Polling Place Location and the Costs of Voting *

Gaurav Bagwe† Juan Margitic † Allison Stashko‡

July 5, 2022

Abstract

We study how distance to one's polling place affects the likelihood of voting, either in-person or by mail. We use a border discontinuity design, with data from over 15 million voters in Pennsylvania and Georgia. The average effect of distance to the polling place on turnout is small, in part because voters substitute to mail-in voting. A one mile increase in distance to polling place decreases the likelihood of voting in a general election by up to 0.99 percentage points. The effect is larger in areas with a heavy reliance on public transportation and in low income areas. Using these estimates, we identify the turnout-maximizing locations of polling places and compute gains to turnout.

JEL Classification: D72, H70, K16
Keywords: Polling Places, Voting Precincts, Costs of Voting

*We would like to thank Laurent Bouton, Garance Genicot, Arik Levinson, Martin Ravallion, Carolina Concha Arriagada, J.J. Naddeo, Arturo Romero Yáñez, and Daniel Valderrama. We also thank participants at the American Political Science Association Annual Meeting, Southern Economic Association Conference, and the University of Utah Political Science Department. Thank you to Alvin Tsang and Sophia Bagley for their excellent research assistance.

†Bagwe and Margitic were doctoral candidates at Georgetown University while working on this project.
‡David Eccles School of Business, University of Utah, e-mail: allison.stashko@eccles.utah.edu.

1 Introduction

The 2020 U.S. elections have intensified debates and policy action on election administration. In 2021 alone, 36 states enacted a total of 96 new election laws.[1] Many of these laws shift the costs and benefits of voting in person or voting by mail. While some states made mail-in ballots available to all voters, others introduced new requirements for voting by mail. Some states expanded the number of early voting days, while others reduced the number of early voting days or removed polling places altogether. It is important to know how voters respond to changes in the costs of voting. This is especially true in the U.S. context, where voter participation is low and election law has become increasingly partisan in nature (Bentele and O'Brien, 2013; Burden, 2017; Hasen, 2002).

State and local government decisions regarding mail-in voting (Meredith and Endter 2016; Meredith and Malhotra 2011; Lockhart et al. 2020; Thompson et al. 2020), early voting, (Kaplan and Yuan 2020), and at-poll voting requirements (Highton 2017; Cantoni and Pons 2022) can each potentially have important consequences for voter participation. Even small changes to the convenience or cost of voting can determine whether or not someone votes, especially in large elections (Gomez et al. 2007; Braconnier et al. 2017).

This paper focuses on a particular cost of voting: the distance from a voter's home to their polling place. Distance to the polling place is an important determinant of voting behavior to study for three reasons. First, this cost of voting cannot be eliminated, unless we remove polling places altogether. The alternative to voting at a polling place is to vote by mail, access to which varies from state to state. Three-quarters of voters chose to vote at polls on election day in the 2018.[2] Even during the pandemic of the 2020 election cycle, 54% of voters cast their ballot in person.[3] Second, the distance to polling place is inherently uneven among the population of eligible voters. It is important to understand if certain populations are systematically disadvantaged by higher costs of voting, in which case politicians may ignore their interests (Avery 2015; Martin 2003). Given the history of voter suppression in the United States, it is especially important to understand if the distance to polling place varies by race of eligible voters. Third, the polling place location is a policy choice, even though it may not be typically thought of as such. State and local lawmakers determine how to divide a state into voting precincts and where to locate a polling places within precincts. A

[1] Brennan Center for Justice. Voting Laws Round-up: December 2021. Available at: https://www.brennancenter.org/our-work/research-reports/voting-laws-roundup-december-2021

[2] Election Administration and Voting Survey: 2018 Comprehensive Report. https://www.eac.gov/sites/default/files/eac_assets/1/6/2018_EAVS_Report.pdf Retrieved November 11, 2020.

[3] Pew Research Center. The Voting Experience in 2020. November 20, 2020. Available at: https://www.pewresearch.org/politics/2020/11/20/the-voting-experience-in-2020/

precise understanding of how polling place locations affect voter participation is the first and most crucial step in determining the optimal allocation of polling places.

Existing estimates of the effect of distance to polling place on voting vary widely. Here we focus on four recent studies that use causal identification strategies.[4] Most recently, Cantoni (2020) uses a border discontinuity approach, which exploits the fact that residents on either side of a voting precinct border are plausibly similar, but differ in their polling place. Using data from nine urban municipalities in Massachusetts and Minnesota, Cantoni finds large effects: a mile increase in distance to polling place reduces turnout by 4 to 12 percentage points. These findings imply that moving a polling place one mile closer to a voter would cause an increase in turnout that is on par with some of the most successful turnout mobilization tactics (Gerber et al. 2017, Green et al. 2013, Enos and Fowler 2018). At the other extreme, Clinton et al. (2020) and Yoder (2018) find null effects of distance to polling places on turnout. Both studies use panels of voters from North Carolina to study how voting behavior changes as polling places change over time. Finally, Brady and McNulty (2011) study precinct consolidations in Los Angeles County using matching methods. They find a small negative effect: a mile increase in distance to the polling place reduces turnout by 1 percentage point.[5] Several studies find an additional search cost associated with a change in polling place location (Amos et al. 2017, Brady and McNulty 2011, Clinton et al. 2020, Yoder 2018).

The varied findings in the existing literature underscore the need for a more comprehensive study of polling places in the United States and their effect on turnout. One hypothesis is that the range of results reflect heterogeneous effects, given that each study uses data from a different location. Moreover, we know that there are large location and election-specific effects on voting behavior (Cantoni and Pons 2022). However, there is also no overlap in methodologies used across studies, making comparisons difficult.

To help fill this gap, we collect information about the distance to polling place and turnout for over 15 million voters in two large swing states, Pennsylvania and Georgia.

[4]Earlier observational studies on the distance to polling place find either a negative association with turnout or null results. Dyck and Gimpel (2005) find that distance to polling place is associated with lower turnout and more mail-in voting in Clark County, Nevada. Haspel and Gibbs Knotts (2005) find a negative relationship between distance to polling place and the likelihood of voting in Atlanta, Georgia. Amos et al. (2017) find a negative association between distance to polling place and voting in person on election day, but this is offset by a positive association between distance to polling place and voting early or absentee. Their study uses data from Manatee County, Florida.

[5]We use Figure 3 from Brady and McNulty (2011) to compute this effect size. There, a change in polling place of one mile causes an estimated 4% reduction on the likelihood of voting. They also estimate that most of this effect (-1.8%) is attributable to search costs of finding a new polling place, so the remaining -2.2% is attributable to distance to polling place. They also report that 55.1% of registered voters voted in the California Gubernatorial race used in their study. Thus, a 2.2% reduction in the likelihood of voting is equivalent to 1.2 percentage points.

For each registered voter, we observe the location of their residence, the location of their polling place, and whether or not they abstain, vote in person, or vote by mail in the 2018 election.

Simply regressing the likelihood of voting on distance to the polling place would likely not allow us to estimate a causal effect. Polling places are typically located in schools and other government buildings that tend to be centrally located within a town. It is likely that eligible voters who live closer to polling places differ from those who live further away in ways that matter for turnout. To estimate the causal effect of distance to polling place on voting behavior, we use the border discontinuity approach of Cantoni (2020). The innovation of Cantoni (2020) is to exploit the plausibly exogenous change in distance to polling place that occurs across the border of adjacent voting precincts. All voters are assigned to a voting precinct, and each voting precinct has a unique polling place. The intuition is that voters who live near to precinct borders are similar, except that they are assigned to different polling places and thus face different distances to the polling place. Any factors correlated with turnout should be continuous at the border, whereas the distance to a polling place is discontinuous at the border. Of the feasible causal identification strategies in this setting, we find this approach to have identifying assumptions that are most plausibly satisfied. In robustness checks, we use other methodologies from the literature, including matching and differences-in-differences with a panel of Pennsylvania voters.

We find that, on average, a one mile increase in the distance to polling place decreases the likelihood of voting at polls by 0.45 to 1.72 percentage points (p.p.). In Georgia, the decrease in voting at polls is matched with an increase in voting by mail of the same magnitude. There, a one mile increase in distance to polling place leads to an increase in the likelihood of voting by mail by up to 1.91 p.p. and a precisely estimated null effect on the likelihood of voting. The substitution to mail-in voting is substantially smaller in Pennsylvania. There, a one mile increase in distance to polling place increases the likelihood of voting by mail by up to 0.23 p.p.. Due to the muted uptake in voting by mail in Pennsylvania, there is a net negative effect of distance to polling place on the likelihood of voting overall. A one mile increase in distance to polling place decreases the likelihood of voting by 1.23 p.p. in the primary election and 0.99 p.p. in the general election. The differences in substitution to mail-in ballots across states is most likely due to differences in absentee voting laws. In 2018, Pennsylvania required an excuse for a voter to vote by mail, whereas any registered voter could request to vote by mail in Georgia.

These main estimates are an order of magnitude smaller than those from the most comparable study, which included urban areas of Massachusetts and Minnesota. However, the average effects mask significant heterogeneity. If we restrict attention to the three largest urban areas in our sample (Atlanta, Philadelphia, and Pittsburgh), then we see larger effects. In Atlanta, a larger distance to polling place induces a decrease in at-poll voting (and an increase in mail-in voting) that is up to 90% larger than the average statewide estimates. Because our methodology closely follows that of Cantoni (2020), location-specific factors and heterogeneous effects likely explain the discrepancies in estimated effect sizes across studies. We then turn to heterogeneous effects analyses to better understand what factors determine sensitivity to polling place locations.

The analysis of the two large states allows us to explore heterogeneous effects by demographic characteristics, party affiliation, economic variables, and transportation. For this analysis we use both individual-level data from voter registration files and characteristics of the Census blocks or block groups that registered voters belong to. Differences in the sensitivity of voting behavior to polling place location are minimal when we compare groups by sex or party affiliation. There is some suggestive evidence that older voters are more likely than younger voters to switch to voting by mail instead of voting in person as distance to polling place increases.

Race and ethnicity are particularly important factors to consider in the context of costs of voting. In recent years, decisions to close or move polling places have come under increased scrutiny due to concerns over voter suppression.[6] Related to these concerns, recent studies show that Black voters are more likely to experience longer waiting times at polls (Chen et al. 2020) and are more likely to have their mail-in ballots rejected (Shino et al. 2021). In Georgia, where we have data about the race of registered voters, we find that Black and Hispanic voters tend to live slightly closer to polling places, on average. We do not find large differences across racial groups in terms of sensitivity of voting behavior to distance to polling place. This evidence alone, however, does not rule out the possibility that polling place changes or closures could have disparate impacts by race or ethnicity.

Education, income, and mode of transportation are all important for understanding how voters respond to distance to polling place. Voters in areas with high reliance on public transportation are particularly sensitive to distance to polling place. For areas in the top quartile of commuting by public transportation, a 1 mile increase in distance to polling place reduces in-person voting by 1.68 p.p. in and by 3.18 p.p. in Georgia. The estimated effect of distance to polling place on voting in person is similar for voters regardless of income and education. However, substitution to mail-in voting is greater in high-income and high-education areas. As a result, in Georgia, distance to polling place reduces overall turnout only among voters in areas with the highest poverty rates.

[6] "The Georgia Governor's Race Has Brought Voter Suppression Into Full View", *The Atlantic.* Retrieved Novermber 6, 2018. "Republican Voter Suppression Efforts Are Targeting Minorities, Journalist Says", *NPR.* Retrieved October 23, 2018.

The estimated effects of distance to polling place on voting behavior tell us how sensitive voters are to a change in the distance to their polling place. But how important are these effects in aggregate? How would turnout respond to realistic changes in polling places? To answer these questions, we need to choose a counterfactual allocation of polling places. We propose the turnout-maximizing allocation of polling places as a useful counterfactual. We numerically solve a planner's problem to maximize turnout in an existing voting precinct by choosing the location of a polling place. The turnout-maximizing polling place location depends on the geographic allocation of voters as well as the estimated sensitivity of turnout to distance to polling place. This exercise is feasible for Pennsylvania only, since in Georgia there is an overall null effect of distance to polling place on turnout.

We find that the polling places used in 2018 in Pennsylvania tend to be located near the turnout-maximizing polling places in their respective precincts. The mean distance between existing and optimal polling places is 0.41 miles in urban areas and 1.04 miles in rural areas. Implementing the optimal polling place would lead to a modest increase in turnout of 0.16 p.p., on average. If Pennsylvania were to implement all optimal polling places, then we estimate turnout would increase by 15,852 votes, a 0.2 p.p. or 0.3% increase. This is on par with the effects of one additional day of early voting (Kaplan and Yuan 2020) or mailings that encourage people to vote (Green et al. 2013). Though these gains are relatively small, margins of victory in recent presidential elections in Pennsylvania have been narrow as well. In Pennsylvania, President Biden won by 80,555 votes in 2020 and former President Trump won by 44,292 votes in 2016.

The counterfactual simulations also have policy relevance. We can use the counterfactual exercise to identify precincts with the largest potential gains to improving polling place location. Among the 92 precincts in the top 1% of gains to turnout rates, the optimal polling place is approximately 4 miles away from the existing polling place. Implementing the optimal polling place in these targeted precincts would increase turnout by an average of 2.5 p.p.. This is on par with more successful voter mobilization tactics, including social pressure (Gerber et al. 2017), canvassing (Green et al. 2013), and large scale advertising campaigns (Enos and Fowler 2018).

Our findings highlight some important lessons for studies of polling places in the future. First, the importance of polling place locations for voter participation will vary depending on the context. If the affected population largely owns cars and drives to work, then a one mile increase in distance to a polling place will have a relatively small effect. On the other hand, a small change to a polling place in an area where people rely on public transportation can have significant effects on turnout. While polling closures in rural areas typically draw media attention,[7] a change to polling location in

[7] "Voting precincts closed across Georgia since election oversight lifted" *Atlanta Journal-Constitution*

an urban area might have an even larger effect on turnout and should be evaluated carefully by election commissions.

Electoral design is also likely important. With two states, we cannot determine how state-level policies affect sensitivity of voting to polling place locations. However, evidence from Georgia suggests that the availability of no-excuse voting by mail is important for mitigating the cost of traveling to a polling place. A longer distance to the polling place makes a voter in Georgia less likely to vote at polls and more likely to vote by mail. In Pennsylvania, by contrast, fewer voters substitute to mail-in voting. Importantly, the average response of voting in person to distance to polling place is similar in both states. Distance to polling place could deter voters from voting at polls, regardless of whether or not there is a convenient alternative method of voting.

2 Institutional Background

We study the 2018 primary and general elections in Pennsylvania and Georgia. In 2018, there were 8.6 million registered voters in Pennsylvania and 6.9 million voters in Georgia.[8] In many states, including Pennsylvania and Georgia, 2018 was a year of historically high turnout for a midterm election (58% and 53 % of registered voters cast ballots compared to 43% and 37% in 2014, respectively). Turnout was much lower for the primary elections (12% for Pennsylvania and 17% for Georgia).

The primary elections took place in May and the general elections on November 6, 2018. Both states elected their governor and all state executives, as well as members of the state legislature and the U.S. House of Representatives. One U.S. Senate seat was up for election in Pennsylvania and none were in Georgia.

Pennsylvania and Georgia differ in a number of election policies.[9] In Pennsylvania, voting in person happened only on election day in 2018; there was no early voting. Voters also required an excuse to vote by absentee ballot at this time in Pennsylvania.[10] In contrast, early voting began in Georgia three-weeks before election day and any voter could request a mail-in ballot up to 180 days before the election.[11]

Both states are divided into voting precincts by local government authorities (either county election commissions or municipal or county heads of government). An accessi-

August 31, 2018.

[8] Voter Registration Statistics from Pennsylvania's Department of State and Voter Registration Statistics from Georgia's Secretary if State Office

[9] Election policies were retrieved from Pennsylvania and Georgia Secretary of State websites and from state election law: Pennsylvania Statutes Title 25 and Georgia Title Code 25.

[10] Pennsylvania introduced early voting and no-excuse mail-in voting in 2019, after the sample period, with Act 77.

