Online Appendix
Polling Place Location and the Costs of Voting

Gaurav Bagwe Juan Margitic Allison Stashko

July 31, 2025

Contents

A Creating Border Segments for the Geographic Regression Discontinuity Specification

In this section we explain how to construct border segments so that the necessary monotonicity condition is satisfied for the Fuzzy RD specification. Recall that using a single-dimensional RD in this setting requires a Fuzzy RD approach. By collapsing a two-dimensional treatment assignment vector (latitude and longitude of the voter) to one dimension (distance from voter to border), we necessarily lose some information about treatment assignment. To be concrete, two voters could be equidistant from the border (same value of the running variable) but differ in distance to the polling place (the treatment variable).

A necessary identifying assumption in the Fuzzy RD setting is monotonicity. While monotonicity is typically an assumption, here we can directly observe the treatment for observations in the data, and we can also compute treatment for hypothetical observations that are arbitrarily close to the cutoff. Monotonicity requires all voters sufficiently close to the border to be closer to the polling place on one side of the border (control) than to the polling place on the other side of the border (treatment). In this setting, we can directly observe violations of the monotonicity condition. Figure A.1 shows one such example. Starting from the southwest end of the border segment, voters that are located very close to the border are closer to the left precinct than the right precinct. The opposite is true at the northeast end of the border: voters along the border are closer to the right precinct. To resolve this issue, we divide the border into two segments, one in which all points along the segment are closer to the left precinct and one in which all points along the segment are closer to the right precinct. By splitting this border into two pieces, monotonicity is satisfied in the limit for both segments. Each precinct acts as the control for one segment and the treatment for the other segment. In order to find the appropriate cut point for the border, we first draw a line that connects the two polling places (the solid purple line in Figure A.1). Any point on the perpendicular bisector of that line is equidistant between the two polling places (dashed purple line). Hence, we cut the border at any point where the perpendicular bisector intersects the border. Every point on the border to the left of the cut point is closer to the left precinct, and every point on the border to the right of the cut point is closer to the right precinct.

Figure A.1: Constructing Border Segments to Satisfy the Monotonicity Condition

Note: This figure shows the locations of voters and polling stations along one of the borders that must be split into two segments in order to satisfy the monotonicity condition for the Fuzzy RD specification. This border is located in Lancaster county, PA and is 1.4 miles long. Voters are shaded according to distance to the polling place. The running variable is distance to the border. The polling place on the left is closer to all points on the border to the left of the cut point, indicated by a purple diamond, and the polling place on the right is closer to all points to the right of the cut point. Hence, by splitting the border into two border segments, we have two valid borders that satisfy monotonicity. To determine the cut point, we draw a line connecting the polling places (solid purple line) and then draw the perpendicular bisector of the first line (dashed purple line). The cut point is where the perpendicular bisector intersects with the precinct border.

B RD Robustness Checks

In this section we show that the estimates for the RD specification in Table 1 are not sensitive to the bandwidth selected. Figure B.1 shows RD Estimates for 10 values of bandwidths, ranging from 1/4th to 4 times the MSE-Optimal bandwidth. In Figure E.1, we show that the non-linear RD estimates from Figure 4 are similar when we limit attention to border segments with first-stage effects in an intermediate range. If the first-stage effects are too small, then we suffer from a weak instrument problem. If the first-stage effects are too large, then we might underestimate effects, given the apparent pattern of increasing (negative) marginal effects.

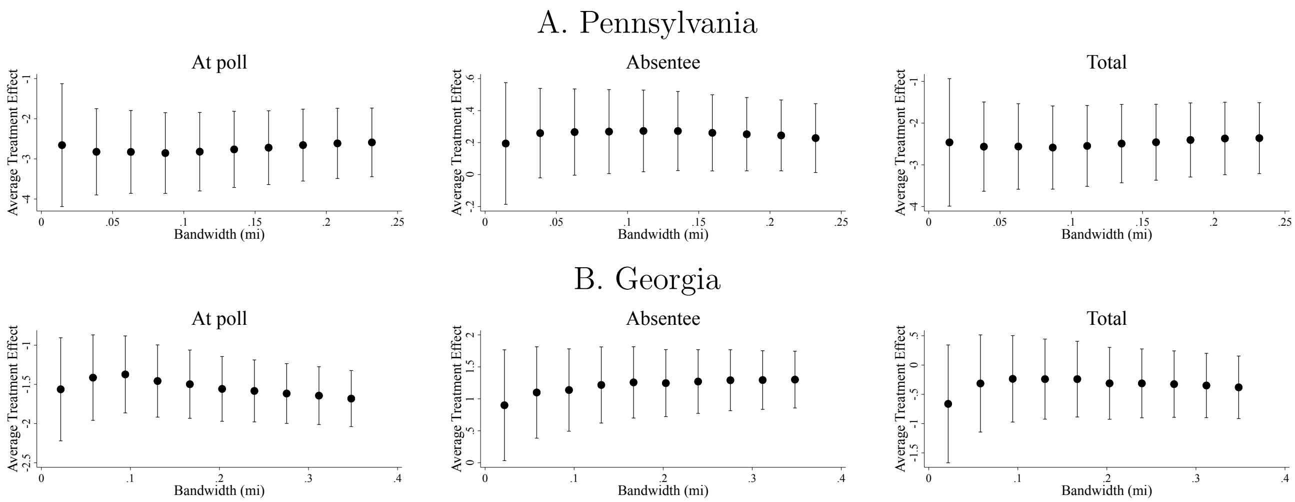

Figure B.1: RD Estimates: Sensitivity to Bandwidth
Note: This figure shows RD estimates from Table 1 of the main text for different values of bandwidths. The minimum bandwidth is 1/4th of the MSE-optimal bandwidth used in Table 1 of the main text and the maximum is 4 times the MSE-optimal bandwidth. Outcome variables indicate whether a registered voter voted in person (At Poll), by mail (Absentee), or by either method (Total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Standard errors allow for clustering at the border segment level.

C Alternative Measures of Distance to Polling Place

In this section we replace travel route distance to the polling place (miles) with Euclidean distance (mi) to the polling place (Table C.1) and with travel time by car (minutes) to the polling place (Table C.2). The estimates indicate that, in Pennsylvania, the likelihood of voting in person falls by 3.8 p.p. per Euclidean mile of distance (similar to travel route distance) and 1.2 p.p. per minute of travel time by car. In Georgia, the likelihood of voting in person falls by 1.8 p.p. per Euclidean mile of distance and 0.8 p.p. per minute of travel time by car. We prefer travel route distance to Euclidean distance since it is more accurate measure of the travel-costs associated with geographic distance to polling place. A limitation of using travel time by car is that we do not know the mode of transportation of all voters.

Table C.1: RD Estimates using Euclidean Distance to Polling Place

	A. Pennsylvania			
	First-stage	Second-stage		
	Euclidean Distance (mi)	At Poll	Absentee	Total
RD Estimate	0.29***	-3.79***	0.36*	-3.43***
	(0.01)	(0.70)	(0.18)	(0.70)
N	2,981,089	2,981,089	2,981,089	2,981,089
Effective N, Left	342,222	564,509	564,509	564,509
Effective N, Right	310,980	500,470	500,470	500,470
Bandwidth	0.04	0.06	0.06	0.06
Outome mean	0.56	57.33	2.28	59.61

	B. Georgia			
	First-stage	Second-stage		
	Euclidean Distance (mi)	At Poll	Absentee	Total
RD Estimate	0.87***	-1.74***	1.41***	-0.33
	(0.02)	(0.33)	(0.43)	(0.50)
N	2,968,151	2,968,151	2,968,151	2,968,151
Effective N, Left	225,954	350,299	350,299	350,299
Effective N, Right	204,683	316,287	316,287	316,287
Bandwidth	0.05	0.08	0.08	0.08
Outome mean	1.22	25.74	28.85	54.59

Note: This table reports RD Estimates for distance to polling place and voting outcomes. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the Euclidean distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for distance and at-poll voting. We use the same bandwidth for all voting outcomes. Standard errors allow for clustering at the border segment level.

