The Effect of Moving Precincts on Voting

Tim Burke	Aakriti Kaul	Jake Mainwaring
School of Information	School of Information	School of Information
UC Berkeley	UC Berkeley	UC Berkeley
timothy.burke@berkeley.edu	aakritikaul@berkeley.edu	jmainwaring@berkeley.edu

Abstract

U.S politics has gained global traction with conversations around policy change, voting rights, and its impact on voting behavior. Past research suggests the costs experienced by voters influence can influence turnout, and that polling place changes could be one such cost. Our research aims to establish the causal impact of polling place changes on voter turnout and more broadly, study how that informs the future of electoral ergonomics. We investigate this effect using a geographic matching method that compares voters who changed polling places with sets of nearby voters who kept their polling place. Our findings suggest that polling place changes decrease turnout by just under 1 percent, with a larger effect as distance to the poll increases.

Keywords: Voting; Political Participation; Election Administration; Statistical Matching

Table of Contents

1. Introduction

Voter participation, based on a multitude of cost factors, is thoroughly discussed with respect to political behavior. Since the Voting Rights Act in 1965, a critical tool in combating racial discrimination in voting, areas with a history of imposing race-based burdens to voting were subject to federal oversight or "preclearance" when changing election rules. However, in the 2013 *Shelby County v. Holder* decision, the Supreme Court overturned this provision. This made voters vulnerable to constitutional amendments by state and local governments resulting in stricter photo ID requirements, reduced early voting sites and elimination of same day registration as few of many mechanisms which can be leveraged for partisan advantage. VICE News[1] reported that in the years following the Shelby decision, on average, these newly 'freed' jurisdictions shut down 20 percent more polling stations than others in the country which were not formerly scrutinized.

While public debates continue around changes to voter ID requirements and polling place closures hampering the ability of registered voters to cast their vote, the innate challenge of changing or re-allocating a polling location is not as widely-covered by the media. However, a great deal of scholarship in political science and economics has been dedicated to understanding the costs of voting

[1] https://news.vice.com/en_us/article/kz58qx/how-the-gutting-of-the-voting-rights-act-led-to-closed-polls

and how certain costs can affect voter turnout - which includes polling place changes and closures. Polling place closures contribute to long wait times during voting periods, knowledge of which may make voters apathetic and deter them from showing up to vote. Officials often justify poll closure decisions by citing accessibility issues, expenses or underutilization of locations, although these actions which may well be politically motivated to increase the costs experienced by minorities and supporters of the other party. Such changes are not subject to federal oversight after the Shelby decision, except by new challenges through the courts.

In a seminal paper, Riker and Ordershook (1968)[2] model the individual decision to vote as a rational choice based on the expected benefits and costs of voting. They argue that this does not just consider the expected benefit of a voter's desired outcome and the probability that a single vote will change the outcome, but also the costs of voting and the benefit a voter derives from the act of voting. A polling place change could constitute a material cost for in-person voters, both as a search cost to find the new polling location, a transportation cost if the new location is farther away, and a cognitive cost of no reward for action. This creates a testable hypothesis that voters will turn out at lower rates if their polling place is moved.

[2]Riker, William H., and Peter C. Ordeshook. "A Theory of the Calculus of Voting." American political science review 62, no. 1 (1968): 25-42.

2. Literature Review

Numerous studies have explored if polling place changes affect voter turnout, with results suggesting that polling changes can have both positive and negative effects on turnout depending on the circumstances. Haspel and Knotts (2005)[3] predict voter behavior in the 2001 mayoral race in Atlanta by considering distance to polling place, whether or not the polling place changed, and covariates for voter demographic characteristics. Their logistic regression model finds that the distance to polling place does have a negative effect on turnout, but changing polling places is correlated with a higher propensity to vote. This is likely because Atlanta increased the number of polling places (reducing distances for many voters) and sent letters about the change (effectively reminding them about the election). Brady and McNulty (2011)[4] find the opposite result during the consolidation of polling places for the 2003 Los Angeles gubernatorial recall election. They demonstrate that polling places were reassigned as-if-randomly and the election can be treated as a natural experiment. Using a statistical matching model, they estimate that voters were slightly less likely to turn out if their polling places were changed, with a marginally larger effect on Democratic

[3] Haspel, Moshe, and H. Gibbs Knotts. "Location, location, location: Precinct placement and the costs of voting." The Journal of Politics 67.2 (2005): 560-573.
[4]Brady, H. E., & McNulty, J. E. (2011). Turning out to vote: The costs of finding and getting to the polling place. American Political Science Review, 105(1), 115-134.

voters. Taken together, these studies support the hypothesis that changes in polling places can affect turnout.

However, neither of the above studies addresses the more concerning cases where polling places are closed in a nonrandom manner for partisan gains. Amos, Smith and Ste. Claire (2017)[5] examine the controversial polling place closures in Florida's Manatee County during the 2014 General Election, when an outspokenly partisan Republican Supervisor of Elections closed 38% of the polling places. After showing that minorities and Democrats were more likely to be re-assigned, they use a multinomial logistic regression to find that voters are significantly less likely to vote in-person if their polling place changed. Clinton, Eubank, Fresh and Shepherd (2019)[6] examined the within-voter effect of changes in polling places across the 2008, 2012, and 2016 presidential elections in North Carolina, a state with a history of voter suppression. Using an OLS model with individual fixed effects, they find evidence of a significant reduction in in-person voting both due to the search costs and the increased travel costs associated with polling place changes. However, this effect appears to be offset by voters switching to early voting.

Overall, the potential negative effects of polling place closures are well-established in the literature. Nevertheless, the universe of studies in this field

[5]Amos, B., Smith, D. A., & Claire, C. S. (2017). Reprecincting and voting behavior. Political Behavior, 39(1), 133-156.
[6] Clinton, J. D., Eubank, N., Fresh, A., & Shepherd, M. E. Polling Place Changes and Political Participation: Evidence from North Carolina Presidential Elections, 2008-2016.

with strongly identified causal mechanisms is far sparser. Following the lead of Amos et al and Eubank et al, we examine changes in polling places and distances to the polls to estimate the causal impact of poll closures. We focus on the state of North Carolina as an instructive example - the state was previously covered by the Voting Rights Act and has experienced recent controversies about partisan interference in elections. Without any clear natural experiment, we instead seek to compare voters who are similar in all respects except for whether or not their polling place changed. We use geographic proximity as a measure of similarity - voters who live "close" to one another should be should, in aggregate, be similar in key respects. By comparing voters who had a polling place change to nearby ones who did not, we hope to create a causal estimate of the effect of the polling change.