[11] Georgia introduced no-excuse absentee voting in 2005 with House Bill 244. In 2021, the state added an ID requirement to vote by absentee ballot with Senate Bill 202.

ble location within each precinct, typically a school, library, police station, or church, is chosen as the polling place location by local authorities. Barring emergencies, polling places must be announced no less than 60 days prior to an election in Georgia and 20 days prior to an election in Pennsylvania. This means that voters may register to vote before knowing exactly where their polling place will be located. Neither state uses same-day voter registration, so voters must register several weeks in advance of the election date. Importantly for our identification strategy, if voting in person, each voter may only vote on election day at the polling place for the precinct in which they reside.

Election day polls are open from 7am to 8pm in Pennsylvania and from 7am to 7pm in Georgia. If a registered voter has voted before in Pennsylvania, they do not need to bring identification. Georgia requires voters to show photo identification when voting in person.

We do not have the ability to assess the effect of these electoral policies with only two states and one cross-section of the data. However, the policies are potentially important for understanding the substantive differences in our findings for Pennsylvania and Georgia.

3 Data

From the Pennsylvania Department of State and Georgia Secretary of State, we obtain voter registration files, which include a unique voter identification, address, and voting precinct for each registered voter within the state. We merge this information with the voter history files, which records whether or not a registered voter voted in each election as well as their method of voting (at-polls or absentee).

We obtain the locations of polling places from Georgia's Secretary of State's website and from Pennsylvania's state-run polling place look-up website.[12] Next, we geocode the polling place locations and registered voter addresses using the Address Locator provided by ArcGIS.[13] The distance to polling place for each individual is measured as the Euclidian distance in miles between the voter's address and the polling place address.

It is important to note that we only observe registered voters in the voter registration files, not eligible voters. At the individual level, we can estimate the effect of distance to polling place on voting, conditional on an individual already being reg-

[12] https://www.pavoterservices.pa.gov/Pages/PollingPlaceInfo.aspx. and https://sos.ga.gov/index.php/elections. Retrieved October 2018 and July 2020.

[13] This address locator uses interpolation to locate addresses, meaning it has the latitude and longitude of the endpoints of every street. It then interpolates the latitude and longitude of the specific address based on the street endpoints.

istered to vote. There may be selection bias in this estimate if distance to polling place also affects the likelihood that an individual registers to vote. Hence, to complement the individual-level analysis, we estimate the effect of distance to polling place on block-level turnout, by aggregating votes at the Census block level and measuring turnout as a fraction of the voting-age population. Blocks are the smallest statistical area used by the Census, corresponding to roughly the size of a city block. At the block level, we have information on the voting age population, a proxy for voting eligibility, from the 2010 Census. The main outcome of interest at the block level is turnout: the total votes per voting age population. Cantoni (2020) similarly aggregates to the block-level, noting that the benefit of avoiding selection bias comes at the cost of less precise measurement of distance.[14]

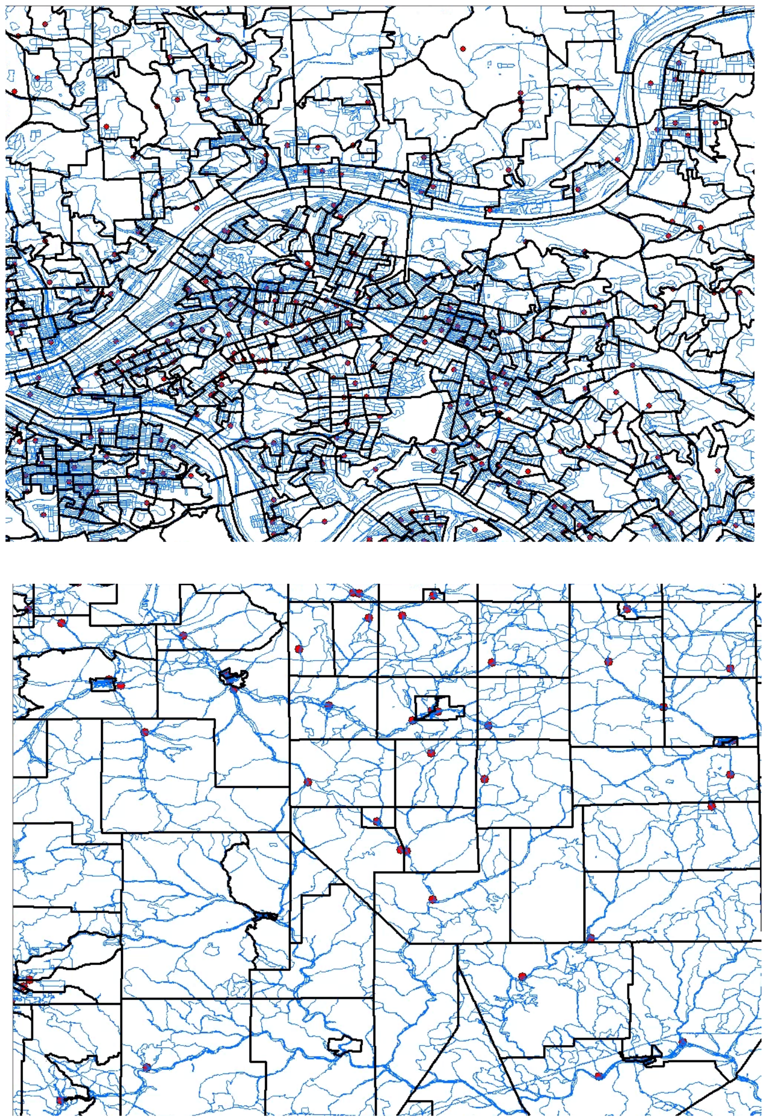
To aggregate the data to the block level, we assign each registered voter to the census block that contains their geolocated address. We likewise assign each voter to a unique voting precinct, called a voting district by the Census. Distance to polling place at the block level is measured as the average distance to polling place for all registered voters in the block.[15] To give a sense of the the geography and scale of blocks relative to precincts, Figure 1 shows voting precincts, census blocks, and polling places in an urban and rural area in Pennsylvania.

Finally, the block-level data is merged with Census data on race, ethnicity, gender and age using block identifiers. Other covariates of interest that may be correlated to both turnout and distance to polling place include car ownership, mode of travel to work and commute time, income, and unemployment. These variables are available at the block-group-level and tract-level from the American Community Survey (see Section Online Appendix A for details).

[14]Cantoni (2020) also pairs this analysis with analysis at the parcel-level. A parcel is a unit of land, typically containing one household. We don't have parcel data for important parts of Pennsylvania (e.g., Philadelphia) nor Georgia. The benefit of Parcel-level analysis is that the measurement of distance is more precise. The drawback is that there is no data on the voting age population, so parcels are assumed to have roughly the same voting age population.

[15]Blocks are relatively small so that there is little variation in distance to polling place within a block. The average standard deviation of distance to polling place within a block is 0.1 miles or 160 meters. We find similar results when we measure the distance to polling place from the block centroid instead.

Figure 1: Maps of Precincts, Census Blocks, and Polling Places



Note: The upper map shows an area of Pittsburgh, PA, (population 302,407) while the lower map is an area of Jefferson County, PA (population 43,804). Bold black lines are precinct boundaries. Thin blue lines are Census block boundaries. Red dots are polling places. Both maps cover an area of roughly 75 square miles (194 square kilometers).

4 Empirical Framework

4.1 Identification Strategy

We estimate the effect of distance to polling place on the likelihood of voting at the poll, voting by absentee ballot (i.e., voting by mail), and voting by either method. Estimating this causal effect presents several challenges. Simply regressing outcomes on distance to polling place is not credible since polling places are non-randomly located. Local election officials are supposed to choose convenient and accessible locations for polling places. Schools and other public buildings are most frequently selected as polling places. Voters who live close to polling places may therefore differ systematically from voters who tend to live far away from polling places in ways that are important for turnout. These differences might be unobservable or not adequately measured due to the aggregated nature of some of the covariates. For example, adults who choose to live close to a school may tend to have school-age children, and therefore belong to a demographic group with a relatively low turnout rate (Wolfinger and Raymond 2008). We therefore use an identification strategy that exploits discontinuities in distance to the polling place at the borders of voting precincts. Intuitively, two neighbors who live on opposite sides of a voting precinct border should be comparable in dimensions that may affect voting behavior, but differ in their assigned polling place because they happen to be on opposite sides of the precinct border.

We first apply this methodology using the individual-level data to estimate the effect of distance to polling place on the likelihood of voting, conditional on being registered. Each individual is assigned to the nearest precinct border and is only included in the sample if they reside within 0.05 miles (161 feet or 81 meters) of the border.[16] Additionally, we restrict attention to segments of precinct borders that to do not overlap with other important boundaries that might cause residents to sort on either side of the border. Voting precinct boundaries included in our sample do not overlap with school district boundaries, town or county boundaries, nor boundaries for state or federal congressional districts. Figure 2 shows the voting precinct border segments that are included and excluded in the samples for Georgia and Pennsylvania.

To limit noise from imprecise geocoding, we only consider individuals that are less than 10 miles from their polling place (2.44 standard deviations from the mean in Pennsylvania and 3.09 standard deviations from the mean in Georgia).[17] Finally, we

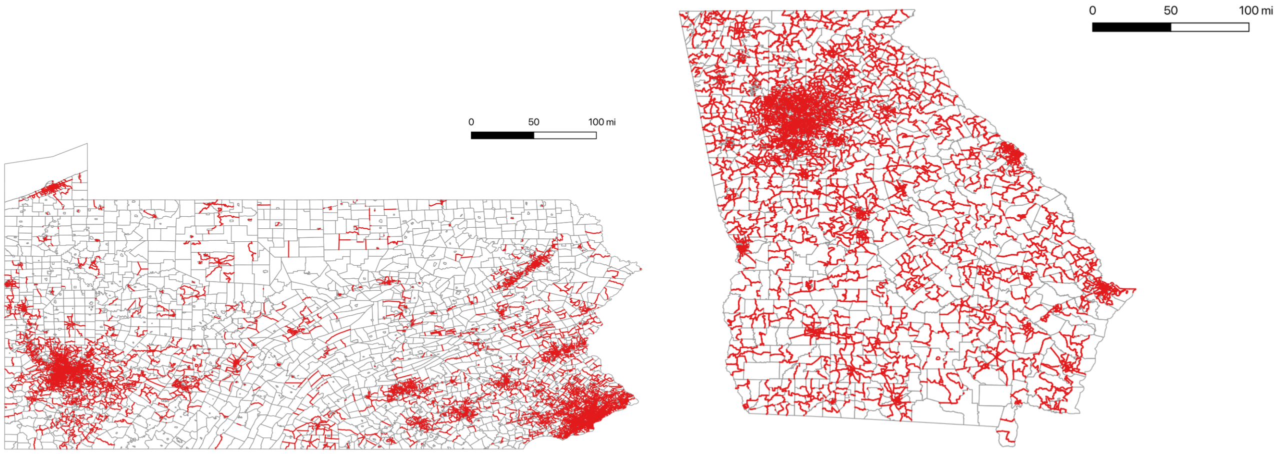
[16] We obtain 2010 voting tabulation district boundaries from the Census, whereas the polling places were used in 2018 elections. There may be some differences between boundaries reported to the Census in 2010 and implemented by the states in 2018. To account for this, we only include a border in the sample if the voters living near the border are also assigned to two different polling places in 2018.

[17] In Online Appendix B we show that estimates are not sensitive to the sample selection criteria by including voters who live more than ten miles away from the polling place and by incrementally including

discard border segments that include both rural and urban areas since these blocks are unlikely to satisfy identifying assumptions. This selection leaves us with 1,704,797 voters (20.7% of all voters) located in 9,921 border segments in Pennsylvania and 495,641 voters (7.1% of all voters) located in 3,945 border segments in Georgia. On average, each border segment has 355 voters in Pennsylvania and 358 voters in Georgia.

Figure 2: Voting precinct borders

A. Pennsylvania B. Georgia



Note: These maps show voting precinct borders in Pennsylvania and Georgia. Only voters living near the borders highlighted in red are included in the regression samples. These red borders do not overlap with the borders of school districts, towns, counties, state legislative districts, or federal congressional districts.

We use a similar empirical design for the block-level dataset to estimate the effect of the distance to polling place on turnout. Each block is assigned to the nearest voting precinct border and is only included in the sample if its centroid is within 0.2 miles (1056 feet or 322 meters) of the voting precinct border. Conditioning on the distance from the block centroid to the border allow us to only consider those blocks where most of the population lives close to the border. We use the same selection of border segments as in the individual-level data. This selection leaves us with 111,782 blocks (26.5% of all blocks) in Pennsylvania and 35,137 (12.1% of all blocks) blocks in Georgia. On average, each border has 27 blocks in Pennsylvania and 22 blocks in Georgia.

The main empirical specification can be written as follows:

$$vote_i = \delta_{s(i)} + \beta distance_i + \mathcal{P}_i'\rho + \mathcal{X}_{b(i)}'\iota + \epsilon_i \quad (1)$$

where $vote_i$ indicates whether or not registered voter i voted. We also separately

voters that live further away from the precinct border, up to 0.5 miles.

consider an indicator for whether a voter cast a ballot in person at polls and an indicator for whether a voter cast an absentee ballot by mail. For ease of interpreting small coefficients, we have $vote_i$=100 if voter i votes and $vote_i = 0$ if voter i does not vote. The variable $distance_i$ is the distance in miles between a voter's residence and polling place. The coefficient β is then interpreted as the effect of a one mile increase in distance to polling place on the likelihood of voting, measured in percentage points. The variable $\delta_{s(i)}$ is a fixed effect for a segment of a precinct border, where each voter i is assigned to a unique segment $s(i)$. The specification includes individual-level controls ($\mathcal{P}_i$) and block- or block-group-level controls ($\mathcal{X}_{b(i)}$).[18]

At the block-level, the estimating equation is:

$$turnout_b = \delta_{s(b)} + \beta distance_b + \mathcal{X}_b' \iota + \epsilon_b \tag{2}$$

where $turnout_b$ is the percent of voting-age population in block b that votes, $distance_b$ is the average distance to the polling place among registered voters in the block, and $\mathcal{X}_b$ is block-level observables.[19]

The identifying assumption in both of our specifications is that all unobservable factors affecting the likelihood of voting are uncorrelated with distance to polling place across and along border lines, conditional on observables.

Following Cantoni (2020), we relax this identifying assumption with a second set of estimating equations that include county-specific controls for latitude and longitude:

$$vote_i = \delta_{s(i)} + \beta distance_i + \alpha_{c(i)} lat_i + \gamma_{c(i)} lon_i + \mathcal{P}_i' \rho + \mathcal{X}_{b(i)}' \iota + \epsilon_i \tag{3}$$

$$turnout_b = \delta_{s(b)} + \beta distance_b + \alpha_{c(b)} lat_b + \gamma_{c(b)} lon_b + \mathcal{X}_{b(i)}' \iota + \epsilon_b \tag{4}$$

In this less parsimonious model, the effect is identified using only the discontinuities in distance to polling place at the border. The identifying assumption is that any variables that affect turnout, apart from distance, are continuous at all points of the precinct border. Because voters on either side of the border are assigned to different

[18] At the individual level, we include indicators for being a registered Democrat, registered Republican, female, and belonging to age groups 30-49, 50-64, and 65 and up. At the block and block-group level, we include population, voting age population, percent Black, percent Hispanic, median household income, percent without high school diploma, cars per household, percent who commute to work by walking, percent with commute time to work less than five minutes, and percent with commute time to work greater than 60 minutes.

[19] We include the block-level covariates as in the individual-level analysis. We also include percent of registered voters that are Democrats and percent of registered voters that are Republicans by aggregating the individual-level data. We replace age and sex indicators with percent of the block that is ages 30-49, ages 50-64, and ages 65 and up and percent female.

polling places, there is a discontinuity in the variable $distance_i$ at the border.