Table C.2: RD Estimates using Travel Time by Car to Polling Place

	First-stage	Second-stage		
	Travel Time (min)	At Poll	Absentee	Total
A. Pennsylvania				
RD Estimate	0.93***	-1.17***	0.11**	-1.06***
	(0.03)	(0.21)	(0.05)	(0.21)
N	2,981,089	2,981,089	2,981,089	2,981,089
Effective N, Left	511,440	773,874	773,874	773,874
Effective N, Right	456,411	680,113	680,113	680,113
Bandwidth	0.06	0.09	0.09	0.09
Outome mean	2.55	57.33	2.28	59.61
B. Georgia				
	First-stage	Second-stage		
	Travel Time (min)	At Poll	Absentee	Total
RD Estimate	1.77***	-0.86***	0.73***	-0.13
	(0.06)	(0.13)	(0.16)	(0.19)
N	2,968,151	2,968,151	2,968,151	2,968,151
Effective N, Left	378,308	712,191	712,191	712,191
Effective N, Right	341,221	631,023	631,023	631,023
Bandwidth	0.09	0.16	0.16	0.16
Outome mean	4.15	25.74	28.85	54.59

Note: This table reports RD Estimates for distance to polling place and voting outcomes. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Travel time (min) measures the travel time by car on the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for distance and at-poll voting. We use the same bandwidth for all voting outcomes. Standard errors allow for clustering at the border segment level.

D Additional Election Outcomes

Table D.1 reports RD estimates for 2018 primary elections and the 2016 general election in Pennsylvania.

Table D.1: RD Estimates: Additional Election Outcomes

	A. Pennsylvania, 2018 Primary			
	First-stage	Second-stage		
	Distance (mi)	At Poll	Absentee	Total
RD Estimate	0.37***	-1.87***	0.08	-1.79***
	(0.01)	(0.39)	(0.08)	(0.39)
N	2,981,089	2,978,208	2,978,208	2,978,208
Effective N, Left	618,275	617,788	617,788	617,788
Effective N, Right	546,474	545,996	545,996	545,996
Bandwidth	0.07	0.07	0.07	0.07
Outome mean	0.87	19.12	0.41	19.53
	B. Georgia, 2018 Primary			
	First-stage	Second-stage		
	Distance (mi)	At Poll	Absentee	Total
RD Estimate	1.05***	-0.68***	0.32***	-0.36
	(0.03)	(0.20)	(0.12)	(0.25)
N	2,968,151	2,968,151	2,968,151	2,968,151
Effective N, Left	421,714	421,714	421,714	421,714
Effective N, Right	380,270	380,270	380,270	380,270
Bandwidth	0.10	0.10	0.10	0.10
Outome mean	1.90	12.12	4.34	16.46
	C. Pennsylvania, 2016 General			
	First-stage	Second-stage		
	Distance (mi)	At Poll	Absentee	Total
RD Estimate	0.38***	-2.02***	0.24	-1.78***
	(0.01)	(0.51)	(0.15)	(0.50)
N	2,981,089	2,981,089	2,981,089	2,981,089
Effective N, Left	490,687	490,687	490,687	490,687
Effective N, Right	437,878	437,878	437,878	437,878
Bandwidth	0.05	0.05	0.05	0.05
Outome mean	0.87	67.89	2.95	70.84

Note: Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for distance and at-poll voting. We use the same bandwidth for all voting outcomes. Standard errors allow for clustering at the border segment level.

E Nonlinear Effects

In this section we report estimates shown in Figure 4 in the main text (Tables E.1 and E.2). Next, we show nonlinear estimates using average distance among observations in the treatment precinct to define bins of border segments, rather than the observations in the control precinct (Tables E.3 and E.4). Using these sub-samples, the estimated effect of distance on the likelihood of voting at polls is larger for the bottom quartile than using the sub-samples based on control precincts in Figure 4. However, this can be attributed to the smaller overall variance in distance among the bottom quartile (since distance among the control precinct is less than distance in the treatment precinct, by construction). Hence, while the coefficients differ, the average change in the likelihood of voting at-polls per standard deviation in distance to polling place is similar regardless of how one determines the quartiles of border segments: a 1.4 p.p. vs. 1.5 p.p. reduction in Pennsylvania (Tables E.1 and E.3) and a 1.6 p.p. reduction in Georgia (Tables E.2 and E.4).

In all non-linear estimates, we pool across borders that vary in the first-stage effect. In light of the evidence of nonlinear effects, one concern is that border segments with large first-stage effects overestimate the (negative) treatment effect. A separate concern is that border segments with small first-stage effects could generate bias due to a weak identification problem. To address these concerns, we re-estimate nonlinear effects with the specification of Figure 4 using sub-samples that exclude segments with the largest and smallest first-stage effects. In particular, we consider the sub-sample of border segments within the 10th and 90th percentiles and within the 25th and 75th percentiles (Figure E.1). Last, we report estimates of non-linear effects for the Border FE specification in Tables E.5 and E.6.

Table E.1: RD Estimates: Nonlinear Effects in Pennsylvania

	First Quantile (mean d_i=0.35, st. dev.=0.34)			Second Quantile (mean d_i=0.54, st. dev.=0.54)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-7.68***	0.17	-7.51***	-2.23**	0.66*	-1.57
	(1.45)	(0.27)	(1.47)	(1.11)	(0.35)	(1.05)
N	745,271	745,271	745,271	745,196	745,196	745,196
Outome mean	51.77	1.57	53.34	55.72	1.91	57.64
Bandwidth	0.07	0.07	0.07	0.10	0.10	0.10
	Third Quantile (mean d_i=0.90, st. dev.=0.65)			Fourth Quantile (mean d_i=1.72, st. dev.=1.04)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-1.54	-0.02	-1.57	-0.89	0.36	-0.53
	(1.08)	(0.31)	(1.06)	(0.79)	(0.27)	(0.80)
N	744,653	744,653	744,653	745,969	745,969	745,969
Outome mean	59.76	2.50	62.27	62.04	3.14	65.18
Bandwidth	0.13	0.13	0.13	0.10	0.10	0.10

H_0 Quartile 1 = Quartile 2: At polls, $p-$value< 0.01; Absentee, $p-$value=0.27; Total, $p-$value< 0.01.
H_0 Quartile 1 = Quartile 3: At polls, $p-$value< 0.01; Absentee, $p-$value=0.64; Total, $p-$value< 0.01.
H_0 Quartile 1 = Quartile 4: At polls, $p-$value< 0.01; Absentee, $p-$value=0.62; Total, $p-$value< 0.01.

Note: This table reports RD estimates for four sub-samples. The RD sample is divided into quartiles based on the average distance to the polling place among observations in the control precinct of a border segment. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level.

Table E.2: RD Estimates: Nonlinear Effects in Georgia

	First Quantile (mean d_i=1.01, st. dev.=0.85)			Second Quantile (mean d_i=1.41, st. dev.=0.86)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-2.83***	2.45***	-0.38	-0.98*	1.08	0.10
	(0.62)	(0.77)	(0.95)	(0.51)	(0.71)	(0.81)
N	740,819	740,819	740,819	739,584	739,584	739,584
Outome mean	26.76	25.66	52.42	26.11	27.70	53.81
Bandwidth	0.13	0.13	0.13	0.12	0.12	0.12
	Third Quantile (mean d_i=1.92, st. dev.=1.01)			Fourth Quantile (mean d_i=3.25, st. dev.=1.66)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-1.04**	0.86	-0.18	-1.19***	0.74*	-0.44
	(0.43)	(0.68)	(0.74)	(0.33)	(0.41)	(0.46)
N	744,972	744,972	744,972	742,776	742,776	742,776
Outome mean	25.17	30.26	55.43	24.94	31.77	56.71
Bandwidth	0.16	0.16	0.16	0.10	0.10	0.10

H_0 Quartile 1 = Quartile 2: At polls, $p-$value=0.02; Absentee, $p-$value=0.19; Total, $p-$value=0.70.
H_0 Quartile 1 = Quartile 3: At polls, $p-$value=0.02; Absentee, $p-$value=0.12; Total, $p-$value=0.87.
H_0 Quartile 1 = Quartile 4: At polls, $p-$value=0.02; Absentee, $p-$value=0.05; Total, $p-$value=0.95.

Note: This table reports RD estimates for four sub-samples. The RD sample is divided into quartiles based on the average distance to the polling place among observations in the control precinct of a border segment. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level.