Keele and Titinuik (2014)[7] present an interesting approach to this problem in their study of geographic boundaries in regression discontinuity models. Their work closely mirrors our proposed approach to estimate geographically located effects while controlling for compound treatments, incorporating geographic distance metrics that can overcome the drawbacks of one-dimensional "naive" distance and investigating spatial variation around the discontinuity point. Their framework informs our approach in multiple ways. However, given the difference

[7] Keele, Luke J., and Rocio Titiunik. (2015). "Geographic Boundaries as Regression Discontinuities." Political Analysis 23 (1): 127–155.

in our core definition of a geographical boundary, our proposed methodology incorporates excerpts of their GRD design learnings.

Our design is primarily informed by a paper by Stuart (2009),[8] who details a number of matching methods for causal inference. Our approach most closely resembles a combination of “Variable Ratio Matching” and “Caliper Matching”. Variable Ratio Matching compares each treated unit to a variable number of control group units. This lends itself well to our research problem because some treated individuals live near more “control” voters (those with the same polling location) than others. However, this method also complicates the analysis. Ming and Rosenbaum (2001)[9] go into more detail on variable ratio matching calculations. Caliper Matching[10] compares treatment units to control units that are within some predefined width of the treatment unit’s propensity score. Our model employs geographic location defined by latitude and longitude as a propensity score, rather than using some other continuous covariates.

One other methodological consideration for our research is sampling with or without replacement. Each treated voter is compared to a number of control units. In our design, it’s possible that the same voter could be included in multiple control sets if their polling place did not change but they live near multiple people

---

[8] Stuart E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science : a review journal of the Institute of Mathematical Statistics*, *25*(1), 1-21.
[9] Ming, K., Rosenbaum, P. (2001) A Note on Optimal Matching With Variable Controls Using the Assignment Algorithm, Journal of Computational and Graphical Statistics, 10:3, 455-463
[10] Rubin, D. B., & Thomas, N. (2000). Combining propensity score matching with additional adjustments for prognostic covariates. Journal of the American Statistical Association, 95(450), 573-585.

whose did. Stuart notes that this type of matching with replacement makes inference more complicated (see also: Dehejia and Wahba[11]).

3. Data on Voters, Voter History, and Polling Places

To identify individual voter patterns within a neighborhood, with respect to voting activity being affected by a polling place change, we collect individual-level data on every registered voter in North Carolina. This data is made accessible, free of cost, by the North Carolina State Board of Election (NCSBE) at www.ncsbe.gov. Records available to download begin in 2005. Voter data is linked to an individual via NCID - a unique identifier that provides identity management and access control to North Carolina state-owned resources. A given voter cannot be assigned multiple NCIDs, except in the edge case that they move out of state and then back in between elections. Voters who are removed from the voter rolls also can have their NCID reassigned to another voter. The data contain few duplicate NCIDs, and those that are found are dropped from our analysis. While the data from these electoral snapshots is anonymized, it is detailed enough to include factors such as address, voter status, and demographic information on race, ethnicity, age and gender. In addition to voter information, the NCSBE tracks voter history, a record of every vote cast by a given

[11] Dehejia, R. H., & Wahba, S. (1999). Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. Journal of the American Statistical Association, 94(448), 1053-1062. doi:10.1080/01621459.1999.10473858

NCID, and polling location information. We supplement the voter information files with voter history to learn the voter turnout rate, as well as polling location information to determine the distance of a given voter from his or her polling place. As a departure from existing literature, we focus our analysis not on a panel design to compare the same voter across years, but on a comparison of treatment voters in the 2016 general election with geographically proximate control voters. The full details of the model are discussed in Section 4.

Voter Information

The NCSBE maintains snapshots every few months with data on each registered North Carolina voter with information including age, race, gender, home address and voter registration status. The 2012 snapshot contains 11,352,660 rows of voter records, while the 2016 snapshot contains 7,449,896 records. These numbers include voters who were recently removed from the voter rolls - such voters are not assigned a polling place and do not always contain address information, as their voter registration was removed for reasons such as having an inaccurate address or passing away. Such voters are removed from our analysis.

Voter History

The voter history file holds a row record for each vote that was cast across elections in the given year and past years. Information includes the election description, the party the voter was registered as and the method of voting (such as absentee or in-person), county name, and precinct name. We supplement our final dataframe by performing a left join of these details onto the voter information file, matching records using the NCID. Limiting to the 2016 election and NCIDs that are not duplicated, this file contains 4,767,975 records.

Polling Place Listings

Election polling place listings made available by NCSBE include data on election date, county identification number, county name, polling place identification number and name, precinct name and street address. A given voter's polling place can be uniquely identified by the election year, county and precinct. The 2012 and 2016 polling place files contain a combined 5,412 unique polling locations, which we use to identify the location of each voter's polling place in a given election year. We noted that there is a mismatch between the 2012 voter information records and the 2012 polling place listings - we identified 131 precinct-county combinations that existed in the voter information file but did not exist in the polling place records for 2012. All such mismatched pairs were found in the 2011 polling information,[12] suggesting that legacy names remained in the data. We used the addresses given in the 2011 data for these cases.