4.2 Summary Statistics and Balance Tests

In 2018, there were 7,014 polling places in Pennsylvania and 2,340 polling places in Georgia. Voters on average live 0.93 miles away from their polling place in Pennsylvania and 1.66 miles in Georgia. For the 2018 general election, the turnout was 59% of registered voters in Pennsylvania and 55% of registered voters in Georgia. In Georgia, 29.76% of those registered to vote chose to vote by mail whereas only 2.23% did the same in Pennsylvania.

Comparing the regression samples with the whole states, we find some important differences. Mainly, urban areas are over-represented in the samples: 87% of the sample is urban in Georgia and 98% in Pennsylvania, versus 75% and 80% for the whole sample, respectively. This is mainly because small towns are more likely to have only one voting precinct, and we remove voting precinct border segments that overlap with town borders. However, there are sufficient border segments in rural areas within each state sample in order to detect differences between urban and rural areas (210 segments in Pennsylvania and 1380 segments in Georgia). Full summary statistics are reported in Appendix A.1

To see how distance to polling place varies for voters within each state, we regress distance to polling place on political, demographic, and socioeconomic variables in Appendix A.2. In both states, polling places tend to be farther away for registered Republicans than for registered Democrats. Polling places are closer on average for younger registered voters, for females, and in areas where people have lower educational attainment, shorter commute times, and fewer cars per household. In Georgia, where we observe race and ethnicity of registered voters, polling places are closer to Black and Hispanic registered voters than to white registered voters, on average.

To test for balance within each border segment, we include border fixed effects and county-latitude/longitude controls. After including border-fixed effects, most of the correlations between distance to polling place and covariates are statistically insignificant. However, 2 of the 16 the variables are statistically significantly correlated with distance to polling place in Pennsylvania (indicators for age 65 and up and percent of the Census block-group that walks to work) and 3 of the 16 are in Georgia (indicator for voter being a registered Democrat, Female, and average number of cars per household in the Census bock-group). To be cautious, we include all covariates in Table A.2 in our preferred specification to control for observable differences within border segments.

5 The Effect of Distance to Polling Place on Voting

We begin with a discussion of the average effect of distance to polling place on likelihood of voting in Pennsylvania (Table 1) and Georgia (Table 2). We report the coefficient on distance to polling place for several outcomes: likelihood of voting at the polling place, likelihood of voting by absentee ballot, and the likelihood of voting by either method for both primary and general elections. In each table, Panel A reports coefficients from estimating the outcomes with controls and precinct fixed effects only, Panel B reports coefficients from estimating Equation 1, and Panel C reports coefficients from estimating Equation 3.

In Panel A of Tables 1 and 2 we observe small but statistically significant negative correlations between distance to polling place and the likelihood of voting at polls. Comparing these point estimates to those in Panels B and C, we see that correlations between voting and distance to polling place using within-precinct variation tend to *understate* the effects of distance to polling place on voting behavior.

From Panel B of Table 1, a one mile increase in distance to polling place in Pennsylvania is associated with a decrease in at poll voting of 1.35 p.p. in primary elections and 1.23 p.p. in general elections. The effects are similar in Georgia. A one mile increase in distance to polling place is associated with a decrease in at poll voting of 0.46 p.p. in the primary election and 1.71 p.p. in the general election (Panel B of Table 2).

In Georgia, the negative effect of distance to polling place on voting at polls is compensated for by a positive effect on the likelihood of absentee voting. A one mile increase in distance to polling place is associated with a 0.44 p.p. increase in absentee voting in primary elections and a 1.87 p.p. increase in absentee voting in general elections (Table 2, Panel B). Overall, there is a precisely estimated null effect of distance to polling place on the likelihood of voting ($\beta = -0.02$, $SE = 0.02$ for primary elections and $\beta = 0.16$, $SE = 0.19$ for general elections).

In Pennsylvania, there is substantially lower take-up of absentee voting in response to a larger distance to polling place such that distance to polling place has an overall negative effect on the likelihood of voting. A one mile increase in distance to polling place is associated with a 0.12 p.p. to 0.23 p.p. increase in absentee voting. The effect of distance to polling place on the likelihood of voting in Pennsylvania is -1.23 for primary elections and -0.99 p.p. for general elections. It is important to remember that Pennsylvania and Georgia differed in their requirements for voting by mail in 2018. Pennsylvania required voters who request absentee ballots to provide an excuse while Georgia did not. Without variation in no-excuse absentee voting policies during the

Table 1: The effect of distance to polling place on likelihood of voting: Pennsylvania

	Panel A: Precinct FE and Controls					
	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance (miles)	-0.9691***	-0.0051	-0.9742***	-0.4393***	0.0986***	-0.3407***
	(0.0536)	(0.0082)	(0.0548)	(0.0648)	(0.0168)	(0.0666)
N	6922990	6922990	6922990	6949983	6949983	6949983
y variable mean	18.31	0.38	18.69	56.43	2.15	58.58
R^2	0.059	0.007	0.061	0.057	0.011	0.065
	Panel B: Border FE and Controls					
	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance (miles)	-1.3523***	0.1177*	-1.2346***	-1.2255***	0.2326***	-0.9929***
	(0.2261)	(0.0611)	(0.2190)	(0.2786)	(0.0808)	(0.2669)
N	1529592	1529592	1529592	1534873	1534873	1534873
y variable mean	16.15	0.33	16.48	51.57	1.60	53.17
R^2	0.072	0.028	0.074	0.075	0.027	0.083
	Panel C: Border FE with Controls and County-Lat./Lon.					
	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance (miles)	-1.3476***	0.1201*	-1.2274***	-1.1533***	0.2322***	-0.9211***
	(0.2325)	(0.0623)	(0.2247)	(0.2766)	(0.0816)	(0.2640)
N	1529592	1529592	1529592	1534873	1534873	1534873
y variable mean	16.15	0.33	16.48	51.57	1.60	53.17
R^2	0.072	0.029	0.074	0.075	0.028	0.083

Note: Distance to polling place is measured in miles. The dependent variables are indicators for whether or not a registered voter has voted at the polling place, by absentee ballot, or by either voting method. For readability, we multiply the dependent variables by 100 so that the coefficients can be interpreted as percentage point changes in the likelihood of voting. County-Lat./Lon. refers to latitude and longitude controls, interacted with county fixed effects. All regressions include individual-level controls (registered Democrat, registered Republican, age 30-49, age 50-64, age 65 and up, female) and block-level controls (population, voting age population, percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, percent with commute time less than 5 minutes, and percent with commute greater than 60 minutes). Standard errors clustered at the border level are reported in parentheses.

Table 2: The effect of distance to polling place on likelihood of voting: Georgia

Panel A: Precinct FE and Controls

	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance (miles)	-0.2614***	0.2403***	-0.0211***	-1.1373***	1.4199***	0.2826***
	(0.0229)	(0.0227)	(0.0030)	(0.0572)	(0.0665)	(0.0556)
N	5999708	5999708	5999708	5999708	5999708	5999708
y variable mean	12.14	4.55	16.70	25.52	29.52	55.05
R^2	0.677	0.243	0.980	0.033	0.135	0.161

Panel B: Border FE and Controls

	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance (miles)	-0.4572***	0.4379***	-0.0193	-1.7106***	1.8681***	0.1575
	(0.0753)	(0.0755)	(0.0156)	(0.1794)	(0.1935)	(0.1939)
N	445881	445881	445881	445881	445881	445881
y variable mean	11.45	4.12	15.57	24.61	27.36	51.97
R^2	0.694	0.249	0.979	0.049	0.156	0.188

Panel C: Border FE with Controls and County-Lat./Lon.

	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance (miles)	-0.4677***	0.4458***	-0.0219	-1.7159***	1.9143***	0.1984
	(0.0773)	(0.0770)	(0.0155)	(0.1836)	(0.2018)	(0.2020)
N	445881	445881	445881	445881	445881	445881
y variable mean	11.45	4.12	15.57	24.61	27.36	51.97
R^2	0.694	0.250	0.979	0.050	0.158	0.189

Note: Distance to polling place is measured in miles. The dependent variables are indicators for whether or not a registered voter has voted at the polling place, by absentee ballot, or by either voting method. For readability, we multiply the dependent variables by 100 so that the coefficients can be interpreted as percentage point changes in the likelihood of voting. County-Lat./Lon. refers to latitude and longitude controls, interacted with county fixed effects. All regressions include individual-level controls (registered Democrat, registered Republican, age 30-49, age 50-64, age 65 and up, female) and block-level controls (population, voting age population, percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, percent with commute time less than 5 minutes, and percent with commute greater than 60 minutes). Standard errors clustered at the border level are reported in parentheses.

study period, we can not directly test the hypothesis that the availability of no-excuse absentee voting drives the higher uptake of absentee voting in Georgia.[20] However, the estimates are consistent with a larger share of voters substituting to mail-in voting if it's convenient to do so in response to a change in the cost of voting at polls.

The addition of latitude and longitude interacted with county-fixed effects in Panel C does little to change point estimates compared to those in Panel B for Tables 1 and 2. The estimates in Panel C suggest that the effect of a mile increase in distance to polling place on the likelihood of voting in a general election is negative 0.92 p.p. in Pennsylvania with no statistically significant effect in Georgia. For remaining analyses, we will largely focus on the more parsimonious estimating equation without county-specific latitude and longitude controls, as in Panel B.

We find similar results from the block-level analysis, in which we estimate the effect of distance to polling place on turnout, unconditional on voter registration. Tables A.3 and A.4. A one mile increase in distance to polling place is associated with a 0.74 p.p. reduction in turnout in Pennsylvania (Table A.3, panel C). In Georgia, point estimates for voting at polls and by absentee ballot are very similar to point estimates from the individual-level analysis (Table A.4). There is still a statistically insignificant effect of distance to polling place on overall turnout, though the point estimate is less precise than in the individual-level analysis ($\beta = -0.625$, $SE = 0.349$ for the general election, Table A.4 Panel C). The similarity in results from individual-level and block-level analyses suggests that distance to polling place does not meaningfully affect a voter's decision to register to vote. This may be because registering to vote is an infrequent and relatively low-cost action that largely takes place before polling places are known.

Across all specifications, the estimates of the average effect of distance to polling place on overall turnout in Georgia and Pennsylvania are an order of magnitude smaller than those estimated in Cantoni (2020). The same is true if we employ a differences-in-differences approach for a panel of individual voters in Pennsylvania (Section A.3) and if we use a geographic matching approach for block-level analysis (Section A.4).[21] The estimates from Pennsylvania are closer to those in Brady and McNulty (2011), whereas the null estimates in Georgia coincide with findings in Clinton et al. (2020). To reconcile these differences, we study how estimates vary within Pennsylvania and Georgia depending on the context.

[20] The switch from excuse-only to no-excuse absentee voting had no immediate effect on the sensitivity of voters to distance to polling place in the case of Minnesota (Cantoni 2020). In Georgia, no-excuse absentee voting may function differently, since it was a long-standing policy by the time of the 2018 election.

[21] Using the Pennsylvania panel data we can also estimate the effect of a change in polling place location, separately from the change in distance to polling place. The change in polling place may reduce turnout due to search costs, as voters have to identify and locate their new polling place (as in Amos et al. 2017, Brady and McNulty 2011, Clinton et al. 2020, Yoder 2018). We find evidence of search costs in primary elections, but not in general elections (Appendix A.3).

First, 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 area and Minneapolis, Minnesota. Compared to the statewide Pennsylvania and Georgia samples, the census blocks in these areas are higher in population, income, and education. We pool all census-block level data and estimate a propensity score for the likelihood of being in the Massachusetts and Minnesota sample. We use a logit specification and the covariates used by Cantoni (population, income, race, car ownership, and education). Then, we construct a matched sample by selecting the blocks with the highest propensity score so that we have a sample size roughly equivalent to that of Cantoni (2020). For comparability, we consider voting in general elections in midterm years (2018 in Georgia and Pennsylvania, 2014 in Massachusetts and Minneapolis).

Table 3 reports the border fixed effects estimates for full state samples (columns 1 and 2), for the matched samples (column 3) and for the urban areas in Massachusetts and Minnesota (column 4). Point estimates in the matched sample are roughly 6 to 8 times larger than point estimates in the full state samples. The point estimate in the matched sample is -4.49 (SE = 3.50), comparable to -5.44 (SE = 2.48) in the Massachusetts and Minnesota sample. This analysis suggests that the effects of distance to polling place on turnout are specific to the setting and electoral design. Estimates based on small areas are unlikely to generalize to larger areas or to other states.

Table 3: Comparing state samples to urban areas: Border fixed effects regressions

	State Samples		Matched Sample	Urban Areas
	(1) GA	(2) PA	(3) GA and PA	(4) MA and MN
Distance to polling place (mi)	-0.733***	-0.562***	-4.494	-5.435**
	(0.135)	(0.122)	(3.504)	(2.484)
N	84171	165082	1538	1694
y variable mean	54.002	47.584	60.253	38.229
R^2	0.250	0.294	0.478	0.595

Note: The Urban Areas sample is provided by Cantoni (2020) and include data from the Boston, Massachusetts area (MA) and Minneapolis, Minnesota area (MN). The dependent variable is turnout in the 2018 midterm election for Georgia (GA) and Pennsylvania (PA), and the dependent variable is 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.

Next, in Table 4, we estimate equation 1 separately for the three largest cities in our sample: Philadelphia, Pittsburgh, and Atlanta. Voters in these urban areas are significantly more sensitive than the statewide average estimates would indicate. A mile increase in distance to polling place is associated with a decrease in the likelihood of voting in the general election of 2.62 p.p. in Philadelphia and 0.61 p.p. in Pittsburgh. As in the rest of Pennsylvania, there is little evidence of substitution into absentee voting. In Atlanta, Georgia, a mile increase in distance leads to a decrease in turnout of 3.16 p.p. and a compensating increase in absentee voting of 3.28 p.p..

To explore the role of location further, we estimate the effect of distance to polling place on the likelihood of voting separately for each county. Due to limited power within some counties, we do not include additional covariates in each regression. Figure 3 shows maps of Pennsylvania and Georgia, indicating the point estimate for each county and whether or not the coefficient is precisely estimated. Cities with over 100,000 population are also indicated. A first observation from these maps is that voters in urban areas and surrounding suburbs tend to be more sensitive to distance to polling place than voters in rural areas, consistent with results for the largest cities in Table 4. However, there is also large variance in point estimates within rural areas of each state.