Table E.3: RD Estimates: Nonlinear Effects in Pennsylvania, using Distance in Treated Precincts to Define Bins

	First Quantile (mean d_i=0.30, st. dev.=0.18)			Second Quantile (mean d_i=0.50, st. dev.=0.27)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-16.75***	0.64	-16.11***	-3.83**	0.21	-3.62**
	(4.17)	(0.72)	(4.22)	(1.59)	(0.39)	(1.56)
N	744,463	744,463	744,463	745,412	745,412	745,412
Outome mean	51.88	1.48	53.37	54.88	1.93	56.81
Bandwidth	0.09	0.09	0.09	0.11	0.11	0.11
	Third Quantile (mean d_i=0.90, st. dev.=0.48)			Fourth Quantile (mean d_i=1.80, st. dev.=1.17)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-2.19**	0.24	-1.95**	-0.74**	0.25*	-0.50
	(0.92)	(0.26)	(0.90)	(0.35)	(0.14)	(0.35)
N	745,172	745,172	745,172	746,042	746,042	746,042
Outome mean	60.86	2.65	63.51	61.67	3.06	64.73
Bandwidth	0.17	0.17	0.17	0.13	0.13	0.13

H_0 Quartile 1 = Quartile 2: At polls, $p-$value¡.01; Absentee, $p-$value=0.61; Total, $p-$value=0.01.
H_0 Quartile 1 = Quartile 3: At polls, $p-$value¡.01; Absentee, $p-$value=0.61; Total, $p-$value¡0.01.
H_0 Quartile 1 = Quartile 4: At polls, $p-$value¡.01; Absentee, $p-$value=0.60; Total, $p-$value¡0.01.

Note: This table reports RD estimates for four sub-samples. The RD sample is divided into quartiles based on the average distance to the polling place among observations in the treatment precinct of a border segment (versus the control precinct in Figure 4). Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level.

Table E.4: RD Estimates: Nonlinear Effects in Georgia, Using Distance in Treated Precincts to Define Bins

	First Quantile (mean d_i=0.90, st. dev.=0.49)			Second Quantile (mean d_i=1.42, st. dev.=0.65)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-4.24***	2.42	-1.82	-1.98***	1.78	-0.20
	(1.53)	(1.90)	(2.41)	(0.74)	(1.12)	(1.28)
N	741,079	741,079	741,079	742,098	742,098	742,098
Outome mean	26.59	25.32	51.91	25.99	28.60	54.59
Bandwidth	0.12	0.12	0.12	0.11	0.11	0.11
	Third Quantile (mean d_i=2.00, st. dev.=0.88)			Fourth Quantile (mean d_i=3.28, st. dev.=1.87)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-1.42***	1.23**	-0.18	-0.95***	0.81***	-0.15
	(0.37)	(0.52)	(0.58)	(0.20)	(0.27)	(0.28)
N	742,836	742,836	742,836	742,138	742,138	742,138
Outome mean	25.33	30.11	55.44	25.06	31.37	56.44
Bandwidth	0.23	0.23	0.23	0.13	0.13	0.13

H_0 Quartile 1 = Quartile 2: At polls, $p-$value=0.18; Absentee, $p-$value=0.77; Total, $p-$value=0.55.
H_0 Quartile 1 = Quartile 3: At polls, $p-$value=0.07; Absentee, $p-$value=0.55; Total, $p-$value=0.51.
H_0 Quartile 1 = Quartile 4: At polls, $p-$value=0.03; Absentee, $p-$value=0.40; Total, $p-$value=0.49.

Note: This table reports RD estimates for four sub-samples. The RD sample is divided into quartiles based on the average distance to the polling place among observations in the treatment precinct of a border segment (versus the control precinct in Figure 4). Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level.

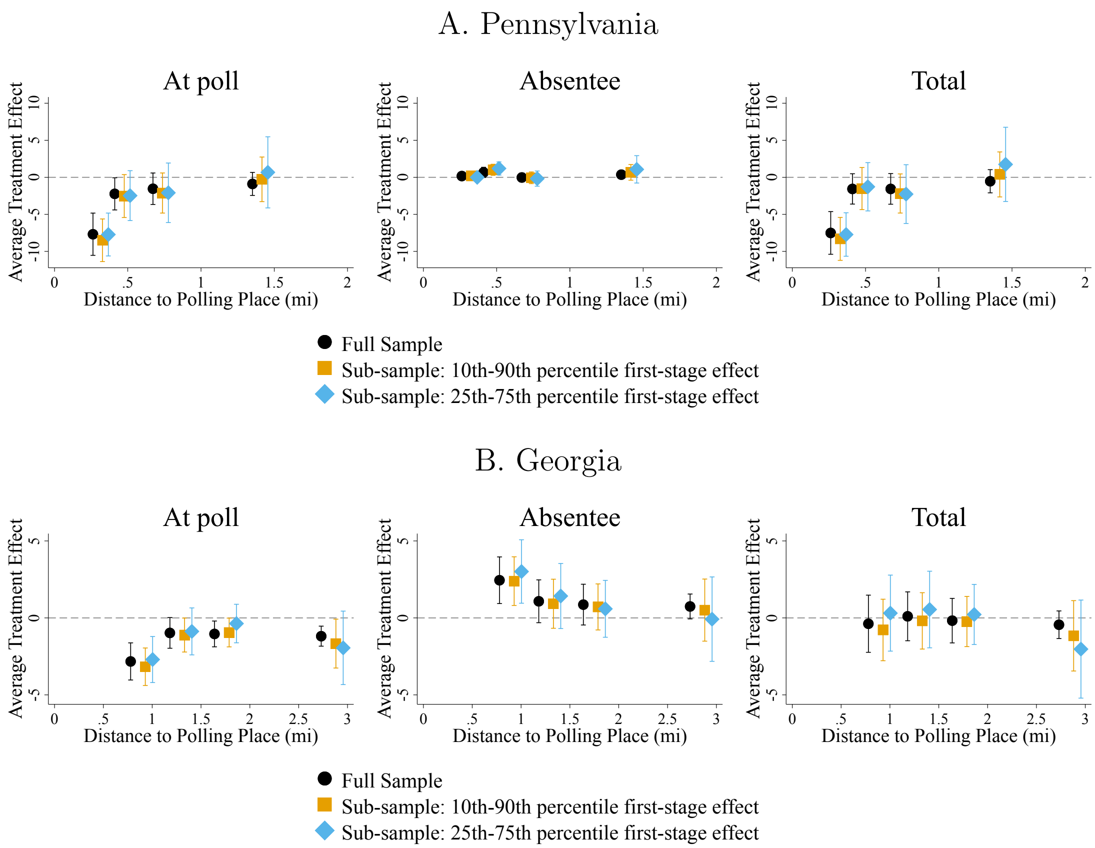

Figure E.1: RD Estimates: Sensitivity to Magnitude of First-Stage Effect

Note: This figure shows RD estimates from Figure 4 of the main text alongside the same estimates for two sub-samples that exclude border segments with the smallest and largest first-stage effects. Outcome variables indicate whether a registered voter voted in person (at Poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level.

Table E.5: Border FE Estimates: Nonlinear Effects in Pennsylvania

	First Quantile (mean d_i=0.32, st. dev.=0.29)			Second Quantile (mean d_i=0.51, st. dev.=0.61)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance to Polling Place (mi)	-1.89***	0.19	-1.69***	-0.95**	0.17**	-0.78*
	(0.60)	(0.15)	(0.56)	(0.48)	(0.08)	(0.44)
N	288,504	288,504	288,504	205,172	205,172	205,172
Outcome mean	50.12	1.39	51.51	51.93	1.54	53.47
R^2	0.11	0.03	0.12	0.12	0.03	0.13
	Third Quantile (mean d_i=0.82, st. dev.=0.64)			Fourth Quantile (mean d_i=1.70, st. dev.=1.09)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance to Polling Place (mi)	-0.14	0.21	0.07	-0.65*	0.03	-0.62*
	(0.59)	(0.16)	(0.55)	(0.35)	(0.10)	(0.35)
N	133,332	133,332	133,332	94,385	94,385	94,385
Outcome mean	55.92	2.12	58.05	59.63	2.69	62.32
R^2	0.12	0.03	0.13	0.11	0.04	0.12

H_0 Quartile 1 = Quartile 2: At polls, $p-$value=0.22; Absentee, $p-$value=0.91; Total, $p-$value=0.20.
H_0 Quartile 1 = Quartile 3: At polls, $p-$value=0.04; Absentee, $p-$value=0.93; Total, $p-$value=0.02.
H_0 Quartile 1 = Quartile 4: At polls, $p-$value=0.08; Absentee, $p-$value=0.36; Total, $p-$value=0.10.

Note: This table reports border fixed effects estimates for for the same sub-samples of border segments used in the nonlinear RD specification (Figure 4 and E.1). Outcome variables indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Each regression includes border segment fixed effects, individual-level covariates (age, sex, registered democrat, registered republican), and block-level covariates (percent White, non-Hispanic, percent Black, non-Hispanic, percent Hispanic, median household income, percent with no high school degree, and percent that walk to work, cars per household). Standard errors allow for clustering at the border segment level.