[12] NCSBE public election data: https://dl.ncsbe.gov/?prefix=ENRS/2011_11_08/

3.1. Data Filtering and Processing

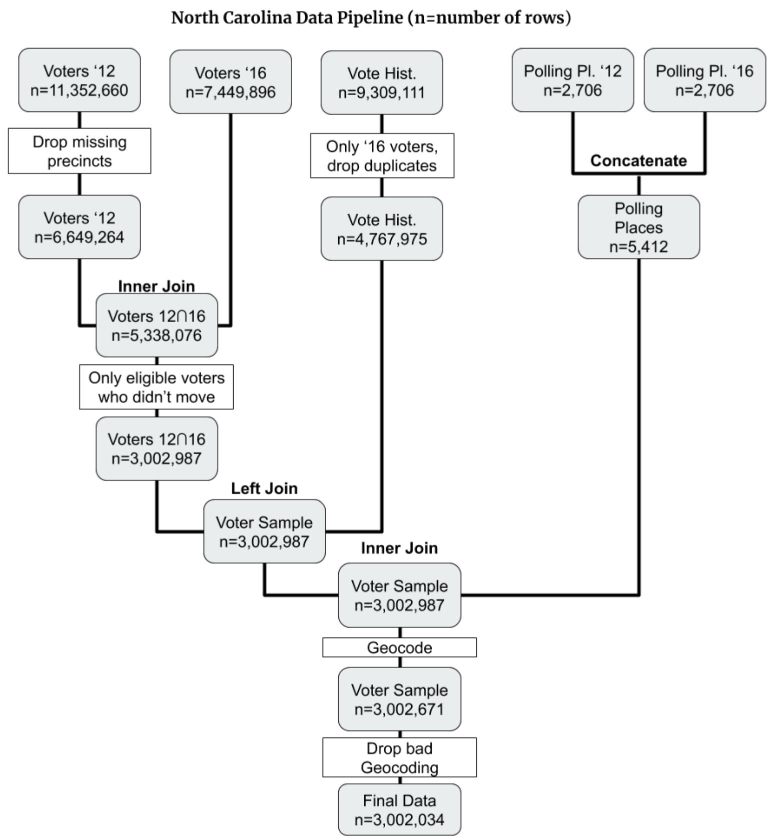

Figure 1: Data Pipeline

Our process begins with the 2012 and 2016 voter information files. We drop any voters who are missing precinct information because we cannot identify the polling locations for these voters and therefore cannot assign them to treatment or control. Cases of missing precincts are almost entirely voters who have been removed from the rolls, although there are a small number of cases of missing data as well. We then inner join the 2012 and 2016 voters so that we can identify which voters might have moved between elections, which would cause a voluntary change in polling place. We removed these cases because the polling place change is voluntary and therefore the causality of any related change in turnout is difficult to establish.

Voters who moved are identified by checking if the house number in their street address changed. Voters who moved are dropped from the data, as are voters who are ineligible to vote. Eligibility is defined as having an active voter status or inactive for the reason of not confirming their address. These are the same eligibility criteria used by the state of North Carolina in determining turnout statistics. We choose not to filter eligibility before joining the 2016 and 2012 data because it is possible that a voter who was ineligible in 2012 may be eligible again in 2016, and we did not want to lose these voters from our sample.

We then augment the voter data with voter history information. Voter history is filtered down to votes cast in the 2016 presidential election, and duplicate NCIDs that are found in the data are dropped. From there, we left join the voter history onto the voter information. Voters who appear in voter history cast a vote in the 2016 election, while those who do not did not vote. This gives us the voting outcome for each voter in our sample. Finally, the combined voter information and voter history is inner joined with the polling place information for each voter. Voter polling place is uniquely identified by the election year, county and precinct for each voter.

Following the joining of all data, we need to identify the geographic location of each voter and polling place address in terms of latitude and longitude. Geocoding addresses is done using the Here API.[13] A small subset of addresses was done using the Google Maps Geolocation API, but the pricing was infeasible for a project of our scale. Following the completion of geocoding, a very small set of addresses could not be found by the API, and were therefore discarded from our analysis.

To identify the treatment and control groups in our study, we took the distance of each voter from their precinct in 2012 and their precinct in 2016. The difference in these two distances is used to determine treatment assignment - voters who had zero distance between the 2012 and 2016 locations are control voters, and those that had nonzero distances are treatment.

Some outliers existed after geocoding - hundreds of voters had apparent changes in polling place distance of a hundred or more miles, which does not appear reasonable. A deeper examination of these cases suggested that the voter addresses were incorrectly geocoded. Based on the distribution of voter distances, we defined outliers to be those who had a polling place distance change of greater than 12 miles, thus dropping 489 observations from our dataset. The model is not sensitive to the exact cutoff we use for outliers.

[13] See: developer.here.com/Geocoding/API

3.2. Comparison of Sample Data and Voter Rolls

Statistic	2016 Voter Rolls	Our Sample
Number of Eligible Voters	6,251,529	3,002,034
Voter Turnout	72.3%	85.6%
Voted as Registered Democrat	28.3%	35.5%
Voted as Registered Republican	24.0%	30.4%
Mean Age	49.3	57.8
Female	52.6%	54.1%
White	68.7%	74.0%
Non-White	31.3%	26.0%

Figure 2: Sample and Voter Roll Statistics

Due to the data processing steps we take, the sample data we use to evaluate our model contains some important differences from the North Carolina voter rolls as a whole. First, only about half of the voters in the 2016 voter rolls are included in our sample. The vast majority of the difference is due to the fact that we only include voters who were in the 2012 voter rolls and did not move addresses in between the elections. We use the 2012 general election rather than the more recent 2014 election primarily because we were interested in capturing the effects of polling place changes after the 2013 Shelby v. Holder Decision. When the federal restrictions were lifted and North Carolina was more free to change polling locations, we expect that these changes would have been made prior to the 2014 midterms. This would be especially true if the changes were made for partisan advantage.

We needed voters to appear in both the 2012 and 2016 election records so that we could identify whether or not polling places changed between years. For the same reason, we removed voters who were missing the information needed to identify their polling place. We dropped voters that moved addresses because they raise issues with the causal inference question our model addresses. Our reasoning is that voters who moved addresses will naturally move polling places, except for them it is a choice rather than the result of a decision by an election administrator. Our model is not causally well-identified if the voter moved polling places by choice - the change in polling place may affect their decision to vote, but so could other unobserved factors that caused them to move addresses in the first place. Therefore these voters could not be included in the analysis.

