The Great Poll Closure - How neighborhood composition is related to polling place changes

Master Thesis

A thesis submitted in partial fulfilment of the requirements for the degree of the Master of Social and Economic Data Science at the University of Konstanz.

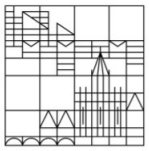
submitted by

Philipp Martin Bosch

01/911259

at

Universität Konstanz



Period of completion: 08.10.2021 - 15.03.2022

1st supervisor: Prof. Peter Selb

2nd supervisor: Prof. Susumu Shikano

Konstanz, 15.03.2022

Abstract

The notion that the act of voting entails costs for individual voters is one of the cornerstones of the study of political participation. How such costs of voting are altered by the closure and relocation of polling places recently became part of an overarching and long-lasting debate about voter suppression in the United States. In this thesis I add to this debate by providing a novel, fine-grained dataset including geocoded polling places covering four general elections from 2012 until the year 2018 in the United States. By constructing the dataset from a raw collection of fragmented addresses and subsequently utilizing it to answer my substantial research question, I provide both a first test of the validity of the data itself and present new insights into potential strategies of poll closures. Contrary to conventional wisdom and several reports from voting rights activist and the media, I do not find evidence of excessive polling place closures in counties with a high share of Black voters. What I do find, is that counties with a history of vote suppression and a high share of Hispanic voters did in fact close polling places at a higher rate. Ultimately, I discuss several approaches how this new dataset could be used to study the interplay of political participation and election administration, going far beyond the study presented here.

Contents

1 Introduction **1**

2 Poll Closures in Context **6**

2.1 A History of Voter Disenfranchisement 6

2.2 Empirical Evidence 14

2.2.1 Poll Closures and Turnout 14

2.2.2 The Politics of Poll Closures 23

2.3 Deducing Hypotheses from previous Evidence 27

3 Data **31**

3.1 Geocoding of Polling Places 32

3.2 Further Variables and Final Dataset 37

4 Empirical Approach **41**

4.1 A comparable Measure of Poll Closures 41

4.2 Modeling Closures between Elections 43

5 Results **46**

6 Discussion **51**

References **54**

Appendices **63**

List of Figures

1 Map of Counties and States formerly covered by the VRA . . 10
2 Sample of raw data 32
3 Percentage of missing Addresses after Geocoding by State . . 35
4 Counties with Polling Places for 2012 and 2016 Election in Sample 36
5 Comparison between the amount of Polling Places per County found in own data and EAVS data 36
6 Distribution of the Share of Black Voters by State 39
7 Map of Counties which were covered by the Voting Rights Act until 2013 40
8 PCVPP Values for two Counties from the Sample 42
9 Distribution of PCVPP Values by State 42
10 Posterior Distribution of Coefficients 46
11 Posterior Distribution of Coefficients 47
12 Posterior Distribution of Random Slope for Share of Hispanic Population in Arizona 48
13 Posterior Distribution of Random Intercept 49
14 Posterior Distribution of Coefficients for Categorical Model . . 50
15 Overview of covered Elections and States 63
16 Overview of covered Elections and States 64
17 Summary Table of Geocoding Process 65
18 Summary Table of Geocoding Process 66
19 Distribution of Share of Hispanic Voting Age Population . . . 67
20 Posterior Distribution of Random Intercept 68
21 Posterior Distribution of Random Intercept 69

1 Introduction

In the 2016 U.S. presidential elections, voters in entirely black neighborhoods waited 29% longer to vote and were 74% more likely to spend more than 30 minutes at their polling place, compared to entirely white neighborhoods (Chen, Haggag, Pope, & Rohla, 2020). Two years later, a report by Pew Trusts states: "In the five years since the U.S. Supreme Court struck down key parts of the Voting Rights Act [in 2013], nearly a thousand polling places have been shuttered across the country, many of them in southern black communities." (Vasilogambros, 2018)

At first glance, it might seem naturally to establish a causal link between these two observations and classify them under the currently much debated umbrella topic "vote suppression" (for an overview, see: Abrams, Anderson, Kruse, Richardson, and Thompson (2020). The aim of this thesis is not to establish this causal relationship. Rather, I provide a new data basis and generate insights about the second statement. Specifically, whether there are racial disparities between counties regarding the closure of polling places in the United States. In other words: do minority voters get targeted with polling place closures based on their place of living?

In order to reason about potential targeting strategies, two aspects about the administrative process of polling place (re-)location are central: First, how are the actors in this administrative process regulated? Second, what are the potential consequences of the closure and relocation of polling places?