Table E.6: Border FE Estimates: Nonlinear Effects in Georgia

	First Quantile (mean d_i=0.96, st. dev.=0.76)			Second Quantile (mean d_i=1.46, st. dev.=0.88)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance to Polling Place (mi)	-2.37***	2.22***	-0.15	-1.11***	1.49***	0.38
	(0.48)	(0.47)	(0.54)	(0.34)	(0.34)	(0.33)
N	98,991	98,991	98,991	78,794	78,794	78,794
Outcome mean	24.93	24.02	48.95	25.41	26.73	52.14
R^2	0.05	0.18	0.22	0.05	0.17	0.20
	Third Quantile (mean d_i=1.96, st. dev.=1.08)			Fourth Quantile (mean d_i=3.65, st. dev.=1.90)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance to Polling Place (mi)	-1.00***	1.49***	0.49	-1.00***	1.11***	0.11
	(0.26)	(0.28)	(0.30)	(0.18)	(0.20)	(0.20)
N	72,432	72,432	72,432	84,899	84,899	84,899
Outcome mean	23.49	28.83	52.31	24.56	31.21	55.77
R^2	0.05	0.19	0.20	0.06	0.18	0.20

H_0 Quartile 1 = Quartile 2: At polls, $p-$value=0.03; Absentee, $p-$value=0.21; Total, $p-$value=0.40.
H_0 Quartile 1 = Quartile 3: At polls, $p-$value=0.01; Absentee, $p-$value=0.18; Total, $p-$value=0.30.
H_0 Quartile 1 = Quartile 4: At polls, $p-$value=0.01; Absentee, $p-$value=0.03; Total, $p-$value=0.65.

Note: This table reports border fixed effects estimates for for the same sub-samples of border segments used in the nonlinear RD specification (Figure 4 and E.2). Outcome variables indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Each regression includes border segment fixed effects, individual-level covariates (age, sex, registered democrat, registered republican), and block-level covariates (percent White, non-Hispanic, percent Black, non-Hispanic, percent Hispanic, median household income, percent with no high school degree, and percent that walk to work, cars per household). Standard errors allow for clustering at the border segment level.

F Block-level Border Fixed Effects Regressions

We estimate the effect of distance to polling place on voter registration rates and turnout rates in each state, using Census block level analysis. For the block-level analysis, we assign each block to a unique precinct using the centroid of the block. We measure distance to polling place as the Euclidean distance between the block centroid and the polling place. Each block is also assigned to a unique border if the border is within 0.05 miles of the block centroid. If the block is near multiple borders, we use the modal border assignment among individual registered voters as the block's unique border assignment

In each block we observe the total voting age population (VAP) from the Census, which is a proxy for the population eligible to vote. We measure the percent registered voters as the number of voters observed in the voter registration file, divided by the VAP of the block. We measure at-poll, absentee, and total turnout as the number of votes divided by the VAP. Because VAP comes from 2010 Census data and voter registration files are from 2018, the turnout rates and registration rates are imperfectly measured. The measures are especially noisy for small blocks. Among blocks with fewer than 50 people, the average number of registered voters is 137% and 114% of VAP in Pennsylvania and Georgia, respectively. The percent registered voters in blocks with population above 50 is 84% and 96% on average. To reduce noise in our outcome measures, we only include blocks with population of 50 or higher. In the final sample of blocks, there are approximately 230 people and 170 registered voters per block, on average.

Table F.1: Block-level Border Fixed Effects Regressions

	A. Pennsylvania			
	Percent Registered	At Poll	Absentee	Total
Distance to Polling Place (mi)	-0.14	-0.50	0.09	-0.41
	(0.73)	(0.46)	(0.06)	(0.47)
N	28,509	28,509	28,509	28,509
Outcome mean	86.03	46.57	1.58	48.15
R^2	0.54	0.65	0.46	0.65
	B. Georgia			
	Percent Registered	At Poll	Absentee	Total
Distance to Polling Place (mi)	-0.54	-1.58***	1.52***	-0.06
	(0.66)	(0.21)	(0.30)	(0.40)
N	9,964	9,964	9,964	9,964
Outcome mean	99.75	26.16	29.84	56.00
R^2	0.41	0.56	0.57	0.54

Note: A unit of observation is a Census Block. Distance to polling place is measured as travel-route distance in miles. Percent registered is the percent of the voting age population (VAP) that is registered to vote. Voting outcome variables measure the percent of VAP that voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the Euclidean distance between a block centroid and polling place. Each regression includes border segment fixed effects, individual-level covariates (age, sex, registered democrat, registered republican), and block-level covariates (percent White, non-Hispanic, percent Black, non-Hispanic, percent Hispanic, median household income, percent with no high school degree, and percent that walk to work, cars per household). Standard errors allow for clustering at the border segment level.

G Balance for Fixed Effects Regression Specifications

Table G.1: Balance for Border Fixed Effects Regression Specifications

A. Pennsylvania

	Female	Age	Democrat	Republican	Median HH Income	Urban
Distance to polling place (mi)	-0.0023***	-0.1862*	-0.0025	0.0022	0.0141	-0.0136***
	(0.0008)	(0.1056)	(0.0017)	(0.0017)	(0.0275)	(0.0030)
N	867,036	882,105	883,056	883,056	878,404	867,036
Outcome mean	0.52	49.12	0.61	0.25	5.02	0.97
R^2	0.35	0.15	0.16	0.17	0.90	0.79

	Percent White	Percent Black	Percent Hispanic	Percent with No HS Degree	Percent that Walk to Work	Cars per Household
Distance to polling place (mi)	0.0012	-0.0025	-0.0011	-0.0009	-0.0005	0.0032
	(0.0022)	(0.0020)	(0.0007)	(0.0007)	(0.0005)	(0.0051)
N	867,036	867,036	867,036	878,752	883,044	769,043
Outcome mean	0.67	0.23	0.08	0.14	0.07	1.34
R^2	0.92	0.94	0.85	0.85	0.90	0.93

B. Georgia

	Female	Age	Democrat	Republican	Median HH Income	Urban
Distance to polling place (mi)	-0.0017***	-0.1439*	-0.0021***	-0.0003	0.0327**	-0.0127***
	(0.0006)	(0.0753)	(0.0007)	(0.0009)	(0.0166)	(0.0023)
N	390,194	404,836	405,001	405,001	404,987	390,194
Outcome mean	0.52	46.06	0.09	0.07	5.77	0.85
R^2	0.48	0.10	0.06	0.09	0.87	0.87

	Percent White	Percent Black	Percent Hispanic	Percent with No HS Degree	Percent that Walk to Work	Cars per Household
Distance to polling place (mi)	0.0018	-0.0014	-0.0015	-0.0020**	-0.0007*	0.0124***
	(0.0021)	(0.0020)	(0.0010)	(0.0008)	(0.0003)	(0.0028)
N	390,194	390,194	390,194	405,001	405,001	362,150
Outcome mean	0.56	0.35	0.07	0.14	0.02	1.74
R^2	0.82	0.84	0.63	0.83	0.82	0.87

Note: This table reports coefficients for the travel route distance to polling place (mi). The outcome variables are the individual-level and census block-level covariates listed in each column. Each regression includes border segment fixed effects. Standard errors allow for clustering at border segment level.

Table G.2: Balance for Matched Pair Fixed Effects Regression Specifications

A. Pennsylvania

	Female	Age	Democrat	Republican	Median HH Income	Urban
Distance to polling place (mi)	-0.005***	-0.456***	-0.007**	0.006*	0.030	-0.026***
	(0.001)	(0.142)	(0.004)	(0.003)	(0.032)	(0.005)
N	3,615,954	3,615,954	3,615,954	3,615,954	3,615,954	3,615,954
Outcome mean	0.52	50.95	0.53	0.33	5.51	0.95
R^2	0.22	0.11	0.16	0.17	0.88	0.68

	Percent White	Percent Black	Percent Hispanic	Percent with No HS Degree	Percent that Walk to Work	Cars per Household
Distance to polling place (mi)	-0.000	0.000	-0.001*	0.000	0.000	0.007
	(0.002)	(0.002)	(0.001)	(0.001)	(0.000)	(0.005)
N	3,615,954	3,615,954	3,615,954	3,615,954	3,615,954	3,615,954
Outcome mean	0.78	0.14	0.06	0.12	0.05	1.56
R^2	0.87	0.90	0.78	0.83	0.81	0.89

B. Georgia

	Female	Age	Democrat	Republican	Median HH Income	Urban
Distance to polling place (mi)	-0.0027***	0.0086	-0.0004	-0.0015	0.0287	-0.0174***
	(0.0007)	(0.0885)	(0.0016)	(0.0012)	(0.0180)	(0.0031)
N	4,069,160	4,069,160	4,069,160	4,069,160	4,069,160	4,069,160
Outcome mean	0.52	47.33	0.09	0.08	5.68	0.89
R^2	0.24	0.08	0.08	0.10	0.82	0.75

Note: This table reports coefficients for the travel route distance to polling place (mi). The outcome variables are the individual-level and census block-level covariates listed in each column. Each regression includes fixed effects for matched pairs of voters. Standard errors allow for clustering at precinct and segment level.