In addition, we also dropped some voters from our sample due to cases of bad data that could not be resolved. A very small number of duplicated NCID values had to be dropped because it was impossible to identify the correct records and voting outcome for these voters. If voter addresses could not be geocoded, they had to be removed as it was impossible to measure distances from them to other voters. Obviously incorrect geocoding also had to be dropped because the distances between them and their neighbors would be meaningless.

We are confident that our sample is the largest possible subset of the voter rolls for which our model can be successfully applied. Nonetheless, there are meaningful differences between the voter rolls and our sample. Figure 2 illustrates these issues. Our sample has a substantially older mean age, higher turnout and a whiter population. Our data filtering steps help explain the difference. First, the choice to compare voters to 2012 data requires that they had been registered in both 2012 and 2016. This automatically filters out any voters who are under the age of 22 in 2016. Not moving addresses between years is also likely to skew towards the older population - younger people are generally less likely to own a home and more likely to move for work or school reasons. The higher voter of our sample may also be a function of age - in aggregate, voter turnout is higher for older voters than it is for young ones. The difference in racial demographics may also be related both to the age of the sample and the requirement that voters not move addresses.

While we believe that our model is still internally valid with the sample of voters we use, it does limit our external validity. For example, it is possible that younger voters do not respond to the costs of polling place changes in the same way that older voters in our sample do. We address considerations with validity in Section 6.

3.3. Statistics and Comparison of Treatment and Control Groups

Statistic	Farther Poll	Closer Poll	Changed Poll	Same Poll
Number of Eligible Voters	88,936	57,673	146,609	2,855,425
Voter Turnout	85.0%	84.9%	85.0%	85.6%
Voted as Registered Democrat	34.5%	32.3%	33.7%	35.6%
Voted as Registered Republican	32.5%	34.0%	33.1%	30.3%
Mean Age	58.1	58.2	58.2	57.8
Female	54.3%	54.1%	54.2%	54.1%
White	74.3%	76.9%	75.3%	73.9%
Non-White	25.7%	23.1%	24.7%	26.1%

Figure 3: Treatment and Control Statistics

In our study, “treatment” is defined as voters who have been assigned a new polling place between the 2012 and 2016 general elections. We assigned treatment as all voters who have a nonzero distance from their 2012 polling address coordinates and their 2016 address coordinates. All other voters in our sample are assigned to the control group. In the table above, “Farther Poll” and “Closer Poll” are subsets of the full treatment group “Changed Poll.”

The treatment and control groups are similar in most respects. In aggregate, the turnout rates between the two groups are nearly identical, and both groups have similar demographics by gender, race and party registration. While poll closures and location changes are not randomly assigned, it does appear that the voters exposed to poll changes were reasonably representative of the voter population as a whole.

Within the treatment group, voters whose polls were moved farther away are also reasonably similar to those who moved closer. The only noticeable difference is in the racial and party demographics - in aggregate, voters who had their poll moved closer to them are slightly more likely to be white and Republican. However, none of these differences are substantial in aggregate.

A deeper dive into the way voters cast their votes reveals some differences between the treatment and control groups. Figure 4 shows voter turnout overall, as well as breaking it out into absentee voting (including early voting and mail-in ballots) and in-person voting. We see that the treatment group has a marginally lower turnout overall and noticeably lower in-person turnout, but have a higher rate of absentee voting than the control group. Eubanks et al[14] noted that having a polling place moved caused voters to substitute absentee voting for in-person voting, which is consistent with this statistic. However, we cannot be certain that the change in polling places caused the observed difference between the treatment and control group voting patterns. We address the impact of polling place changes on absentee voting in Section 6.

[14] Clinton, J. D., Eubank, N., Fresh, A., & Shepherd, M. E., 2019

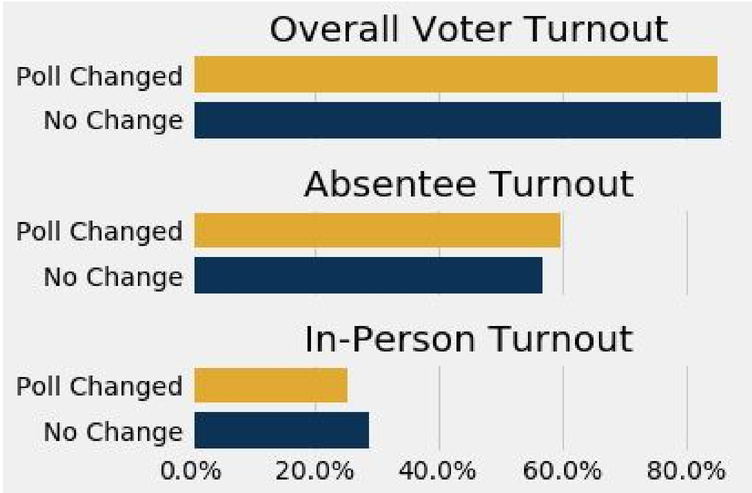

Figure 4: Voter Turnout vs. Different Voting Methods

Figure 5 below addresses the distribution of the change in polling place distance among the treatment group. The change in treatment voters' distance to their polling place is evenly distributed around zero, with a mean of 0.02 miles. The majority of changes in distance are relatively small, and 56% of the treatment group had their polling place moved less than one mile.

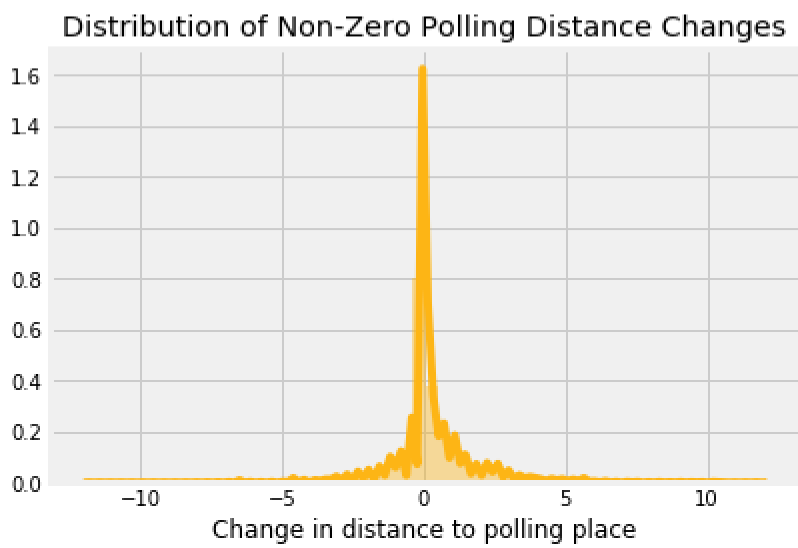

Figure 5: Distribution of Polling Distance

As the long tails of the distribution show, there are voters who had more substantial changes in polling place. Indeed, 2.5% of the sample voters had a change in polling location of greater than five miles. As discussed in Section 3.1, outliers who had a polling distance change of greater than 12 miles were dropped from the sample.