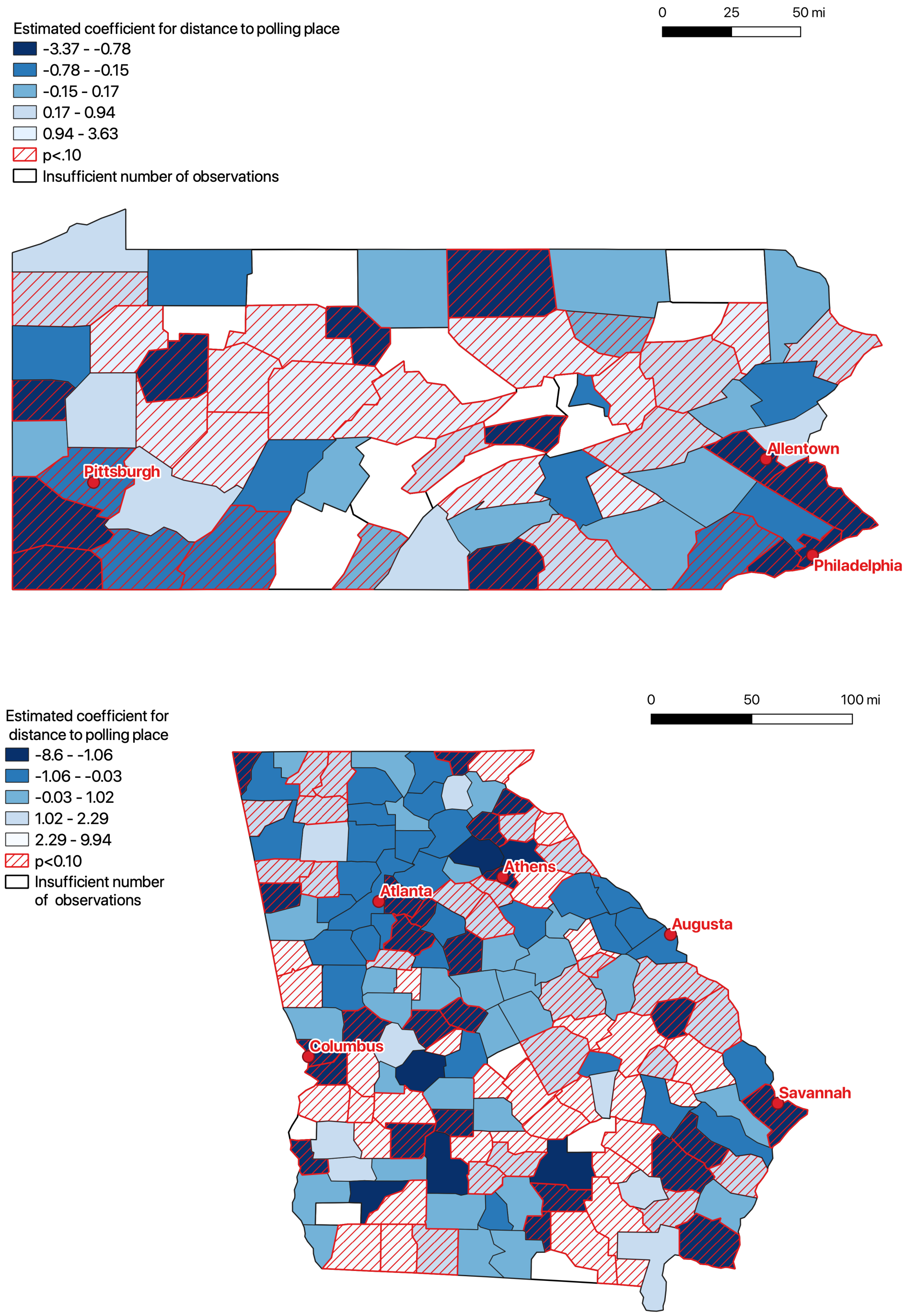
Overall, voters in urban areas appear to be more sensitive to distance to polling place. In line with these findings, the average effects are driven by those who live closer to polling places (see Online Appendix D for estimates of nonlinear effects). Next, we turn to understand which specific factors contribute to the variation in estimated effects across geographic areas.

Table 4: Comparing city samples: Philadelphia, Pittsburgh, and Atlanta

	Panel A: Philadelphia					
	Primary Election			General Election		
	(1) At Poll	(2) Absentee	(3) Total	(4) At Poll	(5) Absentee	(6) Total
Distance (miles)	-2.163***	-0.014	-2.177***	-2.728***	0.113	-2.615***
	(0.520)	(0.026)	(0.521)	(0.645)	(0.070)	(0.638)
N	658435	658435	658435	661231	661231	661231
y variable mean	16.26	0.19	16.45	52.61	1.13	53.74
R^2	0.119	0.016	0.122	0.090	0.021	0.098
	Panel B: Pittsburgh					
	Primary Election			General Election		
	(1) At Poll	(2) Absentee	(3) Total	(4) At Poll	(5) Absentee	(6) Total
Distance (miles)	-0.780***	0.035	-0.745***	-0.736***	0.122*	-0.614**
	(0.203)	(0.029)	(0.196)	(0.269)	(0.068)	(0.277)
N	587367	587367	587367	588804	588804	588804
y variable mean	20.22	0.58	20.79	58.38	2.52	60.90
R^2	0.100	0.019	0.105	0.074	0.022	0.086
	Panel C: Atlanta					
	Primary Election			General Election		
	(1) At Poll	(2) Absentee	(3) Total	(4) At Poll	(5) Absentee	(6) Total
Distance (miles)	-0.894***	0.876***	-0.017	-3.163***	3.284***	0.121
	(0.142)	(0.141)	(0.023)	(0.361)	(0.392)	(0.375)
N	587402	587402	587402	587402	587402	587402
y variable mean	11.55	4.58	16.13	23.30	29.83	53.12
R^2	0.676	0.251	0.975	0.023	0.181	0.188

Note: The Philadelphia sample includes all of Philadelphia county. The Pittsburgh sample includes all of Allegheny county. The Atlanta sample includes all of Fulton county. Distance to polling place measured in miles. Turnout is measured as the number of votes per voting-age population (separately for votes cast at polling places, through absentee ballots, and total). All regressions include Border fixed effects, individual-level controls (registered Democrat, registered Republican, age 30-49, age 50-64, age 65 and up, female) and block-level controls (population, voting age population, percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, percent with commute time less than 5 minutes, and percent with commute greater than 60 minutes). Standard errors clustered at the border level are reported in parentheses.

Figure 3: County-level estimates of the effect of distance to polling place on the likelihood of voting in Pennsylvania and Georgia



Note: The shaded areas indicate the estimated coefficient for distance to polling place from estimating Equation 1. The dependent variable is an indicator for whether or not a registered voter voted (either at polls or by absentee ballot). Regressions include border fixed effects and covariates. Standard errors are clustered at the border level. Counties with coefficients that are statistically significant at at the 10% level are indicated with red lines. The maps are shaded by quantiles of point estimates within each state. All cities with population greater than 100,000 are indicated with red circles.

6 Heterogeneous Effects

In this section we investigate differences in the responses to distance to polling place by demographic, economic, and political factors. When available, we use individual-level variables from voter registration files to estimate heterogeneous effects. We study differences by age, gender, race and ethnicity (available for Georgia only), and party affiliation using this approach. For each of these covariates of interest, we estimate equation 1 using sub-samples defined by a set of mutually exclusive categories.

6.1 Individual Level Heterogeneous Effects

Age. One might expect that younger voters are more sensitive to costs of voting, since they may not have formed the habit of voting (Fujiwara et al. 2016; Plutzer 2002). Newly eligible voters may evaluate the costs of traveling to a polling place or voting by mail, whereas older voters are more likely to continue voting or not voting using the same method as in the past. On the other hand, the oldest voters may be more likely to request an absentee ballot if distance to polling place is larger due to health and accessibility concerns. In Figure 4, we see suggestive evidence that older voters are more sensitive to distance to polling place than younger voters. In both Pennsylvania and Georgia, there is a pattern of increasing sensitivity to distance to polling place among older age groups. In Pennsylvania, voters ages 65 and up are the only age group in which there is a statistically significant increase in absentee voting, consistent with the idea that voters are more likely to be eligible for absentee voting in Pennsylvania in this age group.

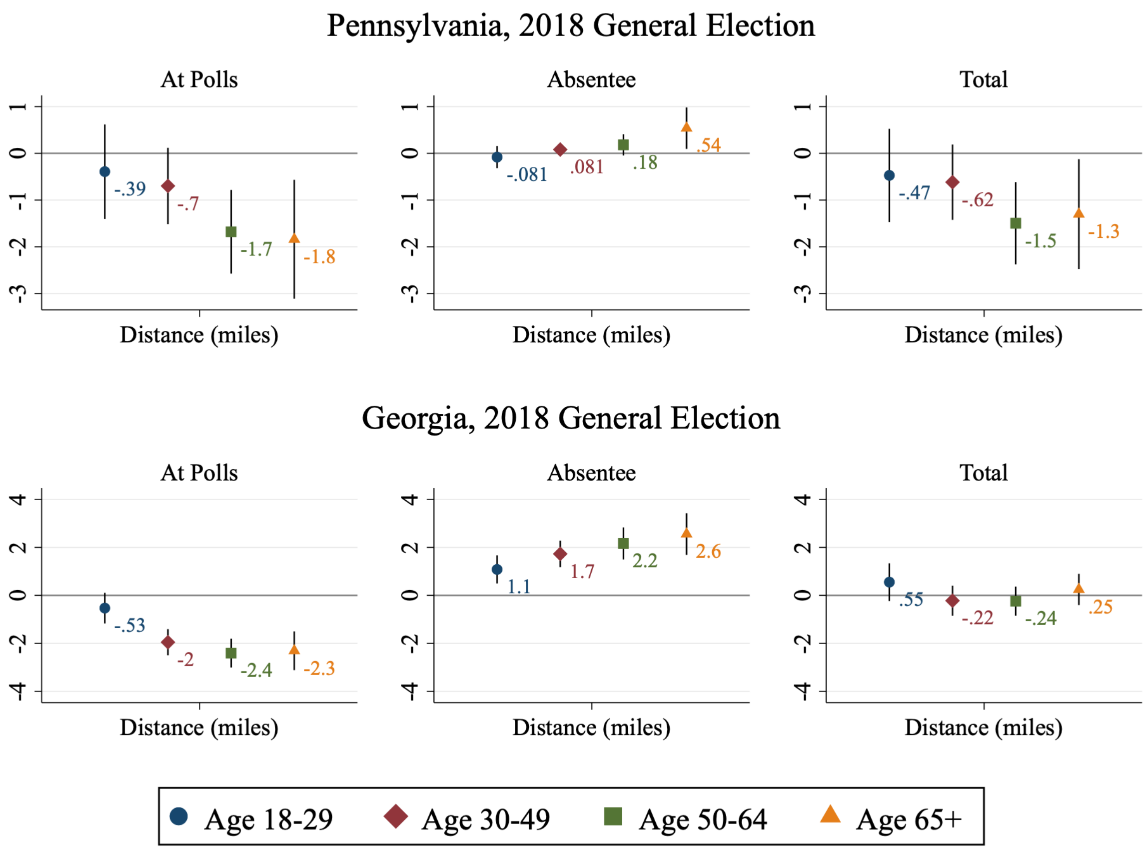
Gender. Women have turned out to vote at a higher rate than men since 1980[22]. However, we find neither large differences in distance to polling place between male and female registered voters (A.2) nor differences in sensitivity to an increase in distance to polling place (Figure 5). This suggests that differential responses to the costs of getting to the polling place do not help to explain differences in voting behavior by gender.

Race and Ethnicity. Race and ethnicity are important to consider in the context of the cost of voting due to both a persistent turnout gap between white and non-white voters (Fraga 2018; Ansolabehere et al. 2021) and long-standing concerns of voter disenfranchisement.

We use the information about race and ethnicity provided in the voter registration data, which is available only for Georgia. In Figure 6, we find no significant differ-

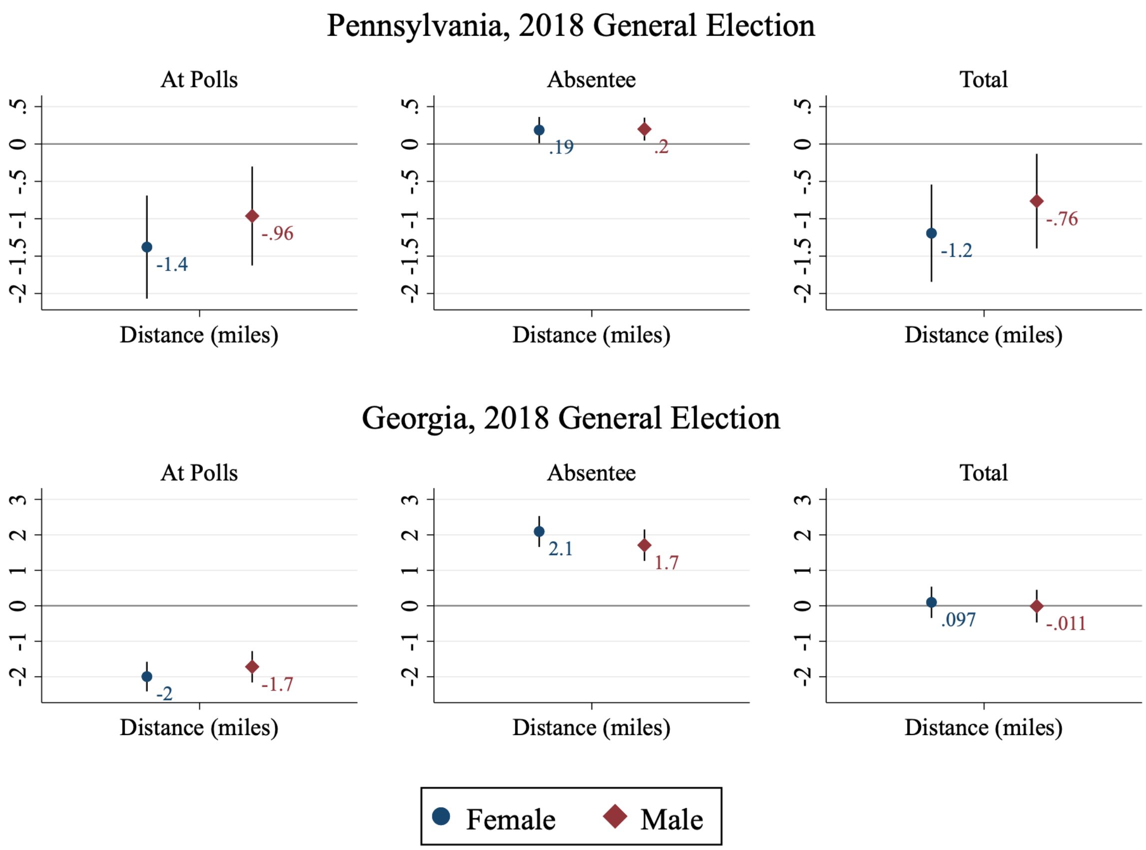
[22] Center for American Women and Politics (CAWP), "Gender Differences in Voter Turnout." Eagleton Institute of Politics, Rutgers University https://cawp.rutgers.edu/sites/default/files/resources/genderdiff.pdf. 2017. Retrieved September 5, 2020.

Figure 4: The effect of distance to polling place on likelihood of voting: by Age



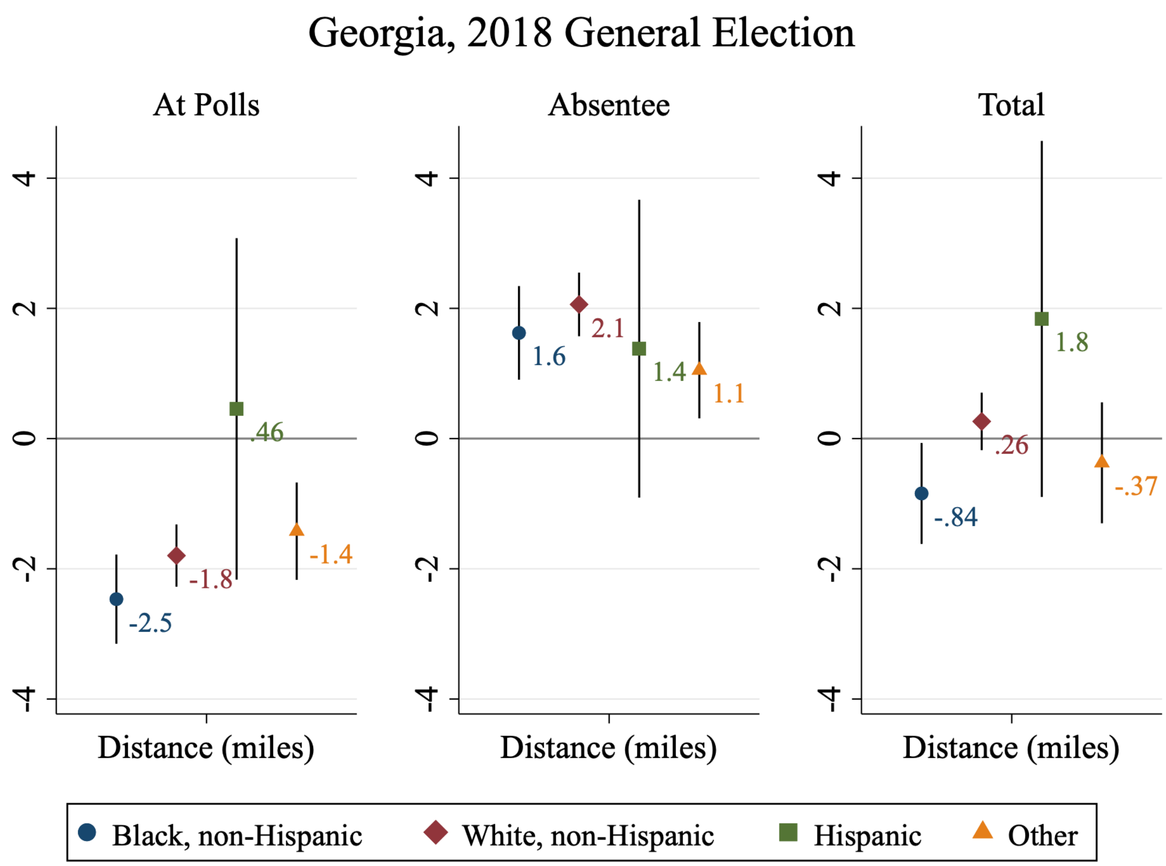
Note: The y-axis measures the coefficient on distance to polling place on the likelihood of voting at polls, by absentee ballot, or by either method. Outcome variables are re-scaled so that coefficients measure the percentage point change in the likelihood of voting. Each symbol represents a point estimate in a separate regression for the sub-sample indicated in the legend. Vertical lines indicate 95% confidence intervals. Standard errors allow for clustering at the border level. All regressions include border fixed effects, individual-level controls (registered Democrat, registered Republican, age 30-49, age 50-64, age 65 and up, female) and block-level controls (population, voting age population, percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, percent with commute time less than 5 minutes, and percent with commute greater than 60 minutes).