H Border FE Estimates for a Sample Matched to Cantoni (2020)

In this section we use the replication data from Cantoni (2020)[1] to reconcile the differences in estimates in his study versus in our Table F.1. We construct a sample of blocks in Georgia and Pennsylvania that are more similar to the urban blocks in the sample of Cantoni (2020), which includes the Boston Massachusetts (MA) area and Minneapolis, Minnesota (MN). Note that Cantoni does not provide individual-level estimates, but reports block-level and parcel-level estimates. A parcel is a unit of land within a city and is smaller than a Census block. We were unable to find state-wide parcel data for Pennsylvania and Georgia. Thus, we compare block-level estimates only. As we do not have the geospatial data used to construct the Cantoni Replication data, we cannot use our RD estimation approach.

To find blocks that are comparable to Boston and Minneapolis, we pool block-level data from PA, GA, MN, and MA and estimate a propensity score for the likelihood of being in the MA or MN sample. We use a logit specification and regress an indicator for the MA or MN sample on distance to polling place, distance squared, and the covariates used by Cantoni (population, income, race, car ownership, and education).[2] Then, we construct a matched sample by selecting the border segments that have the highest average propensity score. In one matched sample, we choose a sample size roughly equivalent to that of Cantoni (2020). In a second matched sample, we choose a sample size four times as large as that of Cantoni (2020). Compared to the statewide PA and GA samples, the Census blocks in the matched samples are higher in population, income, and education. We use turnout rates in general elections of midterm years for the outcome variables (2018 in Georgia and Pennsylvania, 2014 in Massachusetts and Minneapolis).

Table H.1 reports the border fixed effects estimates for full state samples (Columns 1 and 2), for the matched samples (Columns 3 and 4) and for the urban areas in Massachusetts and Minnesota (Column 5). The point estimate from the Massachusetts and Minnesota sample is -5.4 ($SE = 2.6$), as in Table 4, Panel C of Cantoni (2020). For the matched sample of similar size to the Massachusetts and Minnesota sample, the point estimate is large and imprecise (-13.0, SE=10.6). With a matched sample that is four times larger in size, the estimated coefficient is of the same order of magnitude as in Massachusetts and Minnesota

[1] Cantoni, Enrico. 2020. "Replication package for: A Precinct Too Far: Turnout and Voting Costs." American Economic Association [publisher]. Accessed at `https://www.aeaweb.org/journals/dataset?id=10.1257/app.20180306` on 2023-10-23.

[2] Due to the different set of covariates, the estimates in Table H.1 differ from the block-level fixed effects estimates in Table F.1.

(-2.8, SE=1.3), and is larger than the estimates for PA and GA.

Table H.1: Border FE Estimates for Sub-samples that are Observationally Similar to Sample from Cantoni (2020)

	State Samples		Matched Sample (PA and GA)		Cantoni (2020)
	(1) PA	(2) GA	(3) Top 15%	(4) Top 4%	(5) MA and MN
Distance to polling place (mi)	-0.37	-0.59	-2.78**	-13.04	-5.43**
	(0.42)	(0.36)	(1.29)	(10.63)	(2.63)
N	28509	9964	6763	1721	1694
y variable mean	48.151	56.000	44.271	47.292	38.229
R^2	0.640	0.505	0.562	0.471	0.595

Note: The Urban Areas sample is provided by Cantoni (2020) and include data from the Boston, Massachusetts area (MA) and Minneapolis, Minnesota (MN). The Matched Samples include observations near border segments that have the highest average propensity score for the likelihood of being in the MA and MN samples (top 15% of observations and top 4% of observations, with the latter chosen to approximately match the sample size of Column 5). The dependent variable is turnout in the 2018 midterm election for Georgia (GA) and Pennsylvania (PA), and turnout in the 2014 midterm election for MA and MN. All regressions include border fixed effects and the following additional controls: population, median household income, percent non-white, percent with no car, and percent with no high school diploma. Standard errors clustered at the border level are reported in parentheses.

I RD Estimates for Three Large Cities

Table I.1: RD Estimates for Cities

	A. Philadelphia, Pennsylvania			
	First-stage	Second-stage		
	Distance (mi)	At Poll	Absentee	Total
RD Estimate	0.16***	-7.58***	0.49	-7.09***
	(0.01)	(2.72)	(0.40)	(2.73)
N	403,863	403,863	403,863	403,863
Effective N, Left	135,959	135,959	135,959	135,959
Effective N, Right	117,070	117,070	117,070	117,070
Bandwidth	0.05	0.05	0.05	0.05
Outome mean	0.38	53.06	1.06	54.12

	A. Pittsburgh, Pennsylvania			
	First-stage	Second-stage		
	Distance (mi)	At Poll	Absentee	Total
RD Estimate	0.60***	-3.08***	0.31	-2.78***
	(0.14)	(1.04)	(0.26)	(1.00)
N	384,467	384,467	384,467	384,467
Effective N, Left	110,978	110,978	110,978	110,978
Effective N, Right	94,694	94,694	94,694	94,694
Bandwidth	0.08	0.08	0.08	0.08
Outome mean	0.85	54.52	2.60	57.12

	C. Atlanta, Georgia			
	First-stage	Second-stage		
	Distance (mi)	At Poll	Absentee	Total
RD Estimate	0.85***	-2.04***	1.78***	-0.26
	(0.06)	(0.50)	(0.62)	(0.74)
N	1,475,288	1,475,288	1,475,288	1,475,288
Effective N, Left	281,163	281,163	281,163	281,163
Effective N, Right	252,015	252,015	252,015	252,015
Bandwidth	0.12	0.12	0.12	0.12
Outome mean	1.58	26.86	28.73	55.59

Note: This table reports RD estimates for the three largest cities in the sample. The Philadelphia sample includes urban blocks in Philadelphia County, the Pittsburgh sample includes urban blocks in Allegheny county, and the Atlanta sample includes urban blocks in Fulton, DeKalb, Gwinnett, Cobb, and Clayton counties. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for distance and at-poll voting. We use the same bandwidth for all voting outcomes. Standard errors allow for clustering at the border segment level.

J Difference in Differences Estimates

In this section we exploit individual-level variation in distance to polling place across time. We create a panel of voters using voting registration and voter history files from the 2016 and 2018 elections in Pennsylvania. Previous instances of voter registration files (from 2016) were unavailable for Georgia. We estimate two specifications. The first specification estimates the effect of a change in distance to the polling place using within-voter variation:

$$vote_{it} = \beta distance_{it} + \delta_i + \gamma_{ct} + \epsilon_{it}, \tag{J.1}$$

where δ_i are individual fixed effects and γ_{ct} are county-year fixed effects. In a second specification, we disentangle the effects of a change in voter residence versus a change in polling place location:

$$\begin{aligned} vote_{it} =& \beta PL\ moved_{it} + \zeta PL\ moved_{it} \times distance_{it} \\ &+ \mu Voter\ Moved_{it} + \eta Voter\ moved_{it} \times distance_{it} \\ &+ \iota PL\ moved_{it} \times Voter\ moved_{it} + \psi PL\ moved_{it} \times Voter\ moved_{it} \times distance_{it} \\ &+ \delta_i + \gamma_{ct} + \epsilon_{it}, \end{aligned} \tag{J.2}$$

where $PL\ moved_{it}$ is an indicator that equals 1 if voter i in election t is assigned to a polling location different from the one assigned in election $t-1$ and $Voter\ moved_{it}$ is an indicator variable that takes value 1 if voter i in election t has a different home address than during election $t-1$. Because we record some very small differences in locations of polling places, likely due to the geocoding two waves of data separately, we restrict attention to polling place moves that are greater than 0.5 miles. Note that there can only be a change in distance if either the voter or the polling place moves, so we do not identify a coefficient for $distance_{it}$ alone. This specification allows us to identify the effect of distance to polling place separately from the effect of a change in polling place.