4. Methodology

The gold standard for causal questions is the RCT, or randomized control trial, where individuals are randomly assigned to one of a few conditions. True random assignment ensures that individuals are similar across all attributes except for their treatment assignment, which suggests that a difference in outcomes was caused by the treatment itself. Of course, elections are one of many situations where a randomized trial is neither feasible nor ethical. In this case, we should still try and create a treatment/control set that are similar but for the treatment assignment. To do so, we use geographic proximity as a proxy for likeness.

The fundamental assumption of our methodology is that individuals tend to be similar to their neighbors across attributes that affect voter turnout. As an illustrative example, consider two neighbors who live across the street from one another. Perhaps this road was the dividing line for reprecincting, where those who live north of the road were assigned to a new polling place and those who live below the line were not. Assuming these two neighbors are similar on relevant attributes, we can attribute a difference in turnout to the fact that one was assigned to a new polling place.

Rather than this approach, why not simply try and control for the factors above, like race and age? First, our data on each registered voter is limited. We have high-level information on registered voters, but there are many relevant attributes like income or mode of transportation that North Carolina does not provide. More importantly, we are concerned about omitted variable bias. There is a wide range of hard-to-measure variables that determine whether someone will vote in a given election. This could be attitudes toward the candidates, the particular economic climate, the cost of taking off two hours from work on Election Day, and many others. Even with rich voter data, there will inevitably be variables we miss, which limits our ability to make a true apples-to-apples comparison. This is one of the key advantages of our methodology. Instead of attempting to control for everything, we rely on the assumption that, in aggregate, individuals will be similar to their neighbors across measurable and unmeasurable traits.

More specifically, we build control sets for each “treated” individual as defined in Section 3. Figure 6 below provides a visual overview of how these control groups are formed:

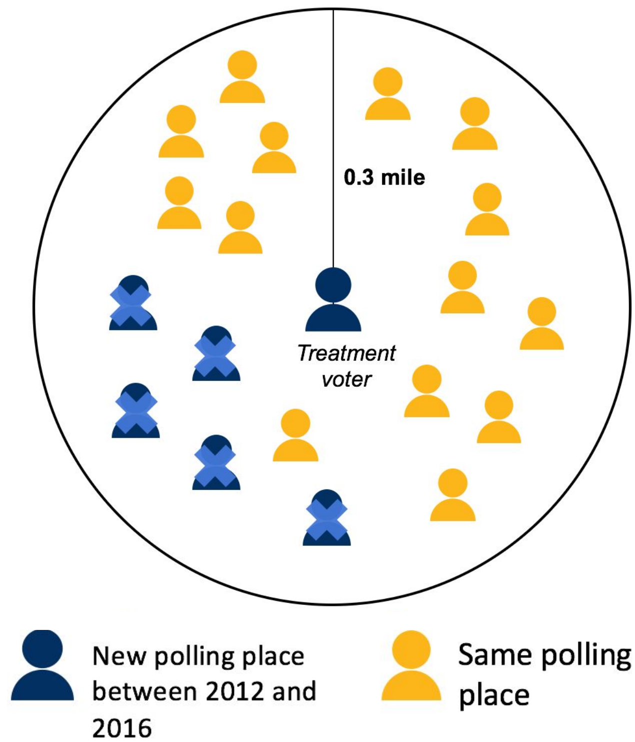

Figure 6: Treatment vs Control formation

Consider a hypothetical voter who was reassigned to a new polling place in 2015. After identifying her as a member of the treatment group, we use geolocation data to best identify the other registered voters who live within a 0.3 mile radius. Distance between voters is defined as the Haversine distance between the latitude and longitude coordinates of each voter's registered address.

The 0.3 mile size of the distance caliper around a given voter is a hyperparameter of our model. To select this hyperparameter, we randomly sampled 70% of treatment voters as "training data" and measured the treatment effect at different caliper sizes. We believed this to be the best way to select the hyperparameter without repeatedly testing our model until we found significance on the whole dataset. If we found a significant effect on the training data that did not generalize to the whole dataset, it would imply that our model was "overfitting" on the sample and not valid. We found that few voters had any neighbors within a small radius such as 0.1 miles, resulting in very large standard errors. The actual effect size of changing polling places declines as the distance caliper is increased, reaching effectively zero at a distance caliper of one mile. We chose 0.3 miles as the best balance of support and effect size, although the treatment effect is close in magnitude and significance at similar caliper sizes.

After finding the set of "neighbors" near our hypothetical voter, we exclude everyone in that radius who was also assigned to a new polling place. Our goal is to create a treatment/control setup, so including other treatment voters in the control set would blur the relationship. At this point, we now have a treatment voter and her control set, which can be seen in Figure 7.

Figure 7: Variable control set per treatment voter

The size of each treatment voter's control set can vary. Many treatment voters do not have *any* control voters within 0.3 miles. For someone who lives in the middle of a region that was reassigned, this is unsurprising - everyone else nearby may have also been reassigned. In this case, the treatment voter does not end up in our analysis, as we cannot construct a reasonable treatment-control setup for him. Of the approximately 135,000 individuals from our sample who received a new polling place assignment, 27,020 of them have at least one "control" individual within 0.3 miles. The median size of these treatment voters' control sets is 11, and the mean is 15.6.