Figure 5: The effect of distance to polling place on likelihood of voting: by Gender



Note: The y-axis measures the coefficient on distance to polling place on the likelihood of voting at polls, by absentee ballot, or by either method. Outcome variables are re-scaled so that coefficients measure the percentage point change in the likelihood of voting. Each symbol represents a point estimate in a separate regression for the sub-sample indicated in the legend. Vertical lines indicate 95% confidence intervals. Standard errors allow for clustering at the border level. All regressions include border fixed effects, individual-level controls (registered Democrat, registered Republican, age 30-49, age 50-64, age 65 and up, female) and block-level controls (population, voting age population, percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, percent with commute time less than 5 minutes, and percent with commute greater than 60 minutes).

Figure 6: The effect of distance to polling place on likelihood of voting: by Race and Ethnicity



Note: Voter-level race and ethnicity data is available in Georgia but not in Pennsylvania. The y-axis measures the coefficient on distance to polling place on the likelihood of voting at polls, by absentee ballot, or by either method. Outcome variables are re-scaled so that coefficients measure the percentage point change in the likelihood of voting. Each symbol represents a point estimate in a separate regression for the sub-sample indicated in the legend. Vertical lines indicate 95% confidence intervals. Standard errors allow for clustering at the border level. All regressions include border fixed effects, individual-level controls (registered Democrat, registered Republican, age 30-49, age 50-64, age 65 and up, female) and block-level controls (population, voting age population, percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, percent with commute time less than 5 minutes, and percent with commute greater than 60 minutes).

ences in how Black non-Hispanic, White non-Hispanic, and Hispanic voters respond to distance to polling place. However, it is important to note that the absence of heterogeneous effects by race does not imply that changes to polling places would be race-neutral. The existing evidence of differential effects of distance to polling place by race is mixed. Cantoni (2020) finds that areas with a larger non-white population are more sensitive to distance to polling place, while Clinton et al. (2020) find that non-white voters are less likely to substitute to early voting in response to polling place changes than white voters. In studying recent polling place and precincting decisions in North Carolina, Shepherd et al. (2021) find no evidence of manipulation of polling place choices that would systematically affect voters differently by race. The mixed findings and subtle average differences across racial and ethnic groups in our data suggest that context-specific factors and electoral design are important to take into account when considering whether or not polling place locations and changes will have disparate racial impacts.

Political Party. Making use of the party affiliation information in voter registration files, we estimate the effect separately for voters registered with the Democrat party, Republican party, and for voters that do not register with either party. In Georgia, there is some suggestive evidence that Democrats are more sensitive than Republicans. However, in both states, differences between Democrats, Republicans, and Other voters are not statistically significant. Here it is important to note that both Georgia and Pennsylvania are swing states, with competitive elections in 2018. The fact that Democrats and Republicans respond similarly to costs of voting may not extend to states with more lopsided support for one political party.

6.2 Block Level Heterogeneous Effects

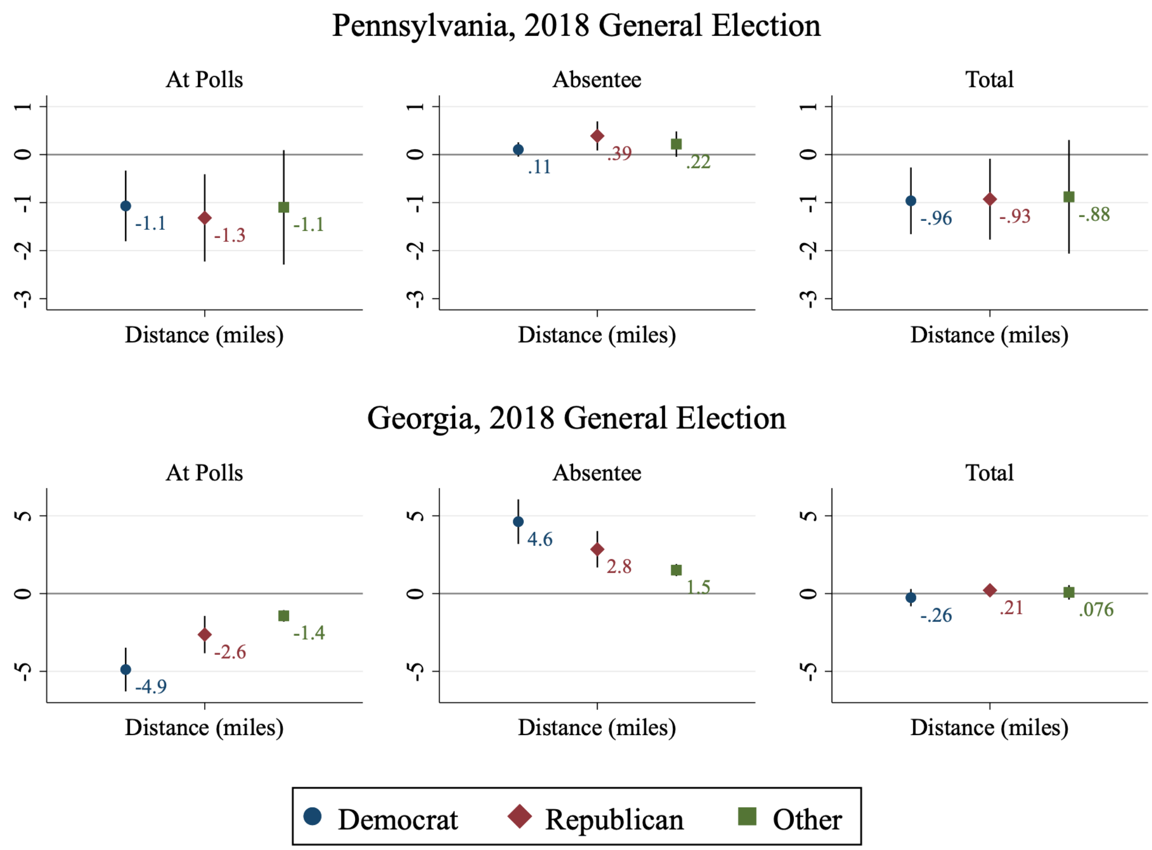
Using data from census block groups, we determine if other correlates of turnout, namely education and income, affect sensitivity of turnout to distance to polling place. We also consider mode of transportation to work, which directly affects the cost of travelling to a polling place. Since we do not observe these characteristics at the individual level, we use continuous block-level measures. Using the block-level data, we first estimate a linear interaction model:

$$turnout_b = \delta_{s(b)} + \beta_0 distance_b + \beta_1 distance_b \times z_b + \beta_2 z_b + \mathcal{X}_b' \iota + \epsilon_b \quad (5)$$

where z_b refers to one of the following moderating variables: percent of adults with a bachelor's degree or higher, median household income, and percent of employed population that uses a car for transportation to work.

In a second specification, we use the approach of Hainmueller et al. (2018) to

Figure 7: The effect of distance to polling place on likelihood of voting: by Party Affiliation



Note: The y-axis measures the coefficient on distance to polling place on the likelihood of voting at polls, by absentee ballot, or by either method. Outcome variables are re-scaled so that coefficients measure the percentage point change in the likelihood of voting. Each symbol represents a point estimate in a separate regression for the sub-sample indicated in the legend. Vertical lines indicate 95% confidence intervals. Standard errors allow for clustering at the border level. All regressions include border fixed effects, individual-level controls (registered Democrat, registered Republican, age 30-49, age 50-64, age 65 and up, female) and block-level controls (population, voting age population, percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, percent with commute time less than 5 minutes, and percent with commute greater than 60 minutes).

estimate the effect of distance to polling place on outcomes separately for four discrete bins of the moderating variable. This binning estimator allows for the marginal effect of distance to polling place on turnout outcomes to vary nonlinearly with the moderating variable. The estimating equation is:

$$turnout_b = \delta_{s(b)} + \sum_{q=1}^{4} \left[\beta_0^q distance_b + \beta_1^q distance_b(z_b - \bar{z}^q) + \beta_2^i(z_b - \bar{z}^q)\right] I_b^q + \mathcal{X}_b'\iota + \epsilon_b \tag{6}$$

where I_b^q is an indicator variable equal to one if block b is in the i-th quartile of the moderating variable, and $\bar{z}^i$ is the median value of the moderating variable for the quartile. We report the marginal effect of distance to polling place at the median value of each quartile of the moderating variable, which is simply β_0^q.

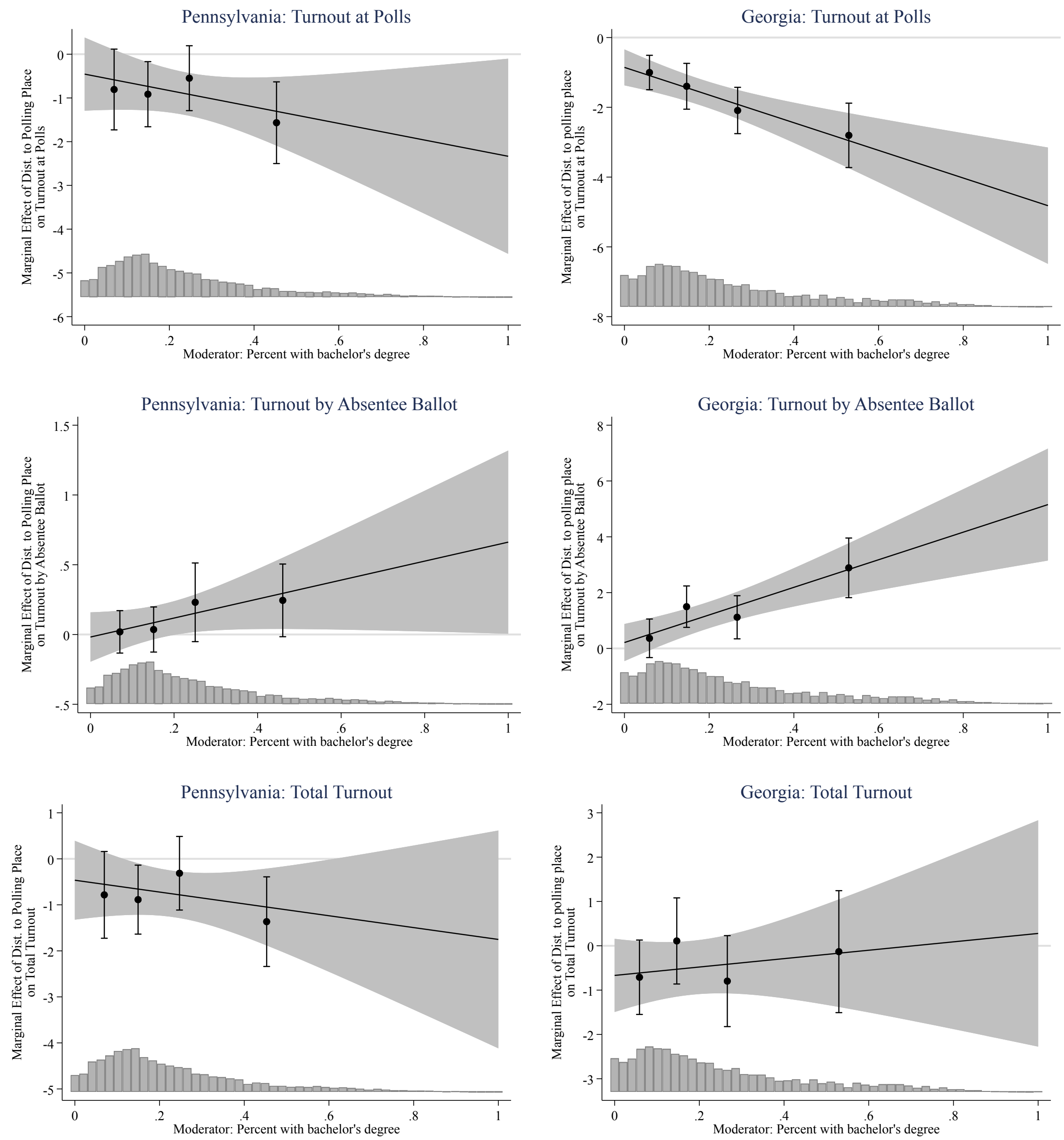
Education. Participation in elections in the United States has historically been greater among those with higher educational attainment (Milligan et al. 2004, Sondheimer and Green 2010). We report how the relationship between turnout and distance to polling place interacts with educational attainment in Figure 8. In this figure and in the remaining heterogeneous effects analyses, we plot the linear function implied by the interaction effect coefficient along with the binning estimator coefficients. Together with these estimates, we plot a histogram of the moderating variable, to avoid misleading extrapolation.[23]

In both Pennsylvania and Georgia, voters in blocks with a higher percent of adults who have earned a bachelor's degree or higher are more likely to substitute away from voting at polls and into voting by absentee ballot, relative to voters who live in areas with a lower percent of adults with a bachelors degree's or higher. Yet, the effect of distance to polling place on turnout does not vary with education. Thus, education moderates how distance to polling place affects the manner of voting, but not the likelihood of voting. The interaction effect between education and distance to polls suggests that there is a higher burden of switching into voting by mail among those with a lower educational attainment. This pattern is consistent with the finding that convenience voting can widen turnout gaps by making it easier to vote for those who are already more likely to vote (Leighley and Nagler 2014).

Income. The costs associated with traveling to the polling place could potentially lead to unequal political representation by income. In theory, the direction of how income may moderate the effect of distance to polling place on turnout is unclear. Low-income voters may have less flexibility to travel long distances on election day.

[23]These figures were created using the *interflex* Stata command (Xu, Hainmueller, Mummolo, Liu. 2017. "Interflex: Stata module to estimate multiplicative interaction models with diagnostics and visualization," Statistical Software Components S458314, Boston College Department of Economics).