In Table J.1 we report estimates for Equation J.1. We estimate precise null effects of a change in distance to polling place on the likelihood of voting in total, at polls, and by absentee ballot. These results are consistent with Clinton et al. (2020) and Yoder (2018). When we separately consider voters who move versus polling places that move, we find a small negative effect of distance to polling place for those who experience a change in polling place but remain in their location and no statistically significant effect for those who moved (Table J.2). The point estimates indicate that if a voter moves, holding distance to polling place constant, then they are more likely to vote, whereas a voter whose polling location place moved, holding distance constant is less likely to vote. The effect of changes in the polling place on turnout is the focus of Clinton et al. (2020) and Yoder (2018). While the size of our difference-in-differences estimates is smaller than those reported in these papers, they are similar in order of magnitude.

Table J.1: Difference in Differences Estimates in Pennsylvania

	At Poll	Absentee	Total
Distance (miles)	0.21***	0.00	0.21***
	(0.06)	(0.01)	(0.06)
N	14673050	14673050	14673050
Outcome mean	63.50	2.68	66.18
R^2	0.75	0.65	0.75

Note: All regressions include individual fixed effects and county-year fixed effects. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Standard errors allow for clustering at the precinct level.

Table J.2: Difference in Differences Estimates in Pennsylvania: Polling Place Versus Voter Location Changes

	At Poll	Absentee	Total
PL Moved	-1.08***	0.14**	-0.95***
	(0.31)	(0.06)	(0.30)
Voter Moved	0.46	0.52***	0.97
	(0.62)	(0.08)	(0.63)
PL Moved × Dist. (mi)	0.25***	-0.03	0.22**
	(0.09)	(0.02)	(0.09)
Voter Moved × Dist. (mi)	-1.28**	0.01	-1.27**
	(0.50)	(0.07)	(0.49)
Voter Moved × PL Moved	-2.94***	-0.04	-2.99***
	(0.70)	(0.11)	(0.70)
Voter Moved × PL Moved × Dist. (mi)	0.45	-0.04	0.41
	(0.53)	(0.08)	(0.52)
N	14673050	14673050	14673050
Outcome mean	63.50	2.68	66.18
R^2	0.75	0.65	0.75

Note: All regressions include individual fixed effects and county-year fixed effects. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Standard errors allow for clustering at the precinct level.

K Heterogeneous Effects

In this section we report the estimated coefficients described in Section 5.3 in the main text. In Table K.1, we report correlations between the covariates of interest and distance to polling place, within the RD sample (columns 1 and 3) and within the effective RD sample (observations within the bandwidth, columns 2 and 4). These correlations help to interpret heterogeneous effects estimates, which are based on sub-samples either defined by individual-level characteristics or block-group characteristics. The goal is to understand how distance to polling place correlates with the selection into each sub-sample. Note that we do not include segment fixed effects for this reason, so correlations reported in Table K.1 do not indicate imbalance within RD segments.

Table K.1: Correlations between Distance to Polling Place and Covariates

	Pennsylvania		Georgia	
Female	-0.01***	-0.01***	-0.00	-0.00
	(0.00)	(0.00)	(0.00)	(0.01)
Age 30-49	-0.02***	-0.02***	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)
Age 50-64	0.03***	0.03***	0.04***	0.07***
	(0.01)	(0.01)	(0.01)	(0.01)
Age 65 up	0.02**	0.03***	-0.03***	0.01
	(0.01)	(0.01)	(0.01)	(0.02)
Democrat	-0.09***	-0.08***	-0.17***	-0.18***
	(0.00)	(0.00)	(0.01)	(0.02)
Republican	0.08***	0.09***	0.08***	0.18***
	(0.01)	(0.01)	(0.01)	(0.02)
Median HH Income (10k)	0.07***	0.06***	0.02***	0.02***
	(0.00)	(0.00)	(0.01)	(0.01)
% No HS Diploma	-0.41***	-0.27***	0.84***	1.41***
	(0.08)	(0.10)	(0.17)	(0.21)
% Walk to work	-0.95***	-0.62***	-3.01***	-3.19***
	(0.08)	(0.07)	(0.42)	(0.51)
White, non-Hispanic			0.27***	0.32***
			(0.02)	(0.03)
Black, non-Hispanic			-0.11***	-0.13***
			(0.02)	(0.03)
Hispanic			0.03	0.01
			(0.02)	(0.03)
Full RD sample	X		X	
Effective RD Sample		X		X
N	2,424,546	1,112,515	2,661,413	648,401
R^2	0.13	0.11	0.04	0.05
Outome mean	0.88	0.69	1.90	1.88

Note: This table reports OLS estimates for the full sample used in RD analysis and for the effective sample of observations that are within the MSE-optimal bandwidths used to estimate voting outcomes in Table 1 (within 0.06 miles in Pennsylvania and 0.09 miles in Georgia). The outcome variable is the travel-route distance between a voter and their polling place. Standard errors allow for clustering at the border segment level.

Table K.2: Heterogeneous Effects: Age

A. Pennsylvania

	Age 18-29			Age 30-49		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-6.70**	0.70**	-6.00**	-1.76**	0.12	-1.64**
	(3.05)	(0.30)	(3.04)	(0.78)	(0.14)	(0.79)
N	446,346	446,346	446,346	971,522	971,522	971,522
Outome mean	33.92	2.97	36.89	50.87	1.04	51.91
Bandwidth	0.09	0.09	0.09	0.09	0.09	0.09
	Age 50-64			Age 65 and up		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-2.64***	0.29	-2.36***	-1.82**	0.36	-1.47*
	(0.72)	(0.18)	(0.71)	(0.92)	(0.35)	(0.82)
N	799,783	799,783	799,783	760,745	760,745	760,745
Outome mean	66.92	1.77	68.69	69.35	4.00	73.35
Bandwidth	0.10	0.10	0.10	0.10	0.10	0.10

H_0 Age 65 and up=Age 18-29: At polls, $p-$value=0.12; Absentee, $p-$value=0.45; Total, $p-$value=0.15.
H_0 Age 65 and up=Age 30-49: At polls, $p-$value=0.96; Absentee, $p-$value=0.53; Total, $p-$value=0.88.
H_0 Age 65 and up=Age 50-64: At polls, $p-$value=0.48; Absentee, $p-$value=0.86; Total, $p-$value=0.41.

B. Georgia

	Age 18-29			Age 30-49		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-0.64	-0.07	-0.70	-1.27***	0.44	-0.82*
	(0.47)	(0.57)	(0.87)	(0.35)	(0.39)	(0.44)
N	674,839	674,839	674,839	1,048,391	1,048,391	1,048,391
Outome mean	19.98	14.86	34.84	28.85	22.87	51.71
Bandwidth	0.19	0.19	0.19	0.16	0.16	0.16
	Age 50-64			Age 65 and up		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-2.13***	1.89***	-0.25	-2.36***	2.99***	0.63
	(0.39)	(0.43)	(0.49)	(0.46)	(0.56)	(0.51)
N	727,922	727,922	727,922	512,119	512,119	512,119
Outome mean	28.72	37.39	66.12	22.98	47.66	70.64
Bandwidth	0.15	0.15	0.15	0.17	0.17	0.17

H_0 Age 65 and up=Age 18-29: At polls, $p-$value=0.01; Absentee, $p-$value<0.01; Total, $p-$value=0.18.
H_0 Age 65 and up=Age 30-49: At polls, $p-$value=0.06; Absentee, $p-$value<0.01; Total, $p-$value=0.03.
H_0 Age 65 and up=Age 50-64: At polls, $p-$value=0.71; Absentee, $p-$value=0.12; Total, $p-$value=0.21.

Note: This table reports RD estimates for sub-samples of border segments. Outcomes are residualized, after removing border segment fixed effects. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. Standard errors allow for clustering at the border segment level. We test if coefficients are equal for pairwise comparisons of sub-samples and report p-values above for each outcome variable. We only report p-values for comparisons to the sub-sample of voters aged 65 and up. p-values are greater than 0.05 for all other pairwise comparisons.

