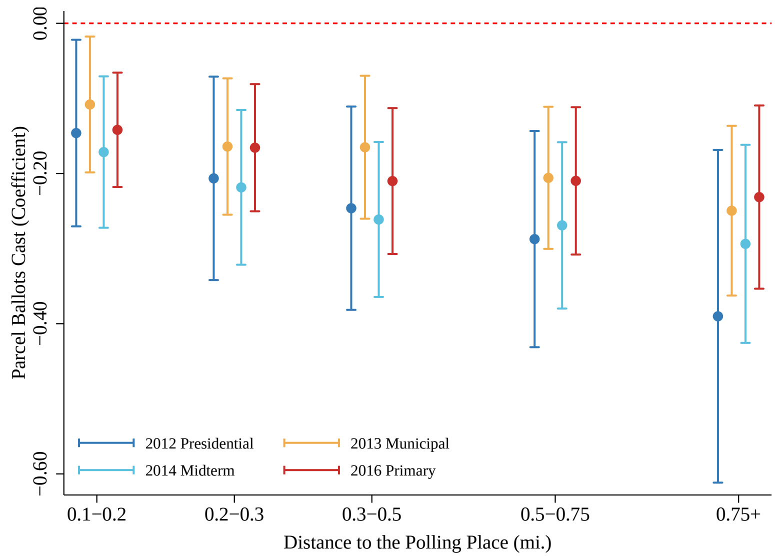


(I) High-School Noncompleters (fraction)

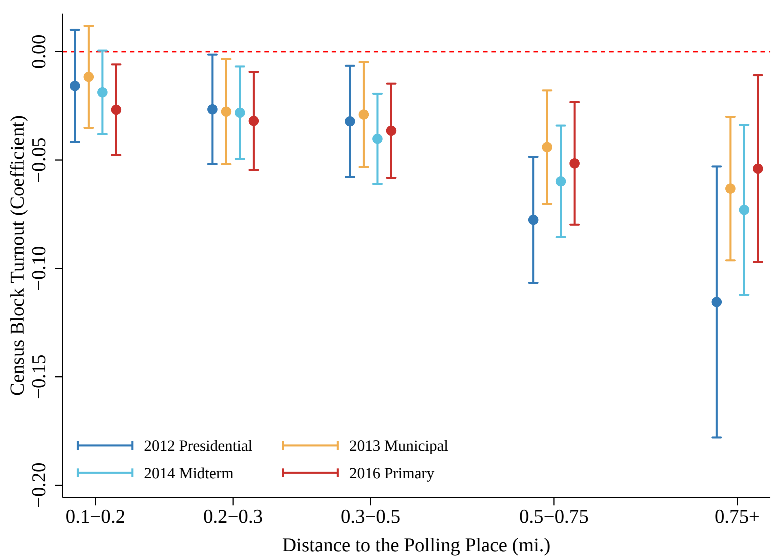
Notes: These figures are constructed in the same way as Figure A9 and plot residualized census block covariates.

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

(A) Non-Linear Effects on Parcel Votes Cast



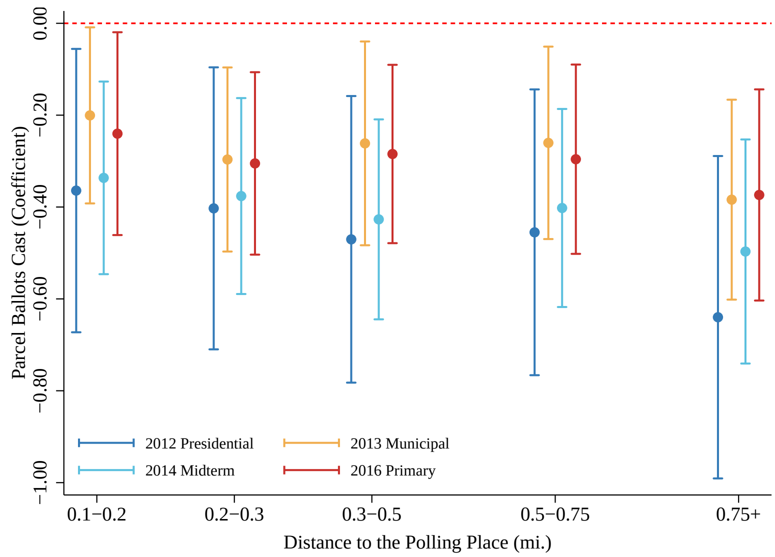
(B) Non-Linear Effects on Census Block Turnout



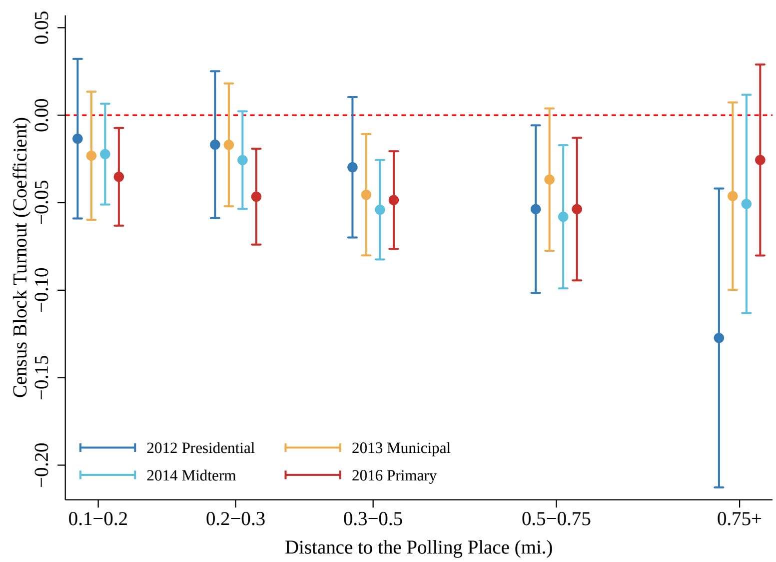
Notes: These figures plot estimates and 95-percent confidence intervals of non-linear treatment effects on the number of ballots cast by parcel residents (Panel A) and block-level voter turnout (Panel B). Each panel reports estimates from four distinct regressions, one for each election. All regressions 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.