Figure 8: Heterogeneous Effects – Education



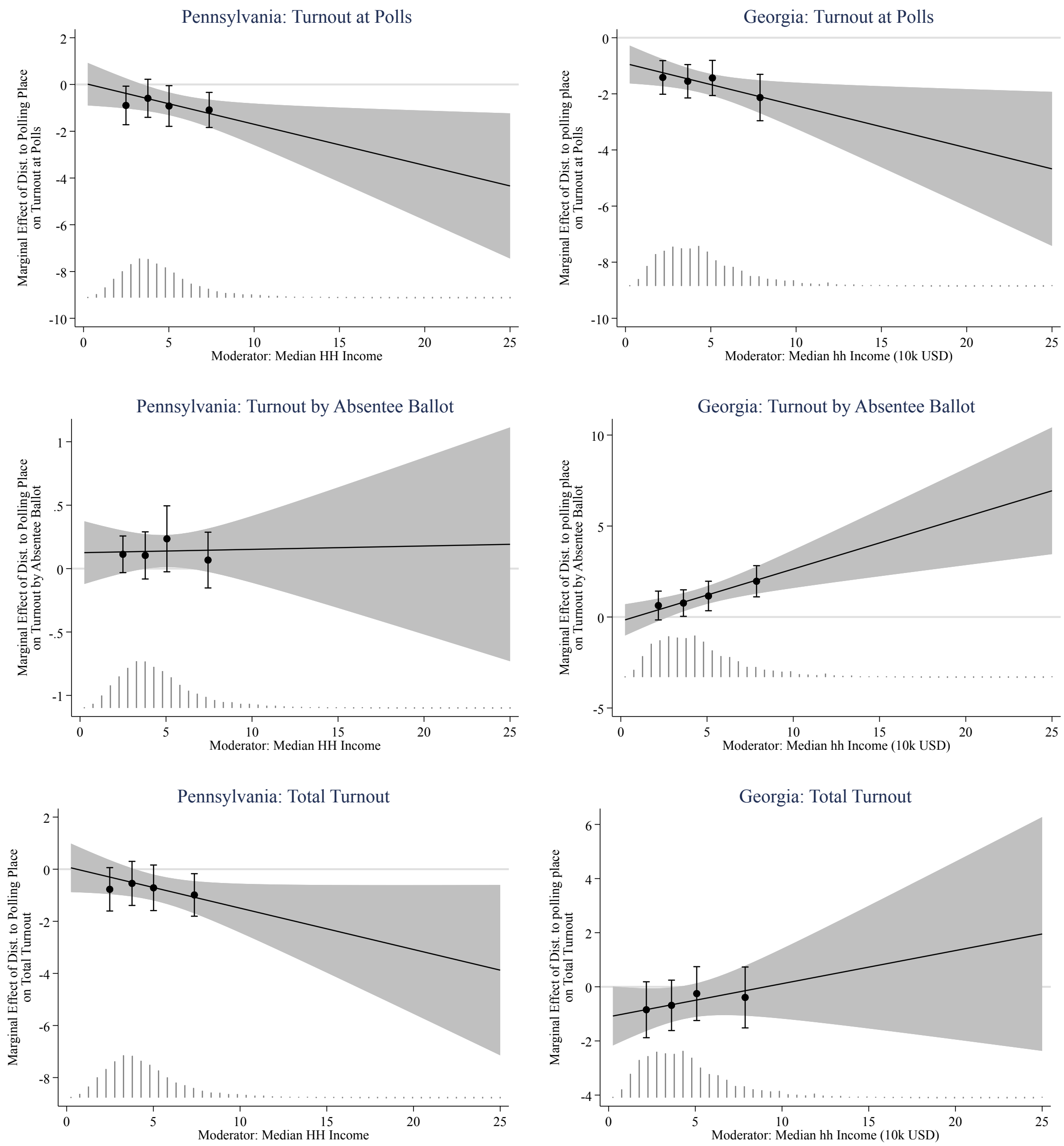
Note: This figure shows the marginal effect of distance to polling place on turnout at polls, turnout by absentee ballot, and total turnout at different levels of the moderating variable – percent of adults with a bachelor's degree or higher. The solid line plots the estimated marginal effect from the linear interaction model. The gray area shows the 95% confidence interval. The black circles and vertical lines represent the binning estimates and 95% confidence intervals. A histogram of the moderating variable is shown along the x-axis in gray. All regressions include border fixed effects and the following controls: percent registered Democrat, percent registered Republican, percent age 30-49, percent age 50-64, percent age 65 and up, percent female, population, voting age population, percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, percent with commute time less than 5 minutes, and percent with commute greater than 60 minutes.

On the other hand, high-income voters may face a higher opportunity cost of taking time from work to vote. In Figure 9, the effect of distance to polling place on voting method varies with median household income in a similar pattern as with educational attainment (Figure 8). Voters in high income blocks are less likely to vote in person and more likely to vote by mail than voters in low income blocks, given the same change in distance to polling place. When it comes to overall turnout, there are no statistically significant differences between the effect of distance to polling place in areas with the highest and lowest income quartiles.

Focusing on median household income may obscure from differences in sensitivity to distance to polling places among the lowest income voters. This is of particular importance because low income voters are underrepresented in U.S. elections (Leighley and Nagler 2014) and elected politicians are less responsive to their preferences (Gilens 2012). In Online Appendix C we show how the marginal effect of distance to polling place varies by poverty rate. Voters in the Census blocks with the high poverty rates are less likely to take up mail-in voting if distance to polling place increases than voters in low poverty areas. This result is most pronounced in Georgia, where distance to turnout rates has a negative effect on overall turnout, only in areas in the highest quartile of poverty rate. Where the poverty rate is 40%, a one mile distance to polling place decreases turnout by 1.17 p.p.. In comparison, there is a statistically and economically insignificant effect of 0.02 p.p. in areas with the lowest poverty rate. The evidence suggests that there are income-specific barriers to the take-up of voting by mail, even in states like Georgia where voters can need no excuse to access an absentee ballot. Further, these barriers can suppress overall turnout in the lowest income areas if polling places become more difficult to access.

Transportation. Finally, the cost of distance to the polling place will vary by mode of transportation. In both Pennsylvania and Georgia, we find that the effect of distance to polling place on turnout at polls is larger in areas where relatively few commuters use cars and is especially large in areas where commuters rely on public transportation (Online Appendix C). For blocks in the top quartile of commuting by public transportation, an additional mile of distance to polling place causes a a 1.68 p.p. decline in turnout in Pennsylvania and a 3.18 p.p. decline in turnout in Georgia. Unlike education and income, mode of transportation does not moderate the effect of distance to polling place on absentee voting. This is consistent with the idea that transportation affects how distance to polling place affects the cost of voting in person, but not by mail. In contrast, education and income appear to affect the cost of taking up mail-in voting, thus affecting the net effect of distance to polling place on turnout through a different mechanism.

Figure 9: Heterogeneous Effects – Median Household Income



Note: This figure shows the marginal effect of distance to polling place on turnout at polls, turnout by absentee ballot, and total turnout at different levels of the moderating variable – median household (hh) income. The solid line plots the estimated marginal effect from the linear interaction model. The gray area shows the 95% confidence interval. The black circles and vertical lines represent the binning estimates and 95% confidence intervals. A histogram of the moderating variable is shown along the x-axis in gray. All regressions include border fixed effects and the following controls: percent registered Democrat, percent registered Republican, percent age 30-49, percent age 50-64, percent age 65 and up, percent female, population, voting age population, percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, percent with commute time less than 5 minutes, and percent with commute greater than 60 minutes.

7 Optimal Polling Places

In the previous empirical sections, we document how the distance to polling place affects in-person voting, voting by mail, and overall turnout. The baseline estimates themselves have a straightforward interpretation: they tell us the effect of a one mile increase in distance to polling place on outcomes. However, from these estimates alone, it is difficult to understand how polling place location choices influence turnout in an election. For instance, any time a polling place is moved, the distance to polling place increases for some voters and decreases for others. To understand the importance of polling place locations for turnout in a realistic election setting, we must first choose a counterfactual allocation of polling places. For this purpose, we propose comparing existing polling place locations to the allocation of polling places that would maximize turnout. We first specify a planner's problem to maximize turnout by choosing the location of a polling place for a given precinct. We then numerically solve for the optimal polling place location for each precinct in Pennsylvania using information about the existing allocation of voters and estimates of how responsive voters are to distance to polling place.[24] From this exercise, we can compare the distances between existing and optimal polling places and the votes gained by implementing the optimal polling places. This simulation exercise has direct policy relevance, since we demonstrate a way to identify potential opportunities for improving voter participation.

7.1 A model of optimal polling place location

In this section, we solve an optimal polling place location problem. A planner chooses where to locate a single polling place within a precinct in order 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 $(x_i, y_i) \subset A$. Let $p(x, y)$ be the mass of voters at any point $(x, y) \subset A$.

Voters decide whether to abstain ($v_i = 0$) or vote ($v_i = 1$). To keep the model

[24] In Georgia, we estimate a precise null effect of distance to polling place on overall turnout. Thus, we can not use our estimates to determine a turnout-maximizing polling place for precincts in Georgia. This is not to say that all potential polling place locations are equivalent in states like Georgia. For this exercise, we abstract from all other costs and benefits of polling place locations apart from the turnout cost of voters' distance to polling place. For instance, there may be differential administrative costs for mail-in versus in-person turnout, and some locations may be more accessible than others.

tractable, we abstract away from the distinction between voting by mail or by absentee ballot. Since the planner's objective is to maximize turnout, the method of voting is a second-order concern. We assume that voter i's utility can be written as follows:

$$u_i(v_i = 0) = 0 + \epsilon_{0i}$$
$$u_i(v_i = 1) = a_i + c(d_i) + \epsilon_{1i}$$

where $c(.)$ is the net benefit of voting as a function of voter i's Euclidean distance to the polling place d_i, a_i denotes voter-specific net benefit of voting unrelated to distance to the polling place, and ϵ_{0i} and ϵ_{1i} are independently and identically distributed extreme value shocks. These assumptions on the functional forms of the utility and error terms translate to convenient logit choice probability functions:

$$Pr_i(v_i = 0) = \frac{1}{1 + e^{a_i + c(d_i)}} \tag{7}$$
$$Pr_i(v_i = 1) = \frac{e^{a_i + c(d_i)}}{1 + e^{a_i + c(d_i)}} \tag{8}$$

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. 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:

$$\max_{\{x^p, y^p\} \subset A} \sum_{i=1}^{N} (1 - Pr_i(v_i = 0))$$

The maximand represents the aggregate precinct-level voter turnout, which is the sum of individual probabilities of voting for all N individuals in the precinct.

7.2 Computational Procedure

Below, we outline the steps of the computational procedure used to estimate the parameters of the model introduced in the previous section along with the optimal polling place locations using these parameters as inputs.

We use the border sample from Pennsylvania to estimate the parameters in equations 5 and 6 using a fixed-effects logit specification. We impose additional structure on a_i and $c(d_i)$ in order to empirically estimate $Pr_i(v_i = 1)$. Specifically, we assume that the components are linear functions of observables that we control for in the main specifications:

$$a_i = \alpha X_i$$

$$c(d_i) = \beta d_i$$

where X_i is a vector of controls for party affiliation, age group, block-level population and block-level voting age-population, and d_i is the distance to the polling place for individual i. Due to the meaningful differences between urban and rural areas (Table 4, Figure 3, and Section 6), we estimate the model separately for urban and rural precincts. We report the estimates of the fixed-effects logit models in Online Appendix E.

Next, we use the estimates to solve for the optimal polling place locations. We solve the unconstrained optimization problem for each existing precinct using the standard Nelder-Mead algorithm. The sufficient conditions for global maxima are included in Online Appendix F.

7.3 Results

Figure 10 shows the locations of existing and optimal polling places for all precincts in Pennsylvania, as well as in an urban and rural area in the same county. This rough look at the raw data and our computed optimal polling places shows that the counterfactual estimates are not too far off from existing polling places. Although the computation does not constrain the location of the optimal polling place to be within existing precinct boundaries, we find that the optimal polling place is located within the same voting precinct as the existing polling place 62.46% of the time.[25] This suggests that there may be gains to re-precincting, that is to re-draw precincts so that they include the optimal polling place.

The mean distance between the existing polling place and the optimal polling place is 0.55 miles. Rural and urban precincts differ in the mean distance between existing and optimal polling places, with an average of 1.04 miles in rural precincts and 0.41 miles in urban precincts (see Figure 11).

The predicted gains to implementing the optimal polling place are small, on average. The mean increase in votes is 1.73 votes per precinct, equivalent to a 0.16 percentage point increase in turnout and a 0.20% increase in turnout.[26] The relatively small magnitude of the change in turnout is perhaps not surprising given that most polling

[25] It is possible that some of the optimal and existing polling places are in different 2010 census voting tabulation districts, though they are in the same precinct used in 2018. Additionally, we drop 62 precincts from this analysis in which the distance to between optimal and existing polling place was greater than 10 miles, due to the high likelihood of geocoding errors for some voters in the precinct.

[26] We use the estimated turnout for both the existing and optimal polling places in order to compute percent changes. The average estimated turnout is 79% of voting age population.

places are located near the optimal polling place and that voter sensitivity to a change in distance to polling place is also relatively small.

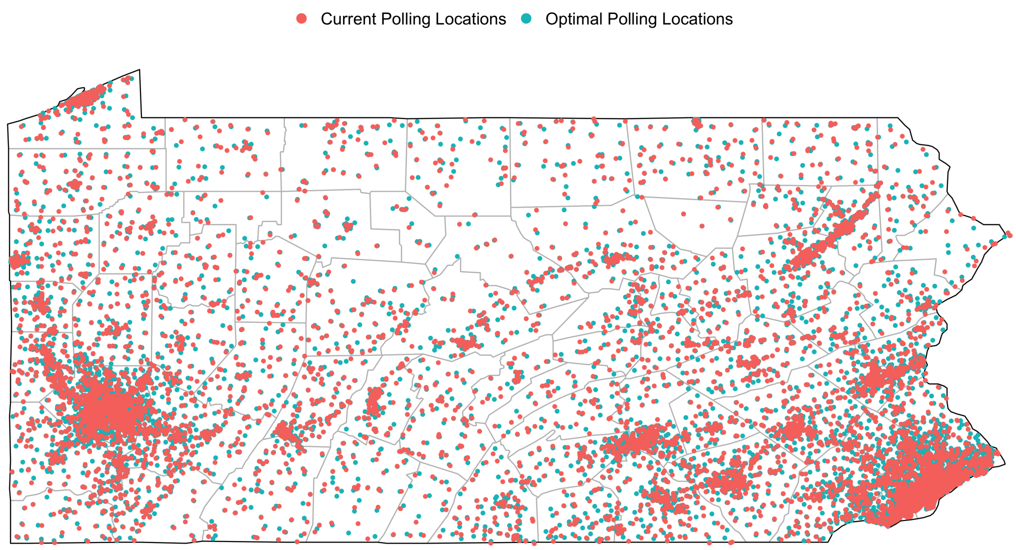
Figure 12 shows the distribution of the predicted change in percentage points of turnout for urban and rural precincts. The average gain in turnout in a rural precinct is 0.26 p.p. and the average gain in turnout in an urban precinct is 0.13 p.p.. Although polling places are further from the optimal location in rural precincts, the gains in turnout from implementing optimal polling places are similar in rural and urban areas (Figure 11). This is because rural voters are less sensitive to a change in distance to polling place than urban voters.

According to these simulations, if Pennsylvania were to implement all of the optimal polling places, then turnout would increase by 0.16 p.p., which is 0.21% or 15,852 votes. Thus, there are modest gains to turnout from implementing the turnout maximizing polling places statewide, holding the number of polling places fixed. For perspective, it helps to compare this counterfactual to other voter mobilization tactics and election policies. In a meta-analysis of 147 field experiments in which eligible voters receive mail encouraging them to vote, the average effect on turnout per mailer is 0.162 p.p. (Green et al. 2013). Kaplan and Yuan 2020 estimate that an additional day of early voting increases turnout by 0.22 percentage points. In swing states like Pennsylvania, these small gains in turnout could be meaningful given the extremely close elections in recent years.

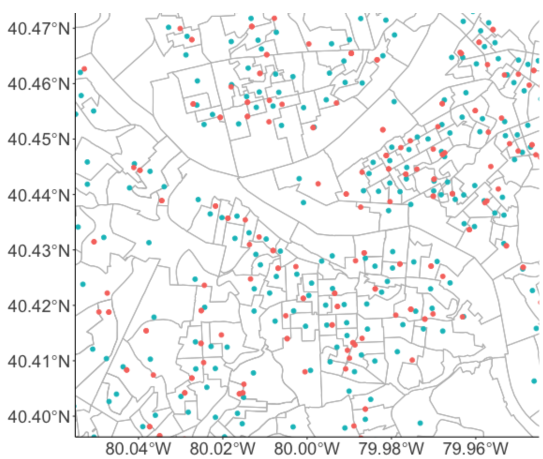
Although we find that the average precinct has small gains to improving polling place locations, we can also use this method to identify those precincts where the effect of implementing the optimal polling place would be substantial. Among the 92 precincts in the 99th percentile of gains from implementing the optimal polling place, the average increase in turnout would be 2.45 p.p. (3.28%). Half of the precincts in the 99th percentile of gains to turnout are in urban areas and half are in rural areas. There are ten precincts that could increase turnout by more than 5% by changing the polling place. In these cases, the estimated effect of implementing an optimal polling place on turnout would be on par with some of the most successful voter mobilization tactics. For instance, in a large-scale field experiment in the U.S., mailings that add elements of social pressure (e.g., the voting history of the resident as well as of their neighbors) increase turnout by 2.2 p.p. (Gerber et al. 2017). The average effect of canvassing across 71 studies is 2.54 p.p. (Green et al. 2013). Finally, Enos and Fowler (2018) estimate that large scale presidential campaigns, including all advertising and canvassing, can increase turnout by 7-8 p.p. in the most heavily targeted states. Compared to costly voter mobilization efforts, relocating a polling place could be a cost-effective way to improve voter participation. The cost effectiveness will also depend on the availability of suitable buildings near to the optimal polling place location.