Both, legislative circumstances and intended consequences (therefore strategies) of the actors involved can be derived from the history of U.S. legislation regarding voting rights. Taking into account these legislative realities and supplying them with recent journalistic observations and scientific evidence, I will demonstrate the potential impact, tactics like polling place closure can have on turnout from minority groups. Thus, I highlight the incentives political and administrative actors have to influence the electoral process and what is more, call attention to the need for a granular, large-scale data set

of polling place locations in the USA, which I supply with this thesis.

Throughout my thesis, I build upon the Economic Theory of Democracy by Anthony Downs to explain why poll closures have an impact on individual turnout. Put simply, Anthony Downs argues that the act of voting is time-consuming and therefore, costly. If the costs of voting surpass the benefits, a rational voter will abstain (Downs, 1957; Riker & Ordeshook, 1968). Concerning polling places and more specifically the location of a polling place, there are in principle three factors that have the potential to increase the costs of voting for individual voters. First, so-called search costs to find out where one has to vote and second, travel costs to the polling place on election day (Brady & McNulty, 2011). Somewhat less consideration so far has been given to a third cost factor. Namely, long lines in front of polling sites due to an excessive amount of voters per polling place (Fausset, 2014). By closing or relocating polling places, authorities are therefore able to intervene directly in the cost equation of individual voters (Haspel & Knotts, 2005). At plain sight, such interventions should affect all voters which are assigned to a new polling place equally. I will show that this is not a given. First, increasing partisan and racial segregation (Bishop, 2009) in the U.S. allows local authorities to precisely target specific voter groups with polling place closures. Second, drawing on the resource model of political participation, different groups of voters are more resilient to external shocks to the costs of voting (Brady, Verba, & Schlozman, 1995). Members of minority groups, on average equipped with less resources to overcome potential shocks to their utility function of voting (Verba, Schlozman, Brady, & Nie, 1993; Leighley & Nagler, 2013), thus suffer the most from the closure of polling places. Particularly long waiting times may harm minority and precariously employed voters most. Election day is only a public holiday in eleven out of fifty states across the United States. For many American voters, time lost standing in line to vote is tantamount to fewer hours on their paycheck.

Research concerning the closure and relocation of polling places so far was mainly concerned with the direct effects on individual turnout. At first glance, this is completely justified. Democracy rests on the notion that representatives of the people are chosen by the people (Cantoni, 2020). Therefore, evidence that administrative processes actively suppress turnout of certain groups of voters touch the core of the democratic system itself. However, speaking normatively, any appearance of manipulation of the electoral process poses a risk to democracy, not only if it is successful and proven (Shepherd, Fresh, Eubank, & Clinton, 2021). Furthermore, the individual decision to cast a ballot or to abstain depends on a myriad of external and personal factors (beyond the classical rational choice model of voting). Additionally, due to the federal system of the U.S., voters from different states often undergo completely different steps to cast a ballot. Thus, isolating the sole effect of polling place closures on individual turnout is only feasible using a sophisticated research design as well as utilizing circumstances which generate quasi experimental conditions.

Existing literature in the realm of voter turnout and poll closures mostly focussed on single states or municipalities to avoid confounding by context factors such as state laws or salience of single elections under observation. Nevertheless, the findings of such studies help to illuminate the motives behind poll closures. Although I will engage with prior research extensively in the following sections, it is crucial to state that the core assumption that the closure and relocation of polling places increases the costs of voting is supported by a host of prior evidence (Gimpel & Schuknecht, 2003; Haspel & Knotts, 2005; Brady & McNulty, 2011; Amos, Smith, & Ste. Claire, 2017; Yoder, 2018; Cantoni, 2020; Clinton, Eubank, Fresh, & Shepherd, 2021; Tomkins et al., 2021). How these costs translate to lower turnout and whether minorities suffer more from these costs thereby depends on the election and jurisdiction observed in each study.

Having established the general process that the relocation and closure of

polling places potentially decreases turnout, scholars more recently investigated whether such poll closure are used as a political tool to target minority voters or suspected voters of the political opponent in general. Contrary to a host of anecdotal evidence from news outlets (e.g., Roth, 2015), findings whether poll closures are used strategically are mixed. Shepherd et al. (2021) find no evidence of strategic targeting in their analysis of voter file and polling place data in North Carolina. The only nationwide study by Squires (2021) on the contrary finds that counties with a high share of minority voters experience both higher rates of poll closures and declining turnout rates. Both studies not only provide valuable insights into the theoretical process of targeting, but also underline the need for a national, fine-grained dataset of polling places. Since such data was missing so far, Squires (2021) falls back on self-reported survey data on polling places on the county level. Such data not only lacks information on the exact location of a polling place but, what is more, is also prone to non-response bias from local election officials who might not want to reveal the closure of polling places in their jurisdiction.