Table K.3: Heterogeneous Effects: Sex

A. Pennsylvania

	Female			Male		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-3.29***	0.42**	-2.87***	-3.09***	0.14	-2.95***
	(0.73)	(0.20)	(0.71)	(0.71)	(0.15)	(0.73)
N	1,114,298	1,114,298	1,114,298	1,318,553	1,318,553	1,318,553
Outome mean	59.81	2.36	62.17	52.71	2.21	54.92
Bandwidth	0.08	0.08	0.08	0.09	0.09	0.09

H_0 Female=Male: At polls, $p-$value=0.84; Absentee, $p-$value=0.26; Total, $p-$value=0.94.

B. Georgia

	Female			Male		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-1.52***	1.48***	-0.04	-1.42***	1.03***	-0.40
	(0.26)	(0.34)	(0.37)	(0.27)	(0.31)	(0.39)
N	1,605,032	1,605,032	1,605,032	1,357,509	1,357,509	1,357,509
Outome mean	26.57	30.68	57.25	24.77	26.72	51.49
Bandwidth	0.16	0.16	0.16	0.16	0.16	0.16

H_0 Female=Male: At polls, $p-$value=0.80; Absentee, $p-$value=0.33; Total, $p-$value=0.51.

Note: This table reports RD estimates for sub-samples of border segments. Outcomes are residualized, after removing border segment fixed effects. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level. We test if coefficients are equal for pairwise comparisons of sub-samples and report p-values above for each outcome variable.

Table K.4: Heterogeneous Effects: Race and Ethnicity

A. Georgia

	White, non-Hispanic			Black, non-Hispanic			Hispanic		
	At Poll	Absentee	Total	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-1.68***	1.56***	-0.13	-1.64***	1.57***	-0.06	-0.33	-0.61	-0.94
	(0.28)	(0.34)	(0.34)	(0.41)	(0.56)	(0.67)	(0.94)	(0.89)	(1.18)
N	1,397,065	1,397,065	1,397,065	1,033,100	1,033,100	1,033,100	100,112	100,112	100,112
Outome mean	29.59	31.29	60.88	22.78	30.17	52.95	25.90	17.18	43.08
Bandwidth	0.16	0.16	0.16	0.14	0.14	0.14	0.17	0.17	0.17

H_0 White=Black: At polls, $p-$value=0.93; Absentee, $p-$value=0.98; Total, $p-$value=0.93.
H_0 White=Hispanic: At polls, $p-$value=0.17; Absentee, $p-$value=0.02; Total, $p-$value=0.51.
H_0 Black=Hispanic: At polls, $p-$value=0.20; Absentee, $p-$value=0.04; Total, $p-$value=0.52.

Note: This table reports RD estimates for sub-samples of border segments. Outcomes are residualized, after removing border segment fixed effects. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level. We test if coefficients are equal for pairwise comparisons of sub-samples and report p-values above for each outcome variable.

Table K.5: Heterogeneous Effects: Party

A. Pennsylvania

	Democrats			Republicans			Independent		
	At Poll	Absentee	Total	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-3.88***	0.33*	-3.55***	-0.99*	0.25	-0.74	-2.96***	0.04	-2.92***
	(0.72)	(0.18)	(0.72)	(0.58)	(0.18)	(0.56)	(1.02)	(0.20)	(1.05)
N	1,550,300	1,550,300	1,550,300	1,005,232	1,005,232	1,005,232	425,557	425,557	425,557
Outome mean	58.40	2.34	60.75	62.67	2.52	65.19	40.78	1.49	42.26
Bandwidth	0.07	0.07	0.07	0.19	0.19	0.19	0.15	0.15	0.15

H_0 Democrat=Republican: At polls, $p-$value<0.01; Absentee, $p-$value=0.73; Total, $p-$value<0.01.
H_0 Democrat=Independent: At polls, $p-$value=0.46; Absentee, $p-$value=0.28; Total, $p-$value=0.62.
H_0 Independent=Republican: At polls, $p-$value=0.09; Absentee, $p-$value=0.45; Total, $p-$value=0.07.

B. Georgia

	Democrats			Republicans			Independent		
	At Poll	Absentee	Total	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-3.18***	2.69***	-0.49	-3.00***	3.19***	0.20	-1.15***	0.87***	-0.28
	(0.86)	(0.91)	(0.62)	(0.68)	(0.72)	(0.33)	(0.23)	(0.28)	(0.33)
N	270,082	270,082	270,082	209,776	209,776	209,776	2,488,293	2,488,293	2,488,293
Outome mean	30.32	66.51	96.83	38.98	58.02	97.00	24.13	22.31	46.44
Bandwidth	0.13	0.13	0.13	0.17	0.17	0.17	0.14	0.14	0.14

H_0 Democrat=Republican: At polls, $p-$value=0.87; Absentee, $p-$value=0.66; Total, $p-$value=0.33.
H_0 Democrat=Independent: At polls, $p-$value=0.02; Absentee, $p-$value=0.06; Total, $p-$value=0.77.
H_0 Independent=Republican: At polls, $p-$value=0.01; Absentee, $p-$value<0.01; Total, $p-$value=0.31.

Note: This table reports RD estimates for sub-samples of border segments. Outcomes are residualized, after removing border segment fixed effects. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level. We test if coefficients are equal for pairwise comparisons of sub-samples and report p-values above for each outcome variable.

Table K.6: Heterogeneous Effects: Median Household Income

A. Pennsylvania

	Below Median HH Income			Above Median HH Income		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-4.30***	0.11	-4.18***	-1.60***	0.40*	-1.21**
	(0.93)	(0.16)	(0.93)	(0.62)	(0.22)	(0.60)
N	1,490,026	1,490,026	1,490,026	1,490,850	1,490,850	1,490,850
Outome mean	50.61	1.51	52.12	64.04	3.05	67.09
Bandwidth	0.07	0.07	0.07	0.08	0.08	0.08

H_0 Below = Above : At polls, $p-$value=0.02; Absentee, $p-$value=0.29; Total, $p-$value=0.01.

B. Georgia

	Below Median HH Income			Above Median HH Income		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-1.48***	1.19***	-0.29	-1.33***	1.20**	-0.13
	(0.32)	(0.38)	(0.47)	(0.36)	(0.54)	(0.54)
N	1,483,652	1,483,652	1,483,652	1,484,445	1,484,445	1,484,445
Outome mean	23.20	25.46	48.66	28.28	32.24	60.53
Bandwidth	0.11	0.11	0.11	0.12	0.12	0.12

H_0 Below = Above : At polls, $p-$value=0.75; Absentee, $p-$value=0.99; Total, $p-$value=0.82.

Note: This table reports RD estimates for sub-samples of border segments. Outcomes are residualized, after removing border segment fixed effects. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level. We test if coefficients are equal for pairwise comparisons of sub-samples and report p-values above for each outcome variable.

Table K.7: Heterogeneous Effects: Percent with No High School Diploma

A. Pennsylvania

	Below Median % with no HS Diploma			Above Median % with no HS Diploma		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-1.89***	0.27	-1.62***	-4.00***	0.28	-3.72***
	(0.62)	(0.20)	(0.62)	(0.86)	(0.19)	(0.85)
N	1,490,468	1,490,468	1,490,468	1,490,621	1,490,621	1,490,621
Outome mean	62.34	3.04	65.37	52.31	1.53	53.84
Bandwidth	0.09	0.09	0.09	0.07	0.07	0.07

H_0 Below = Above : At polls, $p-$value=0.05; Absentee, $p-$value=0.97; Total, $p-$value=0.05.

B. Georgia

	Below Median % with no HS Diploma			Above Median % with no HS Diploma		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-1.40***	1.43***	0.03	-1.50***	1.07***	-0.42
	(0.34)	(0.49)	(0.53)	(0.29)	(0.33)	(0.42)
N	1,484,021	1,484,021	1,484,021	1,484,130	1,484,130	1,484,130
Outome mean	27.68	31.30	58.98	23.80	26.41	50.21
Bandwidth	0.19	0.19	0.19	0.12	0.12	0.12

H_0 Below = Above : At polls, $p-$value=0.83; Absentee, $p-$value=0.55; Total, $p-$value=0.50.

Note: This table reports RD estimates for sub-samples of border segments. Outcomes are residualized, after removing border segment fixed effects. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level. We test if coefficients are equal for pairwise comparisons of sub-samples and report p-values above for each outcome variable.