Figure 10: Optimal and Existing polling place locations in Pennsylvania

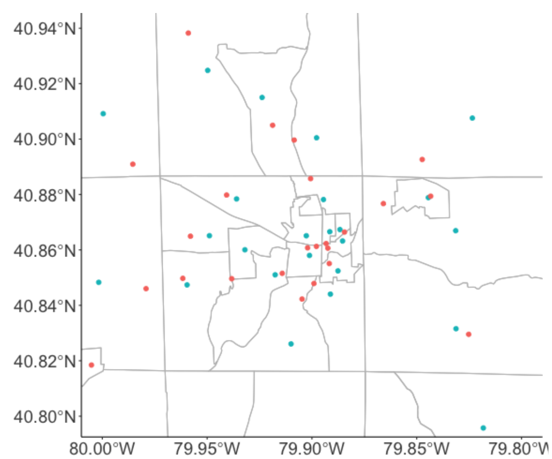
(a) All of Pennsylvania



(b) Urban area in Allegheny County, PA

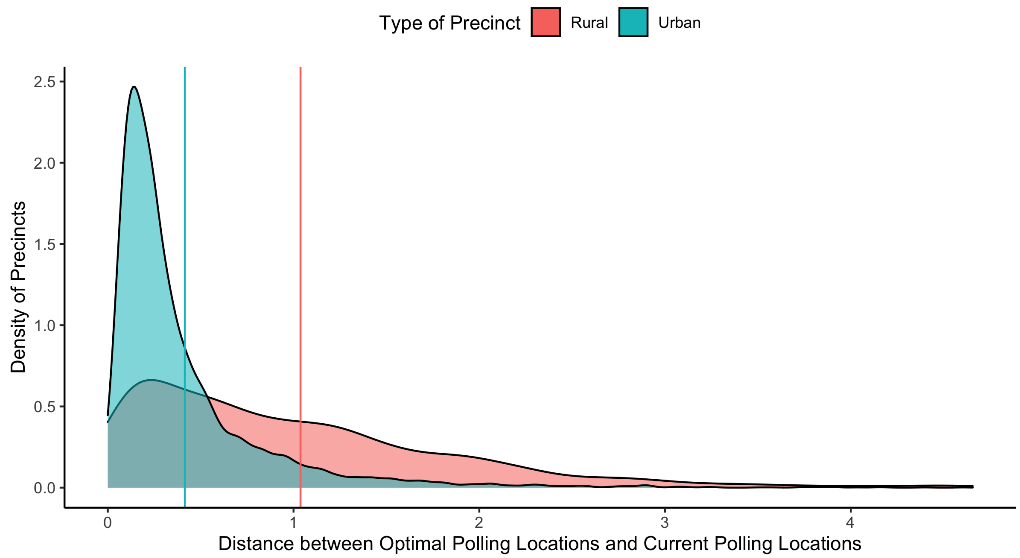


(c) Rural area in Allegheny County, PA



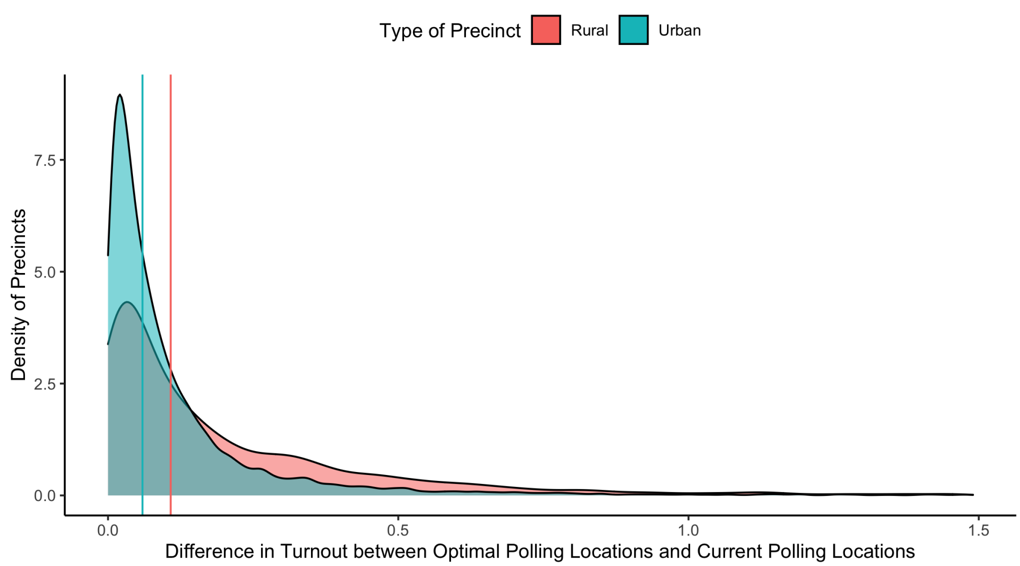
Note: This figure shows the locations of optimal and existing polling places. Panel A shows a map of the state of Pennsylvania, with current polling locations indicated in red and optimal polling places indicated in blue. The gray lines in panel (a) show county borders and the gray lines in panels (b) and (c) show precinct borders.

Figure 11: Distribution of the distance between optimal polling place and current polling place in Urban and Rural precincts



Note: This figure shows the distribution of the distance between the existing polling place and the optimal polling place for rural and urban precincts. Mean values are indicated with vertical lines. For readability, the histograms exclude the top 1% of observations. The maximum distance between optimal and current polling location is 9.92 miles.

Figure 12: Distribution of the difference in turnout (percentage points) under optimal and current polling places in Urban and Rural precincts



Note: This figure shows the distribution of the difference in turnout under the optimal polling place and under the existing polling place for rural and urban precincts. Turnout is measured in percentage points. Mean values are indicated with vertical lines. For readability, the histograms exclude the top 1% of observations. The maximum gains to turnout is 7.02 percentage points.

8 Conclusion

Among the many costs of voting, some are easy to control or ban, like polling taxes, and others are impossible to control, bad weather. The distance to the polling place is both inherent to the voting process and is a matter of policy. This paper is another step toward understanding how polling places affect voter participation.

We study the causal effect of distance to polling place on voter participation and voting method (at polls or by mail) in two large swing states, Georgia and Pennsylvania. On average, there is a small negative effect of distance to polling place on the likelihood that a registered voter goes to the poll to vote (0.45 p.p. to 1.72 p.p. per mile). In Georgia, voters fully substitute to voting by mail, such that there is no net effect of distance to polling place on turnout. In Pennsylvania, substitution to mail-in voting is limited and there is a negative net effect of distance to polling place on the likelihood of voting. An important difference between the states in the 2018 elections is that voters needed no excuse to vote by mail in Georgia and did need an excuse to vote by mail in Pennsylvania. Although we use the same geographical border discontinuity approach as in Cantoni (2020), we find substantially smaller average effects. We reconcile these differences through alternative empirical specifications and extensive analysis of heterogeneous effects. Voters tend to be more sensitive to distance to polling place in urban areas and in areas where cars are used less frequently for commuting or where public transportation is used more frequently. Low income areas and areas with lower educational attainment have a relatively smaller take-up of mail-in voting when distance to polling place increases, which leads to larger declines in turnout as well.

The results highlight some important lessons for studies of electoral design and voter participation in the future. First, it is important to use large datasets to study costs of voting in large elections, where power is needed to detect very small effects. The ability to estimate small effects with precision is especially important in settings where we expect close margins of victory. In the 2020 presidential election, the margin of victory for Joe Biden was less than one percentage point in both Georgia and Pennsylvania. Having a statewide dataset also allows for meaningful descriptive statistics of the distribution of polling places and allows us to explore context-dependent effects of moving polling places. Second, it is important to take context into account when determining the effect of any election design change. Cantoni and Pons (2022) show that the context or location of a voter has influences whether or not they vote. This paper adds that voters' sensitivity to costs of voting is also highly context-dependent.

The findings in this paper can help election commissions that face costly trade-offs in choosing how many polling places to open and where to place them. Our counterfactual exercise shows that the causal estimates from our paper are not merely of academic interest but are highly policy-relevant if local officials want a practical way of increasing turnout. In Pennsylvania, implementing optimal polling places would increase turnout by 0.21%. In future work, one might take into account the differential costs of voting in person and by mail and of opening and closing polling places, in order to determine the cost efficiency of implementing turnout-maximizing polling place locations. With a more nuanced understanding of when voters choose to vote in person, vote by mail, or abstain, election commissions have an opportunity to reduce costs of voting or to make costs of voting more equal across the population.

References

Amos, B., D. A. Smith, and C. Ste. Claire (2017). Reprecincting and Voting Behavior. *Political Behavior 39*(1), 133–156.

Ansolabehere, S., B. L. Fraga, and B. F. Schaffner (2021). The CPS Voting and Registration Supplement Overstates Minority Turnout. *Journal of Politics*, Forthcoming.

Avery, J. M. (2015). Does Who Votes Matter? Income Bias in Voter Turnout and Economic Inequality in the American States from 1980 to 2010. *Political Behavior 37*(4), 955–976.

Braconnier, C., J. Y. Dormagen, and V. Pons (2017). Voter Registration Costs and Disenfranchisement: Experimental Evidence from France. *American Political Science Review 111*(3), 584–604.

Brady, H. E. and J. E. McNulty (2011). Turning out to Vote: The Vosts of Finding and Getting to the Polling Place. *American Political Science Review 105*(1), 115–134.

Cantoni, E. (2020). A Precinct Too Far: Turnout and Voting Costs. *American Economic Journal: Applied Economics 12*(1), 61–85.

Cantoni, E. and V. Pons (2022, April). Does context outweigh individual characteristics in driving voting behavior? evidence from relocations within the united states. *American Economic Review 112*(4), 1226–72.

Chen, M. K., K. Haggag, D. G. Pope, and R. Rohla (2020). Racial Disparities in Voting Wait Times: Evidence from Smartphone Data. *Review of Economics and Statistics*, Forthcoming.

Clinton, J. D., N. Eubank, A. Fresh, and M. E. Shepherd (2020). Polling Place Changes and Political Participation: Evidence from North Carolina Presidential Elections, 2008-2016 *. *Political Science Research and Methods First View*, 1–18.

Dube, A., T. W. Lester, and M. Reich (2010). Minimum wage effects across state borders: Estimates using contiguous counties. *The Review of Economics and Statistics 92*(4), 945–964.

Dyck, J. J. and J. G. Gimpel (2005). Distance, turnout, and the convenience of voting. *Social Science Quarterly 86*(3), 531–548.

Enos, R. D. and A. Fowler (2018). Aggregate effects of large-scale campaigns on voter turnout. *Political Science Research and Methods 6*(4), 733–751.

Fraga, B. L. (2018). *The Turnout Gap.* Cambridge University Press.

Fujiwara, T., K. Meng, and T. Vogl (2016). Habit formation in voting: Evidence from rainy elections. *American Economic Journal: Applied Economics 8*(4), 160–188.

Gerber, A. S., G. A. Huber, A. H. Fang, and A. Gooch (2017). The generalizability of social pressure effects on turnout across high-salience electoral contexts: Field experimental evidence from 1.96 million citizens in 17 states. *American Politics Research 45*(4), 533–559.

Gilens, M. (2012). *Affluence and Influence.* Princeton University Press.

Gomez, B. T., T. G. Hansford, and G. A. Krause (2007). The pepublicans should pray for rain: Weather, turnout, and voting in U.S. presidential elections. *Journal of Politics 69*(3), 649–663.

Green, D. P., M. C. McGrath, and P. M. Aronow (2013). Field experiments and the study of voter turnout. *Journal of Elections, Public Opinion and Parties 23*(1), 27–48.

Hainmueller, J., J. Mummolo, and Y. Zu (2018). How much should we trust estimates from multiplicative interaction models? simple tools to improve empirical practice. *Political Analysis 27*(2), 163–192.

Haspel, M. and H. Gibbs Knotts (2005). Location, location, location: Precinct placement and the costs of voting. *Journal of Politics 67*(2), 560–573.

Highton, B. (2017). Voter identification laws and turnout in the united states. *Annual Review of Political Science 20*(1), 149–167.

Kaplan, E. and H. Yuan (2020). Early voting laws, voter turnout, and Partisan vote composition: Evidence from Ohio. *American Economic Journal: Applied Economics 12*(1), 32–60.

Keele, L. J. and R. Titiunik (2015). Geographic boundaries as regression discontinuities. *Political Analysis 23*(1), 127–155.

Leighley, J. E. and J. Nagler (2014). *Who votes now?* Princeton University Press.

Lockhart, M., S. J. Hill, J. Merolla, M. Romero, and T. Kousser (2020). America's electorate is increasingly polarized along partisan lines about voting by mail during the covid-19 crisis. *Proceedings of the National Academy of Sciences 117*(40), 24640–24642.

Martin, P. S. (2003). Voting's rewards: Voter turnout, attentive publics, and congressional allocation of federal money. *American Journal of Political Science 47*(1), 110–127.

Meredith, M. and Z. Endter (2016). Aging into absentee voting: Evidence from texas. *Working Paper*.

Meredith, M. and N. Malhotra (2011). Convenience voting can affect election outcomes. *Election Law Journal 10*(3), 227–253.

Milligan, K., E. Moretti, and P. Oreopoulos (2004). Does education improve citizenship? evidence from the united states and the united kingdom. *Journal of Public Economics 88*(9), 1667 – 1695.

Plutzer, E. (2002). Becoming a Habitual Voter: Inertia, Resources, and Growth in Young Adulthood. *The American Political Science Review 96*(1), 41–56.

Shepherd, M. E., A. Fresh, N. Eubank, and J. D. Clinton (2021). The Politics of Locating Polling Places: Race and Partisanship in North Carolina Election Administration, 2008–2016. *Election Law Journal: Rules, Politics, and Policy 20*(2), 155–177.

Shino, E., M. Suttmann-Lea, and D. A. Smith (2021). Determinants of rejected mail ballots in georgia's 2018 general election. *Political Research Quarterly 75*(1), 231–243.

Sondheimer, R. M. and D. P. Green (2010). Using experiments to estimate the effects of education on voter turnout. *American Journal of Political Science 54*(1), 174–189.

Thompson, D. M., J. A. Wu, J. Yoder, and A. B. Hall (2020). Universal vote-by-mail has no impact on partisan turnout or vote share. *Proceedings of the National Academy of Sciences of the United States of America 117*(25), 14052–14056.

Wolfinger, N. and W. Raymond (2008). Family Structure and Voter Turnout . *Social Forces 86*(5), 1863–1903.

Yoder, J. (2018). How Polling Place Changes Reduce Turnout: Evidence from Administrative Data in North Carolina. *Working Paper*.