I address these gaps in the previous research by utilizing newly collected data about polling places from the Center for Public Integrity (Rebala et al., 2022). By geocoding partially fragmented addresses of more than 300,000 polling places across thirty-five states and four federal elections in the U.S., I supply a fine-grained dataset on polling place locations. This not only allows me to replicate the study of Squires (2021) with independently collected data, but to provide data that can be further utilized to study changes *within* counties. Hence, this thesis does not only contribute to the debate whether minority voters get targeted with strategies of vote suppression, but also provides a new data set for future research.

As the data from the Center for Public Integrity has not been used before in scientific analysis, my thesis also serves as a first test of the validity of the data. During geocoding and subsequent data preprocessing I encountered some inconsistencies in the data which I point out and call for future research

to investigate. I discuss these caveats extensively in the third section of my thesis when I show the data generation process. Subsequently, with a dataset which has undergone several plausibility checks, I set up a Bayesian multilevel model to answer my research question whether counties with a high share of minority voters experience comparatively more polling closures between elections. My results indicate that counties with a high share of Hispanic inhabitants do in fact experience more closures of polling places between the elections of 2012 and 2016. Somewhat surprisingly and in contrast to many reports from voting rights organizations and news outlets, this effect does not hold in counties with a high share of Black inhabitants.

But before we dive deeper into the dataset and the statistical model, I will give a short introduction to the legislative system of the United States and the ongoing struggle of political participation in the next section.

2 Poll Closures in Context

The purpose of the following section is to equip the reader with fundamental knowledge of the electoral system in the USA. On the one hand, in terms of administrative authority and, on the other, legislative regulations on the federal level. Neither of both can be understood without considering the history of the U.S. voter suppression. Of particular relevance thereby is the introduction of the Voting Rights Act in 1965 and its repeal in 2013. Against the backdrop of these legislative realities, I will subsequently discuss current scientific findings and deduce hypotheses regarding the targeting of minority voters through polling place closures.

2.1 A History of Voter Disenfranchisement

The Voting Rights Act (VRA), passed in 1965, was intended to put an end to decades of minority voter disenfranchisement, primarily in the former confederate states of the south. Mandated by President Lyndon B. Johnson to be "the goddamndest, toughest voting rights act that you can devise" (May, 2013, p. 48) the bill was introduced to enforce the rights of black people initially stated in the 14th and 15th amendments to the constitution of the United States. Known as the Reconstruction Amendments and introduced shortly after the civil war in 1867, these articles declared former slaves to be citizens and granted black men the right to vote (Abrams et al., 2020, p. 1).

Protected by law and the federal troops which were still deployed in the south, black voter turnout boomed in the following years. Nationwide, the percentage of black men who were eligible to vote rose from less than half a percentage point to 80.5 percent by the end of 1867 (Valelly, 2004, p. 3). In the south, where up to two-thirds of eligible black men voted during the height of Reconstruction (Davidson, 1992, p. 10), this translated into political representation as well. As a result, about 15 percent of office holders in the south were black in the year 1872, astonishingly a larger proportion even

compared to year 1990 (Davidson, 1992, p. 10).

However, the backlash from white southern Democrats, for whom the elimination of black suffrage was central (Keyssar, 2009, p. 85), follow swiftly. After the last federal troops left the south in 1877 a new generation of white southerners took control over local and state governments (Abrams et al., 2020, p. 5). With the help of *de facto* and *de jure* (Fraga, 2018, p. 32) barriers to voting, Black voter turnout was supposed to be cut back to a minimum. Additionally, what the states could not accomplish by law, they were more than willing to enforce by sheer violence against black voters (Anderson, 2018, p. 14).

The results of this orchestrated attack on black voting rights in the southern states were resounding. In Louisiana, for example, where "more than 130,000 blacks had been registered to vote in 1896, the figure dropped to a bleak 1,342 by 1904" (Waldman, 2016, p. 85). The same trend emerged in Alabama, where the black voting population shrank from 180,000 to less than 3,000 in just three years (Waldman, 2016, p. 85). Altogether, in the entire southern United States, a mere three percent of the eligible black population was registered to vote by 1940 (Waldman, 2016, p. 88).

Although the Supreme Court intervened and overturned discriminatory legislation such as the grandfather clause (1915) and all-white primaries (1944)[1] in the years between the Reconstruction and the passage of the VRA, states in the south still were able to implement *de facto* barriers to Black voting. This was due to the federal system of the USA, which grants individual states the constitutional power to administer local as well as federal elections and to delegate these powers further down to local authorities, so-called local election administrators (LEOs).

Besides poll taxes and literacy tests, the redrawing of electoral boundaries and the merging of constituencies were among the frequently used instru-

[1] for a detailed introduction to the grandfather clause and white primaries see: Greenblatt (2013) and Marshall (1957)

ments of *de facto* voter suppression. By combining electoral districts with a relative high proportion of Black voters and requiring candidates to run in larger districts, the aim was to "diffuse or dilute the black vote among a sea of whites" (Anderson, 2018, p. 25).