To conduct the analysis, we first used the model described in Equation 1 below:

Equation 1

$$Pr(Vote_i) = \beta_0 + \beta_1 Poll\_Changed_i + \gamma Block + \epsilon$$

This is the simplest specification of our model: β_0 is an intercept term, *Poll_Changed* is a binary variable for whether or not a voter had their poll moved, and we include fixed effects at the *Block* level (every treatment voter and corresponding set of control voters make up a unique block). Each row in the table is a voter who is either in treatment or control, and our outcome variable, $Vote_i$, indicates whether she cast a vote in the 2016 general election.

For additional insight, we split the *Poll_Changed* into two coefficients, one for having polling place moved closer and one for having it moved farther away.

Further, because we know approximately 50% fall into each category, we can deduce the overall treatment of reassignment by taking the average of these two coefficients. Equation 2 describes this model specification.

Equation 2

$$Pr(Vote_i) = \beta_0 + \beta_1 Poll\_Further_i + \beta_2 Poll\_Closer_i + \gamma Block + \epsilon$$

In this model, β_0 is an intercept, β_1 is the coefficient for *Poll_Further* which is a binary variable for whether one's polling assignment was moved farther away. Similarly, β_2 is the coefficient for *Poll_Closer*, a binary variable for whether one's polling assignment was moved closer.

Finally, we examine the effect of changing polling location on absentee voting and in-person voting. Equation 3 and 4 specify these models. All coefficients are identical to those used in Equation 2, with the exception of the outcome variable *Voted_Absentee*, which represents whether the voter cast an absentee ballot, and *Voted_In_Person*, which represents voters who cast ballots in-person. The combination of these two variables covers all ballots cast by voters.

Equations 3 and 4

$$Pr(Vote\_Absentee_i) = \beta_0 + \beta_1 Poll\_Farther_i + \beta_2 Poll\_Closer_i + \gamma Block + \epsilon$$
$$Pr(Vote\_In\_Person_i) = \beta_0 + \beta_1 Poll\_Farther_i + \beta_2 Poll\_Closer_i + \gamma Block + \epsilon$$

A potential criticism of our methodology is that neighbors may not be similar, creating treatment/control groups that are actually different across important attributes and calling into question the causal nature of our work. Of course, this may very well be true for a particular set of neighbors, but we believe this concern as a whole is unfounded. As shown in Figure 3, most broad covariates are similar between groups. One robustness check for the geographical proximity decision we selected was running the same analysis as below, but on the 2012 election instead of 2016. The results of this check are covered in Section 5 and Section 6.

Before discussing the results, it is important to point out a few legitimate limitations of our approach. First and foremost, our sample is not perfectly representative of North Carolina registered voters. As seen in Figure 2, this creates a sample that skews older and whiter than the broader 2016 voter rolls. This is largely a byproduct of the types of people who tend to move residences during a four-year time period. On top of that, individuals under the age of 18 in 2012 are automatically excluded because it is impossible to reassigned between 2012 and 2016 if one was not eligible to vote in 2012. These differences between our sample and the electorate also explain why our overall voter turnout numbers in 2016 are higher than what was observed in the state more broadly.

The unrepresentative nature of our sample is more of a concern for generalizability than for internal validity. We still believe our causal identifier between treatment and control is valid. However, it is possible that the effect of reassignment is heterogeneous across certain groups. For example, being assigned to a farther poll might place more of a burden on minority voters, who are less likely to own a car, than on white voters.[15] In this case, our analysis might underestimate the true treatment effect, since our sample tends to skew whiter and white voters could be affected by reassignment.

Another limitation is that our data is limited to North Carolina. Ideally, we would conduct this analysis across a broader set of states with different characteristics and see the circumstances under which the treatment effect varies. However, it can be costly to get a state's voter data, and voter data from different states does not follow a common format. Because of these challenges, we opted to focus our analysis on North Carolina alone.

5. Results

Overall, we estimate that a change in polling place location decreases voter turnout by just under 1%. This varies by whether the new polling place is farther or closer than one's previous assignment, as well as the added distance an individual must travel to vote.

[15] https://nationalequityatlas.org/indicators/Car_access

Full set of voters

	Dependent variable:
	Voted
Polling location changed	-0.008^{***}
	(0.002)
Fixed effects by block?	Yes
Observations	420,791
Adjusted R^2	0.092
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 1: Turnout by change in distance to poll - change detected

When considering our full dataset, the estimated effect of new polling place assignment is -0.8%. All else equal, this means that being reassigned to a different polling location between 2012 and 2016 leads to a 0.8 percentage point lower voter turnout. This is directionally consistent with previous studies - having to find one's new polling place is an added mental burden that can decrease one's propensity to vote.

As we saw in section 3, the distribution around polling place distance change is centered at zero. In other words, there were a similar number of new polling assignments that *shortened* one's distance to vote as there were assignments that *increased* distance. This raises an important question: does turnout vary based on whether one's new polling place is closer or farther away? If not, this suggests that much of the 0.8% from above can be explained by search costs - the mental burden of determining where to vote - rather than the distance one must travel to do so. To study this, we split the individuals from above into groups based on whether their new polling place is farther or closer than their previous one. This can be seen in the model below:

	Dependent variable:
	Voted
Closer polling location	-0.008^{**}
	(0.003)
Farther polling location	-0.008^{***}
	(0.003)
Fixed effects by block?	Yes
Observations	420,791
Adjusted R^2	0.092
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 2: Turnout by change in distance to poll - Closer vs. Farther change

We can see that impact does not appear to vary by distance. Indeed, the two subsets experience an identical change in turnout of 0.8 percentage points. This is a somewhat surprising result - we would expect a farther polling place to decrease turnout more so than a closer one. Interpreted alone, this might suggest that the search cost of finding a new polling place is the primary driver of decreased turnout. However, one should also consider that most changes in polling place distance are quite small, and the effect of larger changes in distance could be more pronounced.