A Appendix

A.1 Summary Statistics

Table A.1: Summary Statistics

	Pennsylvania				Georgia			
	All		Regression sample		All		Regression sample	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Voting History								
Primary election, at polls	18.41	38.75	16.30	36.94	12.30	32.85	11.50	31.91
Primary election, absentee	0.38	6.16	0.33	5.73	4.60	20.95	4.12	19.87
Primary election, total	18.79	39.06	16.63	37.24	16.90	37.48	15.62	36.31
General election, at polls	56.79	49.54	51.91	49.96	25.79	43.75	24.91	43.25
General election, absentee	2.23	14.78	1.65	12.72	29.76	45.72	27.34	44.57
General election, total	59.03	49.18	53.56	49.87	55.55	49.69	52.25	49.95
Distance to polling place (mi)	0.93	3.72	0.38	0.50	1.66	2.70	1.32	1.28
Democrat	0.48	0.50	0.64	0.48	0.08	0.27	0.09	0.28
Republican	0.38	0.49	0.22	0.41	0.09	0.28	0.06	0.25
Independent	0.14	0.35	0.14	0.35	0.83	0.37	0.85	0.36
Demographics								
Population	145.70	220.38	149.02	200.29	337.02	465.12	298.92	421.26
Voting Age Population	115.55	190.53	120.02	178.01	244.13	332.47	223.55	309.17
Percent urban	0.80	0.40	0.98	0.13	0.75	0.43	0.87	0.34
Percent Black	0.11	0.23	0.27	0.35	0.30	0.33	0.37	0.35
Percent Hispanic	0.05	0.11	0.08	0.16	0.07	0.11	0.07	0.11
Poverty Rate	0.12	0.13	0.19	0.18	0.14	0.13	0.17	0.16
Median hh Income (10k USD)	5.86	2.89	4.75	2.68	5.85	2.89	5.69	3.14
Way to work								
% Car to work	0.85	0.17	0.71	0.23	0.90	0.09	0.87	0.12
% Walk to work	0.04	0.09	0.07	0.12	0.01	0.04	0.02	0.06
% Pub transit to work	0.06	0.12	0.16	0.18	0.02	0.06	0.04	0.08
% Bike to work	0.00	0.02	0.01	0.03	0.00	0.01	0.00	0.02
Time to work								
% time to work 0-5min	0.04	0.05	0.03	0.05	0.03	0.04	0.02	0.04
% time to work 5-60min	0.88	0.08	0.88	0.10	0.88	0.08	0.89	0.08
% time to work 60min plus	0.08	0.08	0.09	0.09	0.10	0.08	0.09	0.08
N	8,245,003		1,704,797		6,980,226		495,641	

Note: For each voting history variable, we observe whether or not a registered voter votes, by method of voting. Each indicator variable is multiplied by 100 to make regression coefficients easier to interpret. Demographic variables are measured at the block, or block-group level and assigned to each individual voter that resides in the geographic area.

Table A.2: Correlates of Distance to Polling Place

	Pennsylvania			Georgia		
	(1) All Individuals	(2) Border FE Sample	(3) Border FE Sample	(4) All Individuals	(5) Border FE Sample	(6) Border FE Sample
Democrat	0.001	-0.000	-0.000	-0.020***	-0.015***	-0.015***
	(0.001)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)
Republican	0.001	-0.001	-0.001	-0.016*	-0.010	-0.010
	(0.001)	(0.001)	(0.001)	(0.009)	(0.007)	(0.006)
age 30 to 49	0.000	-0.001	-0.001	-0.003	-0.006**	-0.005*
	(0.001)	(0.001)	(0.001)	(0.003)	(0.003)	(0.002)
age 50 to 64	-0.001**	-0.002	-0.002	0.002	-0.004	-0.003
	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.003)
age 65 and up	-0.004***	-0.007***	-0.007***	-0.006	-0.012**	-0.010*
	(0.001)	(0.002)	(0.002)	(0.008)	(0.005)	(0.005)
Female	-0.001*	-0.001	-0.000	-0.005***	-0.003*	-0.003**
	(0.000)	(0.000)	(0.000)	(0.002)	(0.001)	(0.001)
Population	0.000***	-0.000	-0.000	-0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Voting Age Population	-0.000***	0.000	0.000	0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Percent Black	0.000	-0.015	-0.014	-0.099**	-0.065	-0.055
	(0.011)	(0.023)	(0.023)	(0.050)	(0.046)	(0.042)
Percent Hispanic	-0.033**	-0.014	-0.013	-0.315***	-0.155**	-0.094
	(0.014)	(0.019)	(0.018)	(0.098)	(0.072)	(0.062)
Median hh Income (10k USD)	0.004	0.004	0.002	0.003	-0.001	0.001
	(0.003)	(0.003)	(0.003)	(0.011)	(0.007)	(0.007)
Percent without high school diploma	-0.037*	-0.013	-0.014	-0.184	-0.056	-0.040
	(0.022)	(0.024)	(0.024)	(0.136)	(0.120)	(0.111)
Cars per Household	0.018	0.003	0.015	0.167***	0.140**	0.119**
	(0.014)	(0.019)	(0.015)	(0.063)	(0.055)	(0.047)
% Walk to work	-0.026	-0.071**	-0.063*	-0.525*	-0.034	-0.164
	(0.029)	(0.034)	(0.034)	(0.318)	(0.334)	(0.308)
% time to work 0-5min	-0.077	0.010	0.010	-1.522***	0.132	0.312
	(0.058)	(0.083)	(0.079)	(0.401)	(0.339)	(0.314)
% time to work 60min plus	0.006	-0.027	-0.027	0.364	0.125	0.069
	(0.023)	(0.049)	(0.049)	(0.232)	(0.177)	(0.164)
Precinct FE	X			X		
Border FE		X	X		X	X
County-Lat/Lon			X			X
N	1,064,884	1,064,884	1,064,884	420,096	420,096	420,096
Dep. variable mean	0.37	0.37	0.37	1.33	1.33	1.33
R^2	0.897	0.768	0.777	0.775	0.836	0.853

A.2 Block-level estimates

Table A.3: The effect of distance to polling place on turnout: Pennsylvania

	Panel A: OLS with Precinct FE and Controls					
	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance to polling place	-1.0123***	-0.0424	-1.0536***	-1.2724***	0.1549	-1.1024***
	(0.2463)	(0.0489)	(0.2553)	(0.3594)	(0.1091)	(0.3785)
N	101771	102362	101769	97768	102313	97758
y variable mean	15.72	0.34	16.04	44.47	1.76	46.04
R^2	0.283	0.094	0.286	0.337	0.155	0.354
	Panel B: Border FE with Controls					
	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance to polling place	-0.5278***	0.0055	-0.5376***	-0.8811***	0.1394**	-0.7551***
	(0.1626)	(0.0343)	(0.1707)	(0.2631)	(0.0687)	(0.2696)
N	101771	102362	101769	97768	102313	97758
y variable mean	15.72	0.34	16.04	44.47	1.76	46.04
R^2	0.300	0.116	0.303	0.356	0.183	0.373
	Panel C: Border FE with Controls and County-Lat./Lon.					
	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance to polling place	-0.5964***	0.0130	-0.5982***	-0.8939***	0.1810***	-0.7369***
	(0.1659)	(0.0295)	(0.1709)	(0.2668)	(0.0591)	(0.2707)
N	101771	102362	101769	97768	102313	97758
y variable mean	15.72	0.34	16.04	44.47	1.76	46.04
R^2	0.302	0.118	0.305	0.358	0.185	0.374

Note: Distance to polling place measured in miles. Turnout is measured as the number of votes per voting-age population (separately for votes cast at polling places, through absentee ballots, and total). County-Lat./Lon. refers to latitude and longitude controls, interacted with county fixed effects. All regressions include the following controls: percent registered Democrat, percent registered Republican, percent age 30-49, percent age 50-64, percent age 65 and up, percent female, population, voting age population, percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, percent with commute time less than 5 minutes, and percent with commute greater than 60 minutes. Standard errors clustered at the border level are reported in parentheses.

Table A.4: The effect of distance to polling place on turnout: Georgia

	Panel A: OLS with Precinct FE and Controls					
	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance to polling place	-0.3210***	0.1532	-0.1152	-1.1242***	1.1374***	0.0390
	(0.1170)	(0.0942)	(0.1226)	(0.1577)	(0.2234)	(0.2591)
N	35979	36167	35920	35428	35101	34693
y variable mean	13.29	5.41	18.52	24.84	27.76	52.17
R^2	0.465	0.298	0.581	0.264	0.315	0.319

	Panel B: Border FE with Controls					
	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance to polling place	-0.5599***	0.3484***	-0.2458	-1.6241***	1.1130***	-0.5348
	(0.1521)	(0.1137)	(0.1577)	(0.2043)	(0.2648)	(0.3326)
N	35979	36167	35920	35428	35101	34693
y variable mean	13.29	5.41	18.52	24.84	27.76	52.17
R^2	0.507	0.356	0.613	0.314	0.373	0.372

	Panel C: Border FE with Controls and County-Lat./Lon.					
	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance to polling place	-0.6100***	0.3889***	-0.2565	-1.7538***	1.1820***	-0.6250*
	(0.1603)	(0.1234)	(0.1734)	(0.2228)	(0.2627)	(0.3492)
N	35979	36167	35920	35428	35101	34693
y variable mean	13.29	5.41	18.52	24.84	27.76	52.17
R^2	0.514	0.366	0.618	0.323	0.386	0.382

Note: Distance to polling place measured in miles. Turnout is measured as the number of votes per voting-age population (separately for votes cast at polling places, through absentee ballots, and total). County-Lat./Lon. refers to latitude and longitude controls, interacted with county fixed effects. All regressions include the following controls: percent registered Democrat, percent registered Republican, percent age 30-49, percent age 50-64, percent age 65 and up, percent female, population, voting age population, percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, percent with commute time less than 5 minutes, and percent with commute greater than 60 minutes. Standard errors clustered at the border level are reported in parentheses.

A.3 Difference in Differences

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. We will explore two specifications. The first specification estimates the effect of a change in distance to the polling place using within-voter variation:

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

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:

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

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 $t-1$. 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 A.5 we report estimates for Equation A.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 A.6). 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 A.5: The effect of distance to polling place on turnout: Difference in Differences Estimation in Pennsylvania

	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Distance (miles)	0.0386	0.0090	0.0475	0.0554	0.0086	0.0640
	(0.0897)	(0.0102)	(0.0902)	(0.0863)	(0.0165)	(0.0830)
N	14406602	14406602	14406602	14504036	14504036	14504036
y variable mean	29.14	0.64	29.78	63.92	2.66	66.58
R^2	0.733	0.608	0.737	0.755	0.657	0.757

Note: Distance to polling place measured in miles. The dependent variables are indicators for whether or not a registered voter has voted at the polling place, through absentee ballot, or through either voting method. All regressions include Individual Fixed Effects and County by Year FE. Standard errors clustered at the precinct level are reported in parentheses.

Table A.6: Difference in Differences Estimates from Pennsylvania:The effect of polling place changes, voter location changes, and distance to polling place on turnout:

	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
PL Moved	-0.6455***	-0.0561**	-0.7016***	0.3856*	-0.0744*	0.3112
	(0.1900)	(0.0236)	(0.1931)	(0.2027)	(0.0409)	(0.1957)
Voter Moved	3.9846***	0.2704***	4.2550***	-2.9912**	0.6797***	-2.3115*
	(0.7635)	(0.0549)	(0.7769)	(1.2586)	(0.1243)	(1.3149)
PL Moved × Dist. (mi)	-0.1944**	0.0141	-0.1803*	-0.1283	-0.0068	-0.1351
	(0.0955)	(0.0108)	(0.0960)	(0.0892)	(0.0192)	(0.0855)
Voter Moved × Dist. (mi)	-0.9182	-0.1555*	-1.0738	-0.2751	-0.0029	-0.2780
	(0.7462)	(0.0828)	(0.7668)	(0.7409)	(0.1740)	(0.7536)
Voter Moved × PL Moved	-1.7024**	0.0155	-1.6870**	-0.2019	-0.0792	-0.2811
	(0.7905)	(0.0632)	(0.8040)	(1.2975)	(0.1371)	(1.3520)
Voter Moved × PL Moved	0.4685	0.1266	0.5951	-0.2272	0.0163	-0.2109
× Dist. (mi)	(0.7066)	(0.0863)	(0.7249)	(0.7186)	(0.1798)	(0.7311)
N	14406602	14406602	14406602	14504036	14504036	14504036
y variable mean	29.14	0.64	29.78	63.92	2.66	66.58
R^2	0.733	0.608	0.737	0.755	0.657	0.757

Note: All regressions include individual voter fixed effects and county-year fixed effects. The dependent variables are indicators for whether or not a registered voter has voted at the polling place, through absentee ballot, or through either voting method. For readability, we multiply the dependent variables by 100 so that the coefficients can be interpreted as percentage point changes in the likelihood of voting. Standard errors clustered at the precinct level are reported in parentheses.

A.4 Matching

In this section we match blocks that are nearest to each other along a precinct border, following the matching methods of Keele and Titiunik (2015) and Cantoni (2020). Specifically, we match blocks that have population centroids within a 0.2-mile buffer (1056 feet or 322 meters) of either side of a precinct border by minimizing the Euclidean distance between the two blocks, with replacement. The matched block pairs should be similar in terms of observed and unobserved characteristics relevant for turnout. The differences in distance to polling place between matched blocks stems only from the precinct border, and is thus plausibly exogenous. Using the restricted set of precinct borders that do not overlap with other important borders (as in the main analysis, Figure 2), there are 85,873 matched block pairs in Pennsylvania and 25,512 matched block pairs in Georgia.

We estimate the following equations using the matched sample:

$$turnout_{bp} = \delta_p + \beta distance_b + \mathcal{X}_b' \iota + \epsilon_b \tag{A.3}$$

where $turnout_{bp}$ denotes the block-level turnout of block b in matched pair p and δ_p denotes matched pair fixed-effects. Since we match our blocks with replacement, a single block b can be matched to more than one block and can consequently be part of multiple matched pairs. However, as noted previously in the literature (Dube et al. 2010 , Cantoni 2020), the repetition of the same unit in multiple pairs along a precinct border can induce a correlation in the residuals across block pairs and borders. To address this concern, we cluster standard errors two-ways by border and precinct.

Panels A and B Table A.7 report the value of β for all elections in our sample for Pennsylvania and Georgia respectively. When compared to the reported estimates of β from the block-level Border FE specification (Panel B of Appendix A.3), the direction and the order of magnitude of the coefficients from Panel A for Pennsylvania are strikingly similar. In particular, all of the coefficients lie within the 95 percent confidence interval of the coefficients reported for the Border FE specification. For Georgia, comparing the results from Appendix A.7 Panel B with estimates from Appendix A.4 Panel B paints a similar picture. Here, all of the coefficients for the general election lie within the 95 percent confidence interval of the coefficients reported for the Border FE specification, while the coefficients for the primary election are similar in their order of magnitude.

Table A.7: The effect of distance to polling place on turnout: Matched Pair FE

	Panel A: Pennsylvania					
	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Mean Dist. to polling location (mi)	-0.4834** (0.2246)	-0.0181 (0.0371)	-0.5103** (0.2261)	-0.9593*** (0.3473)	0.0842 (0.0784)	-0.8900** (0.3518)
N	171746	173888	171740	159220	173700	159202
y variable mean	15.40	0.35	15.72	43.57	1.74	45.10
R^2	0.602	0.506	0.603	0.648	0.537	0.657

	Panel B: Georgia					
	Primary Election			General Election		
	At Poll	Absentee	Total	At Poll	Absentee	Total
Mean Dist. to polling location (mi)	-1.0985*** (0.2227)	0.3430 (0.2459)	-0.7738** (0.3212)	-2.0246*** (0.2936)	0.7375** (0.3202)	-1.3081*** (0.4266)
N	51024	51600	50894	49588	48808	47736
y variable mean	13.39	5.24	18.43	24.98	26.94	51.49
R^2	0.574	0.543	0.573	0.594	0.599	0.612

Note: Distance to polling place measured in miles. Turnout is measured as the number of votes per voting-age population (separately for votes cast at polling places, through absentee ballots, and total). All regressions include Matched Pair Fixed Effects and the following controls: population, voting age population, percent registered Democrat (PA only), percent registered Republican (PA only), percent Black, percent Hispanic, median household income, percent without a high school diploma, percent that walk to work, and indicators for whether travel time to work is less than 5 minutes or greater than 60 minutes. Standard errors clustered at the border-precinct level are reported in parentheses.