This practice of deliberately manipulating district lines is known today as *partisan gerrymandering*. While gerrymandering also constitutes an interference in the democratic process by altering the transformation of votes to seats (Lublin, 1997; Deford, Eubank, & Rodden, 2021), it is important to emphasize how it differs from practices which include polling place closures. Both practices occur at different administrative levels and are associated with different terminology. Therefore, let us quickly take a step back from the historic development of voting rights in order to illuminate the electoral process and competences which the U.S. Constitution specifies.

Gerrymandering is actually just a special (partisan and manipulative) form of a constitutionally mandated process called *redistricting*. It describes the process of redrawing the boundaries of electoral districts in individual U.S. states. *Redistricting* is a necessity, as after each decennial census, a reallocation of Electoral College votes and Senate delegates according to the population of each state is made. This process is also known as *reapportionment*. The procedure to determine the boundaries of each district then rests on the individual states. At present, 37 states resort to state legislatures, four states employ independent commissions and two use partisan commissions. Less populated states such as Alaska, North Dakota, Wyoming, South Dakota, Vermont, and Delaware do not redistrict at all, as they currently only have one congressional district (Squires, 2021).

Reprecincting, on the other hand, describes the process of redrawing the boundaries of election precincts, representing the smallest geographic unit of the electoral system in the United States. It is carried out by local governments (usually counties) and also takes place every ten years after the census or at the beginning of an election year, if deemed necessary by local

authorities. One outcome of *reprecincting* can be the closure, relocation, or opening of polling places, as local governments assign at least one polling place per precinct. Combining several precincts therefore can lead to the closure of polling places. What both practices, *reprecincting* and *redistricting*, have in common is, that they are directly derived from the first article of the U.S. Constitution: "The Times, Places and Manner of holding Elections for Senators and Representatives, shall be prescribed in each State by the Legislature thereof."

Turning back to the history of voting rights in the U.S. before 1965 the field of conflict for minority voters becomes obvious now. On the one hand, the constitutionally guaranteed right to vote, regardless of one's race, and on the other hand, the ability of states and counties to intervene in the election process, equally guaranteed by the Constitution. While it was possible in principle for individual citizens to go to court against individual barriers (which too blatantly violated the Constitution), practically this procedure was undermined by lengthy court proceedings and the intimidation of the plaintiffs (Anderson, 2018).

This only began to change with the increasing pressure of the civil rights movement in the U.S., which was instrumental in the passing of the Voting Rights Act of 1965. What made the VRA such a special and successful piece of legislation was the fact that it reversed the burden of proof. It was no longer up to the individual voter or the federal government to prove whether practices in the states or counties violated constitutional voting rights, but local authorities had to get federal approval in advance of any changes they wanted to make to the way elections were carried out. This was ensured by Section 4(b) ("coverage formula") and Section 5 ("preclearance requirement") of the VRA.

The coverage formula determined which states or local authorities were affected by the provisions of the VRA. This included states and counties which (i) used a test or device to restrict the opportunity to vote or register,

(ii) had less than half of their eligible citizens registered to vote in the federal elections of 1964, 1968 or 1972. States and local governments that were subject to the coverage formula subsequently were required by Section 5 to obtain federal approval of any changes to their election laws. In other words, to receive federal preclearance. In particular, they had the burden of proof that the proposed change would not adversely affect minority voters in the exercise of their right to vote. Although counties and states could "bailout" from coverage, this was difficult to secure in practice (Squires, 2021). Figure 1 provides an overview of the constituencies which were covered by the VRA over the years until it was last extended by congress in 2006.