Farther poll changes

We know that for most treatment individuals, the change in polling place was small. 56% of them were reassigned to a new polling place within a mile of their previous one. In this case, reassignment may have a relatively small effect on turnout. For example, if someone is traveling by car, an additional 0.2 mile drive will likely add little travel time or cost, and it is hard to imagine this being the factor that ultimately leads someone to stay home on Election Day.

This may not be true when a new polling place is substantially farther away. In this case, the costs to vote rise, particularly for voters without a vehicle. We wanted to take a closer look at our treatment individuals with farther reassignments and see whether differential impact is present. To do so, we removed all treatment individuals whose reassignment changed by less than a mile in either direction. Of the 27,020 treatment individuals in our study, we were left with 9,700 whose new polling place was at least 1 mile closer or farther than their previous assignment.

We then ran the two previous models on this reduced dataset. When only including the "Polling location changed" feature, results were comparable. Rather than the -0.8% point estimate from above, we saw a -0.9% estimate (*SE*=0.004, *p<0.05)*. However, the results diverged once we split between individuals with closer and farther polling locations:

	Dependent variable:
	Voted
Closer polling location	−0.00000
	(0.006)
Farther polling location	−0.016***
	(0.005)
Fixed effects by block?	Yes
Observations	170,740
Adjusted R^2	0.044
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 3: Turnout by change in distance to poll over 1 mile

Results are visualized in Figure 8 below. For those whose new location was closer to them by at least one mile, there was no detectable treatment effect. However, for the individuals whose poll was moved away by at least one mile, this led to a -1.6% decrease in turnout. The results are directionally similar to the same model on the full dataset, but the magnitude is substantially greater. This provides further evidence that distance to the polls, in addition to search costs, can explain the decrease in voter turnout.

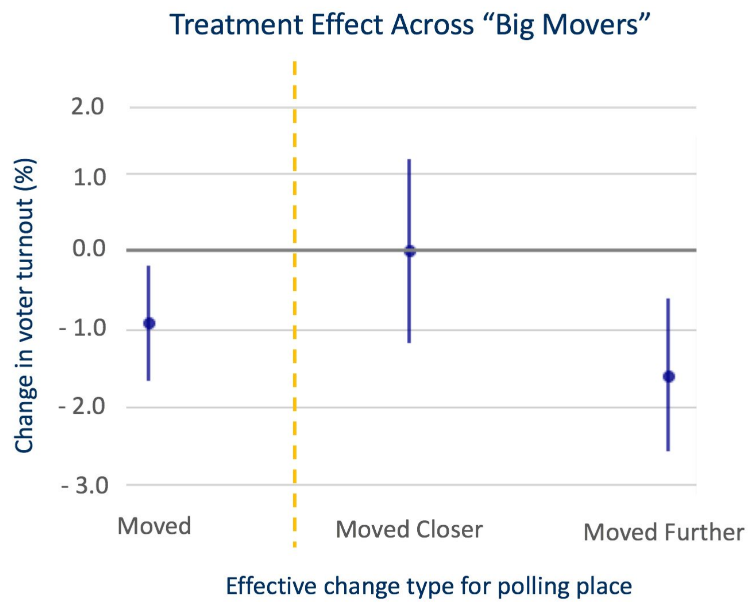

Figure 8: Treatment Effect for Voters with a Large Polling Place Move

Absentee and In-Person Voting

Eubanks et al[16] found that having a polling place change caused voters to change voting methods, essentially substituting in-person voting for absentee voting. This could be a response to the search and travel costs of finding a new polling location and traveling there on election day - it may be more convenient to go to an early voting location or mail in a ballot instead. We examine this question using the models defined in Equation 3 and 4, which examine the effect of poll changes on casting absentee ballots. Results are contained in Table 4 and 5 below.

[16] Clinton, J. D., Eubank, N., Fresh, A., & Shepherd, M. E., 2019

Table 4: Absentee voting by change in distance to poll

	Dependent variable:
	Voted
Closer polling location	0.020***
	(0.005)
Farther polling location	0.017***
	(0.004)
Fixed effects by block?	Yes
Observations	420,791
Adjusted R^2	0.078
Note:	*p<0.1; **p<0.05; ***p<0.01

Table 5: In-person voting by change in distance to poll

	Dependent variable:
	Voted
Closer polling location	−0.028***
	(0.004)
Farther polling location	−0.025***
	(0.004)
Fixed effects by block?	Yes
Observations	420,791
Adjusted R^2	0.070
Note:	*p<0.1; **p<0.05; ***p<0.01

Tables 4 and 5: Absentee and In-Person Voting Models

Changing polling location does appear to have a significant effect on absentee and in-person voting. However, the effect is significant with respect to in-person voting. This is consistent with the findings of Eubanks et al, who see that polling place changes result in voters turning out less in person but potentially substituting it with absentee voting. This effect persists whether the polling place was changed to be closer or further.

Robustness Between 2012 and 2016

To confirm the robustness of our model, we repeat the same model described in Equation 2 on the dependent variable of 2012 voter outcome. This is shown below in Tables 6 and 7. If voters are generally similar to their nearby neighbors in all respects except for having a polling place change in 2016, then they should be effectively the same in prior to 2016. This would result in an insignificant treatment effect. However, our robustness check finds a highly significant effect. This is most likely explained by differential assignment to treatment and control caused by handling ambiguities in the 2012 polling place information. We discuss the implications of this finding in Section 6.

Tables 6 and 7: Robustness Check Using 2012 Turnout

Table 6: 2012 turnout by post-2012 poll change

	Dependent variable:
	Voted
Polling location changed	-0.041^{***}
	(0.004)
Fixed effects by block?	Yes
Observations	55,714
Adjusted R^2	0.059
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

Table 7: 2012 turnout by post-2012 poll change

	Dependent variable:
	Voted
Closer polling location	-0.056^{***}
	(0.006)
Farther polling location	-0.022^{***}
	(0.006)
Fixed effects by block?	Yes
Observations	55,714
Adjusted R^2	0.060
Note:	$^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

6. Discussion

Despite using a new methodology, our findings are generally consistent with prior work. First, being reassigned to a new polling place appears to have a significant impact on turnout across the board. In aggregate, this effect appears consistent whether or not the polling place moved closer or further. However, In the event that one's new polling place is at least a mile closer, there is no detectable effect on turnout, while if distance to polling place changes by a mile or more, the impact is estimated to be -1.6%.