Table K.8: Heterogeneous Effects: Percent that Walk to Work

A. Pennsylvania

	Below Median Commute by Walking			Above Median Commute by Walking		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-1.92***	0.31	-1.61***	-4.24***	0.22	-4.02***
	(0.58)	(0.19)	(0.56)	(0.96)	(0.20)	(0.96)
N	1,489,869	1,489,869	1,489,869	1,491,220	1,491,220	1,491,220
Outome mean	61.50	2.56	64.06	53.15	2.00	55.15
Bandwidth	0.06	0.06	0.06	0.08	0.08	0.08

H_0 Above = Below : At polls, $p-$value=0.04; Absentee, $p-$value=0.76; Total, $p-$value=0.03.

B. Georgia

	Below Median Commute by Walking			Above Median Commute by Walking		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-1.39***	1.57***	0.17	-1.35***	0.72	-0.63
	(0.33)	(0.45)	(0.48)	(0.38)	(0.49)	(0.59)
N	1,481,883	1,481,883	1,481,883	1,486,268	1,486,268	1,486,268
Outome mean	26.38	29.91	56.28	25.11	27.80	52.91
Bandwidth	0.09	0.09	0.09	0.10	0.10	0.10

H_0 Above = Below : At polls, $p-$value=0.93; Absentee, $p-$value=0.20; Total, $p-$value=0.29.

Note: This table reports RD estimates for sub-samples of border segments. Outcomes are residualized, after removing border segment fixed effects. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level. We test if coefficients are equal for pairwise comparisons of sub-samples and report p-values above for each outcome variable.

L Turnout-Maximizing Polling Places

In this section, we solve an optimal polling place location problem. We assume that the planner's objective is to maximize aggregate turnout. Of course, in practice, officials who are responsible for selecting polling place locations might have other objectives such as representativeness of the electorate that might also factor into the planning problem. However, for the purposes of this exercise, we think that the straightforward objective of turnout-maximization provides a useful benchmark.

A planner chooses where to locate a single polling place within a precinct to maximize aggregate turnout. We assume that the planner knows how voters are distributed across the precinct. Each voter decides whether to vote or abstain from voting. Importantly, the location of the polling place affects voting decisions only through the cost of traveling to the polling place to vote in person.

We model a precinct, A, as a compact two-dimensional space, $A \subset R^2$. The planner chooses coordinates for the polling place location, $(x^p, y^p) \subset A$. There are N eligible voters distributed across precinct A. A voter i is located at $(x_i, y_i) \subset A$.

Voters decide whether to abstain ($v_i = 0$) or vote ($v_i = 1$). To keep the model tractable, we abstract away from the distinction between voting in person or voting by mail. Since the planner's objective is to maximize turnout, the method of voting is a second-order concern. We also assume that voting is a function of Euclidean distance to the polling place, rather than travel route or travel time, since it is difficult to simulate counterfactual travel routes. We assume the a voter's likelihood of voting is takes the following form:

$$Pr_i(v_i = 1) = a_i + c(d_i)$$
$$Pr_i(v_i = 0) = 1 - Pr_i(v_i = 1)$$

where $c(.)$ is the cost of voting as a function of voter i's Euclidean distance to the polling place d_i, a_i denotes the voter-specific likelihood of voting if the distance-related cost of voting were zero.

The planner's optimization problem is to pick a set of geographical coordinates for the polling place location (x^P, y^P) that solve the following:

$$\max_{\{x^p, y^p\} \subset A} \sum_{i=1}^{N} Pr_i(v_i = 1), \tag{L.1}$$

equivalently,

$$\max_{\{x^p, y^p\} \subset A} \sum_{i=1}^{N} a_i + c(d_i(x^P, y^P)), \tag{L.2}$$

where $d_i(x^p, y^p)$ is the Euclidean distance between voter i's location (x_i, y_i) and the new polling location (x^p, y^p). The maximand represents the expected precinct-level voter turnout, which is the sum of individual probabilities of voting for all N individuals in the precinct.

Computational Procedure.

We assume that each voter faces a constant marginal cost of distance to polling place, but allow for heterogeneity across voters. In particular, we use non-linear RD estimates, similar to those in Table 4, to capture the fact that voters who that are located close to a polling place tend to be more sensitive to changes in distance to polling place. We measure distance between a voter and each candidate location using Euclidean distance instead of travel route distance, for tractability. Table L.1 reports RD estimates for the four quartiles of voters, based on distance to polling place in the control precinct.

For each registered voter, we draw a marginal cost of distance from a Normal distribution with mean and variance that match the point estimates and standard errors for the relevant quartile in Table L.1. By allowing for heterogeneity in the marginal cost of voting within each bin, we reflect the relative precision of each point estimate. To summarize, the cost of distance is modeled as follows:

$$c(d_i(x^p, y^p)) = c_i \cdot d_i(x^p, y^p) \text{ with } c_i \sim \mathcal{N}(\beta_d^{q(i)}, \sigma_d^{q(i)}), \tag{L.3}$$

where $\beta_d^{q(i)}$ is the point estimate of the marginal effect of distance on the likelihood of voting for a voter i that is initially in quartile q and $\sigma_d^{q(i)}$ is the variance implied by the standard error of the estimate. The cutoffs that define the four quartiles of Euclidean distance to the polling place are 0.2 miles, 0.3 miles, and 0.5 miles.

To find the turnout maximizing optimal polling locations, we solve a constrained optimization problem for each existing precinct using the standard Nelder-Mead algorithm. The optimal polling location is constrained to fall within an approximately 7-mile box centered around the current polling place location. We do not constrain the optimal polling place location to be within the precinct boundaries. We prefer to ignore this constraint because if, for a fixed set of voters, the optimal location falls outside of the existing precinct borders, then it is a sign that the precinct borders are not optimal.

Table L.1: RD Estimates: Nonlinear Effects in Pennsylvania, using Euclidean Distance to Polling Place

	First Quantile (mean d_i=0.21, st. dev.=0.21)			Second Quantile (mean d_i=0.36, st. dev.=0.36)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	-11.24***	0.19	-11.05***	-4.67***	0.72*	-3.95**
	(2.11)	(0.36)	(2.14)	(1.63)	(0.42)	(1.58)
N	745,220	745,220	745,220	745,197	745,197	745,197
Outome mean	51.41	1.55	52.96	55.70	1.93	57.63
Bandwidth	0.06	0.06	0.06	0.08	0.08	0.08
	Third Quantile (mean d_i=0.56, st. dev.=0.41)			Fourth Quantile (mean d_i=1.12, st. dev.=0.74)		
	At Poll	Absentee	Total	At Poll	Absentee	Total
RD Estimate	0.44	-0.13	0.31	-1.47	0.61*	-0.85
	(1.42)	(0.47)	(1.34)	(0.95)	(0.35)	(0.95)
N	745,140	745,140	745,140	745,532	745,532	745,532
Outome mean	60.34	2.55	62.89	61.84	3.10	64.94
Bandwidth	0.11	0.11	0.11	0.09	0.09	0.09

H_0 Quartile 1 = Quartile 2: At polls, $p-$value= 0.01; Absentee, $p-$value=0.33; Total, $p-$value= 0.01.
H_0 Quartile 1 = Quartile 3: At polls, $p-$value< 0.01; Absentee, $p-$value=0.59; Total, $p-$value< 0.01.
H_0 Quartile 1 = Quartile 4: At polls, $p-$value< 0.01; Absentee, $p-$value=0.40; Total, $p-$value< 0.01.

Note: This table reports RD estimates for four sub-samples. The RD sample is divided into quartiles based on the average distance to the polling place among observations in the control precinct of a border segment. Distance to polling place is the Euclidean distance between a voter and their polling place, measured in miles. Voting outcomes indicate whether a registered voter voted in person (at poll), by mail (absentee), or by either method (total). Outcomes are scaled so that coefficients are measured in percentage points. Distance (mi) measures the travel route distance between a voter and their assigned polling place. Bandwidths are MSE-optimal for at-poll voting in each bin. We use the same bandwidth for all voting outcomes within a bin. Standard errors allow for clustering at the border segment level.

In the second counterfactual exercise, we use the same model of voting to find the optimal building locations. We compute predicted aggregate turnout for each candidate building in a precinct and select the building with the highest turnout rate. Finally, in the third counterfactual exercise we select the pair of public buildings in each precinct that maximize the precinct's turnout,simulating the effect of doubling of the number of polling places per precinct. For precincts that have two or fewer candidate buildings, we compute the turnout for the existing candidate buildings.

In all exercises, we exclude 24 precincts (0.3% of all precincts) due to likely geocoding errors: 20 precincts had an average distance to existing polling place greater than 10 miles, and 4 precincts with small population sizes had one or zero registered voters.