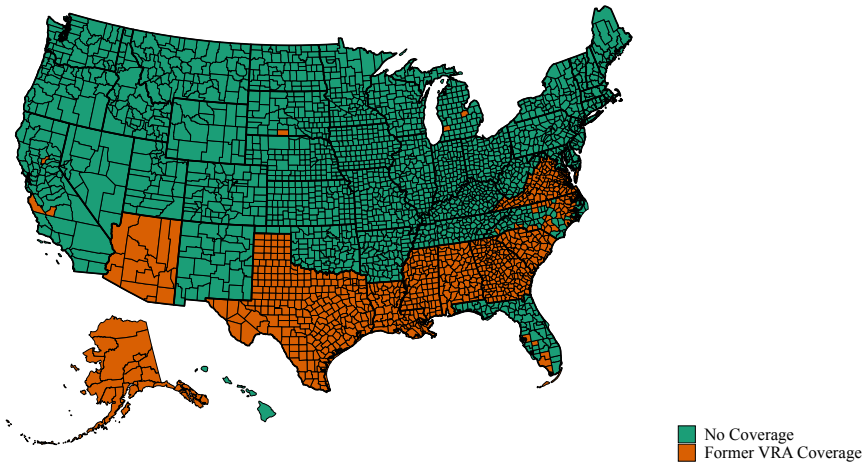


Figure 1: Map of Counties and States formerly covered by the VRA

Taken together, these strict measures ensured that the VRA actually turned out to be as effective as President Lyndon B. Johnson ordered it to be. Black registration and turnout surged in the covered constituencies. For example, "in Mississippi, black registration went from less than 10% in 1964 to almost 60% in 1968; in Alabama, the figure rose from 24% to 57%. In the region as a whole, roughly a million new voters were registered within a few years after the bill became law, bringing African American registration to a record 62%." (Keyssar, 2009, p. 212)

These descriptive findings are backed by scientific evidence regarding the positive causal effect the VRA had on registration and turnout (Fresh, 2018) as well as the political representation of minority voters (Schuit & Rogowski, 2017). Arguably, what best describes the importance the VRA possessed for American democracy as a whole, is the fact that many scholars simply do not consider the U.S. being a democracy prior to its implementation in 1965 (Mickey, 2015; King, 2017). Against this backdrop, the Supreme Court's decision to repeal the Voting Rights Act in its landmark ruling of *Shelby County v. Holder* (2013) becomes all the more significant.

Shelby County, like the entire state of Alabama, was subject to the coverage formula after 1965. Therefore, it was prohibited from implementing any form of redistricting without first obtaining federal approval. In 2010 the commissioners of Shelby County challenged this notion and subsequently brought the case to the Supreme Court. In its ruling in 2013, the Supreme Court, meanwhile equipped with a conservative majority, in principle followed the argumentation of the plaintiff county. The VRA, specifically the coverage formula, would unfairly single out southern states and constituencies based on outdated facts. Barriers to voting that existed in the 1960 and 1970 would no longer exist, and thus the act unduly interfered with the constitutional rights of individual states and counties. To put it ironically, the VRA's great success provided its critics with the arguments to ultimately strike it down. By ruling the coverage formula as unconstitutionally, the Supreme Court automatically rendered the complete Voting Rights Act moot.

The reaction to the ruling from several southern states, now freed from federal oversight, followed promptly. On the very day of the verdict, the state of Texas enacted a strict Voter ID law, which had previously been rejected by preclearance. North-Carolina swiftly followed with strict voter registration requirements that federal courts subsequently found to "target African Americans with almost surgical precision" (Ingraham, 2016).

Not quite as prominently discussed by the public, however, were the changes that took place at the local level through the practice of *reprecincting*. Leading up to the 2016 presidential election, the first federal election in fifty years held without protection of the VRA, 43% of formerly covered counties closed polling places according to an investigation by The Leadership Conference Education Fund. Titled with "The Great Poll Closure" the report states that voters in counties which were no longer under pre-clearance had at least 868 fewer polling places to cast their ballot in 2016, compared to past elections (Simpson, 2016). Nationwide, more than 3,000 polling places had been closed between 2012 and 2016, according to reports in news outlets (Astor, 2018).

To summarize, the current situation of voter suppression in the U.S. can be regarded as a result of "the ebbs and flows of the federal government's intervention in the South" (Abrams et al., 2020, p. 1). In this sense, the repeal of the VRA in 2013 represents a low point of such ebbs. In the years after *Shelby*, scholars argued that states in the U.S. are in danger of becoming "laboratories of authoritarianism" (Levitsky & Ziblatt, 2018, p. 2) or more recently "laboratories of democratic backsliding" (Grumbach, 2021). News outlets and scholars further pointed to the 2016 presidential election to prove that black turnout got under pressure once again by oppressive state legislation (e.g., Anderson, 2018). On a national level, black turnout declined for the first time in twenty years and fell from 66.6% in 2012 to 59.6% (Krogstad & Lopez, 2017). In some regions, like Wisconsin, this decline was even more pronounced. In a state that Donald Trump narrowly won by 27,000 ballots, black turnout dropped below 50% compared to 78% four years before, when Barack Obama ran for president (Wilson, 2017). Most commentators blamed this drop in turnout on the lack of popularity Hillary Clinton had among minority voters. While this reasoning seems plausible, there were also dissenting voices which attributed the "disappearing minority voter" (Anderson, 2018, p. 1) to systematic legislative vote suppression and

called it the 2016 "campaign's most misunderstood story" (Anderson, 2018, p. 1).