These findings are particularly illuminating when considering search costs and travel costs. Prior work has established the idea of a search cost in finding one's new polling place. North Carolina's poll changes between 2012 and 2016 provide an interesting venue for studying search cost, as the number of closer and farther poll changes are approximately equal. Therefore, we can attribute much of the 0.8% decrease in turnout from Table 1 to search costs, because on average, voters are not traveling farther than they were in 2012. However, when we study voters who saw a more significant poll change, the effects are either 0% or -1.6%, depending on whether it became closer or farther. In the first case, this suggests that the shorter distance to the polls is potentially offset by the search cost of finding a new location. For the latter group, the additional mile of travel only exacerbates an already existing cost, leading to an even larger treatment effect.

It is also interesting to see that the effect of polling place changes is significant for in-person voting and absentee voting. This is consistent with the findings in other research. The reason behind the effect is open to debate, but may be a response to the search and travel costs imposed by a change in polling places. Early voting or mail-in ballots offer an alternative and potentially less costly voting option when compared to going in-person to a new polling place on election day.

Finally, the significance of results in 2012 (prior to the change in polling place occuring) raises the possibility that there are problems in the assignment of treatment and control due to ambiguities from missing data. Because 2012 polling place location information did not provide polling places for all voters in the 2012 voter rolls, we found matching locations to the missing values in 2011. It is possible that these locations were not all accurate. If bad locations caused us to observe false changes in polling location (or no change when one did occur) then it would result in units being falsely assigned to treatment or control. Such differential assignment could cause the appearance of significance in 2012, as well as potentially altering the 2016 results. We are conducting further research to determine the source of this outcome.

Of course, reassigning polling locations is often a necessary step in administering elections, whether it is due to an increase in absentee voting, a shortage of volunteers to run the polls, or simply because an old polling place is no longer a viable option. That being said, state officials should keep these findings in mind when making these decisions. Voting is already a substantial cost for many people, and a seemingly small change - like a poll reassignment that is a mile farther than before - could be the deciding factor in whether to turn up on Election Day.

7. Conclusion and Future Research

Our findings directionally follow prior work and support the discussion of important questions for further research. The causal effects we identify provides a comprehensive start for future questions into aspects such as differential effects between groups - are young voters deterred more easily than elders when a polling place is moved? For campaigns that are targeting the millennial vote, does this have an influence on campaign strategy? Furthermore, there remains an open question about the specific effects of seemingly non-partisan changes in polling places and voter ID laws on minority communities. Do these voting process changes have a larger causal effect on minorities, effectively violating the spirit of the Voting Rights Act? What is the interaction effect of minorities who have limited access to resources when their polling place is changed or closed? While our research is limited to North Carolina, it would be interesting to run the same model on other preclearance states such as Georgia and observe any differences and map them to state-specific legislation.

Understanding voting behavior overall is crucial for candidates, campaign teams, policymakers and the general public as well. Bruter and Harrison[17] claim that voters' interaction with electoral arrangements impacts their political psychology - one possible factor that leads to "upsets" in the outcomes that pundits predict. Electoral ergonomics incorporates behavioral economic principles and psychology within political actions and is defined as the "optimisation of all relevant electoral procedures and mechanisms to provide the best possible electoral experience for voters". Decisions made about the election process

[17] Bruter, Michael & Harrison, Sarah. (2017). Understanding the emotional act of voting. Nature Human Behaviour. 1. 0024. 10.1038/s41562-016-0024.

influence who turns up to vote, what voting method they choose, and how they feel about the potential choices, not just of candidates but also in the steps of the election process. All choices are relevant, meaningful and have outcomes that should be studied to inform future decisions. As highlighted in prior research, changing where voters go to vote may introduce additional search costs and inconvenience to the individual, likely affecting travel time to the polls which eventually influence decisions regarding how, where and who to vote for. Being cognizant of outcomes from this research and similar experiments in political science could be helpful when local officials analyze factors to make polling place closure and change decisions.

8. Acknowledgements

We would like to thank Professor D. Alex Hughes, University of California - Berkeley and Professor Micah Gell-Redman, University of Georgia for their support and insights through the course of this research project. We would also like to extend our gratitude to Jacque Blaeske for providing support from the NCSBE.

9. References

[1] Dyck, J. J., & Gimpel, J. G. (2005). Distance, turnout, and the convenience of voting. Social Science Quarterly, 86, 531–548.

[2] Haspel, M. and Knotts, H. G. (2005), Location, Location, Location: Precinct Placement and the Costs of Voting. Journal of Politics, 67: 560-573. doi:10.1111/j.1468-2508.2005.00329.x

[3] Blais, Andre. 2000. *To Vote or Not to Vote?* Pittsburgh, PA: University of Pittsburgh Press

[4] Squire, P., Wolfinger, R., & Glass, D. (1987). Residential Mobility and Voter Turnout. American Political Science Review, 81(1), 45-65. doi:10.2307/1960778 https://www.gsb.stanford.edu/insights/can-polling-location-influence-how-voters-vote

[5] Joslyn,N.,et al. Distance traveled to polling locations. Are travel costs imposed equally on party members? The Social Science Journal (2018) https://doi.org/10.1016/j.soscij.2018.12.006

[6] Stuart E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science : a review journal of the Institute of Mathematical Statistics*, *25*(1), 1-21.

[7] Bruter, Michael & Harrison, Sarah. (2017). Understanding the emotional act of voting. Nature Human Behaviour. 1. 0024. 10.1038/s41562-016-0024.

[8] Ming, K., Rosenbaum, P. (2001) A Note on Optimal Matching With Variable Controls Using the Assignment Algorithm, Journal of Computational and Graphical Statistics, 10:3, 455-463

[9] Dehejia, R. H., & Wahba, S. (1999). Causal Effects in Nonexperimental Studies: Reevaluating the Evaluation of Training Programs. *Journal of the American Statistical Association, 94*(448), 1053-1062. doi:10.1080/01621459.1999.10473858