This section outlined the long history of vote suppression and equipped the reader with knowledge about the electoral arena, as well as initial evidence of repressive legislation in the years following 2013. It demonstrates that the debate about vote-suppression in the U.S. is far from settled. By analyzing new data on the location of polling places in the U.S., this thesis contributes to elucidating a precondition of the debate of vote-suppression. Namely, whether minority communities are targeted with polling place closures in a systematic manner or if the evidence from news outlets listed above are just idiosyncrasies.

The next section serves to bridge the gap between poll closures and vote-suppression by providing scientific evidence surrounding the location and closure of polling places.

2.2 Empirical Evidence

Although the dependent variable of interest in my thesis is the closure and relocation of polling places, it is crucial to address the subsequent effects such closures have on turnout. Not only does this evidence justify the need for an exploration of where these closures occur, but it also helps to clarify the presumed process of strategic targeting, which I assume in this thesis. Due to the decentralized nature of the US electoral system, such strategic targeting lies in the hand of more than 8,000 local election officials (LEOs) (Pettigrew, 2017). Whose true motivations, however, are unobservable. Scientific evidence about the effect of poll closures therefore allows drawing conclusions about the motivations of election officials. In other words: examining whether poll closures occur strategically is only reasonable against the backdrop of the potential effects such closures have.

2.2.1 Poll Closures and Turnout

The basic idea of why something as seemingly mundane as the location of a polling place influences the likelihood of voting traces back to Anthony Downs' economic theory of democracy. Downs argues that the act of voting entails certain costs and benefits for individual voters. If the costs outweigh the benefits, rational voters will refrain from voting (Downs, 1957; Riker & Ordeshook, 1968). There are three conceivable ways how the closure and relocation of polling places during reprecincting increase the costs of voting for individual voters. First, so-called search costs, which describe research costs to find out about the location of the new polling place. Second, longer travel times (also referred to as transportation costs) to the new polling site if it is located further away or less accessible compared to the previous polling site (Brady & McNulty, 2011; Shepherd et al., 2021). Third and last, increased waiting time in front of the polling place as the remaining polling places have to serve more voters after the consolidation of precincts (Fausset, 2014). Search and transportation costs are thereby regarded as

affecting the first-order outcome if an individual decides to cast a ballot, whereas long lines and waiting times are regarded as second-order effects (Brady & McNulty, 2011). This is due to the fact that long waiting times in front of a polling place only influence the probability of casting a ballot *once* a voter shows up to the polls (Shepherd et al., 2021). While this reasoning seems plausible considering individual voters and elections in isolation, the distinction in reality might not be as clear-cut. First, information about long lines is also available to voters through other channels (e.g., social media) apart from first-hand experience on election day and second, research has shown that long waiting times at the polls decrease turnout in subsequent elections (Pettigrew, 2021). Thus, it follows that reprecincting in theory impacts the individual decision to show up at a polling place on election day directly and through a multitude of mechanisms.

In the years since Downs introduced his rational choice approach of voting, the model has seen major revisions and extensions to incorporate concepts such as obligation, altruism, psychic gratification or citizen duty on the benefit side of the equation (Riker & Ordeshook, 1973). Additionally, models of political participation have been introduced that go beyond the classical rational choice approach. On the one hand, psychological models, which attribute the decision to cast a ballot to internal motivational attachments and on the other hand, sociological models, which describe political participation as a product of social networks surrounding individuals (Campbell, Converse, Miller, & Stokes, 1960; Fiorina, 1976; Washington, 2006; Gerber, Green, & Larimer, 2008; DellaVigna, List, Malmendier, & Rao, 2017). What all these models agree on is the basic notion that voting is costly (Akee, Copeland, Costello, Holbein, & Simeonova, 2018). In the context of poll closures, one theoretical model is particularly suitable for explaining why some citizens turn out to vote and others abstain. Namely, the resource model of political participation (Brady et al., 1995). According to this model, individual resources of voters help them to overcome the costs of voting. Key resources

linked to turnout include education (Sondheimer & Green, 2010), health (Burden, Fletcher, Herd, Jones, & Moynihan, 2017), information (DellaVigna & Kaplan, 2007; Holbein, 2016; Lassen, 2005), civic skills (Verba, Schlozman, & Brady, 1995), time (Holbein & Schafer, 2017), and income (Schlozman, Brady, & Verba, 2018). Each of the mentioned resources is closely related to race in the United States. Therefore, the initial authors of the resource model claim "that differences among [African-Americans, Anglo-Whites and Latinos] with respect to participation can be attributed almost entirely to the unequal political resources at their disposal" (Verba et al., 1993, p. 494).

Applied to the situation of racial segregation and partisan sorting in the U.S. (Nall, 2015; Levendusky, 2009; Rothstein, 2017; Bishop, 2009) this implies that citizens who live closer together, also share similar resources with respect to political participation. Against the backdrop of historical vote-suppression and in connection with the ample discretion local election officials have over the way elections are carried out in the U.S., a clear picture emerges about the potential poll closures can have on turnout of minority groups.

Although Anthony Downs already pointed out in his initial work that "the returns from voting are usually so low that tiny variations in cost may have tremendous effects on the distribution of political power" (Downs, 1957, p. 266) most work from political scholars (like the resource model) have focused on the conditions which foster turnout and mostly neglected the cost side of the equation. However, the emergence and accessibility of geographic information system software (GIS) in the early 2000s also propelled scientific work which scrutinized the costs implied by the availability and accessibility of polling places.

In a first approach, Gimpel and Schuknecht (2003) used the geolocation of over 300 polling sites for the 2000 presidential election in three counties of Maryland and measure the distance from each polling site to the population centroid of each associated precinct. Lacking individual data on voters, they

were only able to draw ecological inferences about turnout on the level of precincts. They found that distance has the strongest impact in suburban precincts where polling places are relatively close. For rural precincts where the distance to the polling place is between 6 miles (ca. 10 km) and 10 miles (ca. 16 km), turnout rates were less affected.

Building upon this evidence, Haspel and Knotts (2005) connected voter file data with polling place locations in the city of Atlanta and showed that distance to the polling place has a significant negative impact on the decision to vote during the 2001 mayoral race. More surprisingly, they also found that moving a polling place actually increases turnout. The authors argue that this might be due to the fact that in Atlanta the number of polling places actually increased for the observed election and therefore, polling places in some cases moved closer to voters. Hence, increased information costs about the new polling place were offset by diminishing transportation costs.

Brady and McNulty (2011) subsequently scrutinize such shocks of moving polling places and thus go beyond the pure study of distance on turnout. The authors utilize the consolidation of precincts in Los Angeles County during California's 2003 gubernatorial recall election as a natural experiment. They find that changing polling places resulted in a drop in election day turnout of 3.03 percentage points, which was partially offset by an increase in absentee voting of 1.18 percentage points for affected voters. The design of the study thereby allows the authors to disentangle search and transportation costs and whether increasing costs rather cause voters to abstain or switch to mail ballots. Following the results of Brady and McNulty (2011) 40% of the potential voters who were dissuaded from voting in person on election day switched to mail ballots whereas the remaining 60% decided to abstain completely. The authors conclude: "People make a decision about whether to vote based on the increased search costs from having their polling place moved, and if they decide to vote, they choose absentee or polling place voting based on both search and travel costs." (Brady & McNulty, 2011,

p. 128)

Notably, although Brady and McNulty (2011) begin their paper with an anecdote about strategic poll closure during an election in Houston, the validity of their approach (the natural experiment) rests on the assumption that there is no significant difference between the group of registered voters affected by polling place changes and the control group of registered voters whose polling place remained stable. Through various balance checks they ensure that polling place changes in fact occurred as-if random and not in a manner consistent with partisan targeting. Turning to partisan effects regarding turnout of different groups of voters, the authors find that registered Democrats were more sensitive to changes compared to registered Republicans, but only to a negligible extent (0.22 percentage points difference in turnout between both groups). Nonetheless, the authors point out that although "partisan effects are small enough that they can probably be ignored when polling places are essentially changed randomly, as in Los Angeles, but they are large enough that they could be used by an unscrupulous politician or registrar to manipulate an election" (Brady & McNulty, 2011, p. 128)

Picking up this point, Amos et al. (2017) investigate reprecincting in Manatee County, Florida prior to the 2014 general (midterm) election. The authors choose this case explicitly as "there is evidence that the permanent reprecincting and reduction in the number of polling stations in Manatee County in early 2014 was not done dispassionately, with blind disregard of the partisan (or racial or ethnic) makeup of existing precincts" (Amos et al., 2017, p. 137). For example, "minorities, Democrats, and younger registered voters were disproportionately more likely to be reassigned to a new polling place" (Amos et al., 2017, p. 153). Regarding turnout, the assumptions of the authors about partisan and electoral targeting are also confirmed. In general, registered voters who were assigned to a new polling place were less likely to cast a ballot. The group of registered Latino voters were most affected by this. Notably, Amos et al. (2017) find similar results as Brady and McNulty

(2011) in terms of abstention and switching to "convenient" modes of voting. About 60% percent of the voters who were dissuaded from voting in person on election day, choose to abstain and the remaining 40% switched to early in-person voting or to vote by mail.

Having established that pure distance and changes in polling place location affect turnout and are not always carried out in a technocratic and nonpartisan nature, scholars subsequently scrutinized whether this effect materialized also in a larger context and not only in specific counties or second-order elections.

Thereby, the state of North Carolina offers an appealing test case to scholars for several reasons. First, North Carolina belongs to the so-called swing-states in which elections are often decided by razor-thin margins. For example, Donald Trump just narrowly won the state by 1.34% (74.483 votes) in 2020. Changes in turnout of the magnitude Brady and McNulty (2011) and Amos et al. (2017) found, can therefore be decisive for the outcome of an election. What is more, the decision where to open and close polling places is in the hands of partisan election boards in North Carolina counties (North Carolina's equivalent to LEOs). These election boards are appointed by the Governor. Finally, North Carolina also has a long history of vote-suppression, which is why 40 of the 100 counties in the state were under coverage of the Voting Rights Act until its repeal in 2013. Additionally, from an empirical perspective, the availability of individual-level turnout data in the form of voter files and information about polling places and precincts from official resources make North Carolina a prime example to study the direct effects of polling place changes on turnout.

Yoder (2018) and Clinton et al. (2021) take advantage of these circumstances and observe the effects of polling place changes on turnout during the three presidential elections of 2008, 2012 and 2016 (Clinton et al., 2021) as well as every primary and general election from 2006-2016 (Yoder, 2018). Yoder (2018) finds "that changing a voter's polling place location causes a 1

to 2 percentage point decline in general election turnout likelihood". In contrast Clinton et al. (2021) demonstrate that albeit polling place changes do decrease in-person election day voting by 0.7 percent, "this effect is almost completely offset by substitution into early voting". The authors theorize that "this is because voters whose polling places change location receive notification mailers, offsetting search costs and priming them to think about the election before election day, driving early voting" (Clinton et al., 2021, p. 800). Although Clinton et al. (2021) find that non-white voters are less likely to substitute Election Day voting with early voting when they are confronted with polling place changes, they do not find statistically significant evidence that there is an overall negative turnout effect for minority voters.

The contradictory results between Clinton et al. (2021) and Yoder (2018) are thereby, according to the authors, due to the incomplete sample of counties, used in the study of Yoder (2018). By incorporating primary elections, Yoder (2018) only obtained full precinct data from 35 counties. Using the full sample of North Carolina counties from Clinton et al. (2021) and considering only presidential elections, Yoder (2018) arrives at similar results as Clinton et al. (2021).

Research that goes beyond single states is scarce, with only two notable exceptions by Tomkins et al. (2021) and Cantoni (2020). In their yet unpublished preprint, Tomkins et al. (2021) use census blocks as geographic discontinuities and utilize data from several[2] U.S. states. The design allows them to study the effects of both relative distance to the polling place and shocks introduced by polling place change. In doing so, Tomkins et al. (2021) compare voters who reside in the same block but are assigned to different polling places during the presidential elections of 2012 and 2016. In line with Clinton et al. (2021) in North Carolina, the authors find that both increased

[2]Although the authors contacted all 50 states to collect data about precincts and polling places, only 18 responses met the basic data quality criterion by the authors. Moreover, only six states supplied enough information regarding modes of turnout to study the effects of shocks and distance on different types of voting.

transportation costs and search costs induced by polling place changes, decrease in-person election day voting, but the effects on overall turnout are completely offset by increases in early in-person or mail-in voting.

Most notably, these results are in stark contrast to the study of Cantoni (2020) who also utilizes geographic discontinuities between voting precincts. Covering nine municipalities in Massachusetts and Minnesota between 2012 and 2016, Cantoni (2020) finds that "a 1 standard deviation (0.245 mile) increase in distance reduces ballots cast by 2 to 5 percent across four elections. During non-presidential elections, effects are three times larger in high-minority areas than in low-minority areas." (Cantoni, 2020, p. 61)

Synthesizing the existing evidence and deducing general insights is challenging due to the varying context and design each study employs. However, the core assumption that distance to the polling place and getting assigned to a new polling place increase the costs of casting a ballot on election day is confirmed by all mentioned papers. The exact extent and how much of these costs are offset by switching to "convenient" methods of voting varies by design of the study, geographical entity under observation, type of election and year of election. As Amos et al. (2017) put it in their study about Manatee County, Florida while describing the challenges of modelling turnout in the decentralized electoral landscape of the United States: "With the advent of more convenience voting opportunities in the form of early voting and absentee mail ballots, the ability of scholars to specify the possible costs for voters when choosing to cast a ballot under seemingly innumerable permutations is a technical nightmare." (Amos et al., 2017, p. 141)

However, some insights can be extracted which are also of interest when examining whether minority neighborhoods are exposed unduly to these costs. First, it is crucial that voters are given the opportunity to substitute in-person voting on Election Day with other "convenient" ways of voting and have access to information on their choices in the event of polling place changes. This is of particular relevance given that since 2013, states in the

U.S. often "bunch" together different election laws and regulations. For instance, voting-ID laws get tightened and the period in which early-in-person voting is possible is abridged. Following the evidence laid out above, cutting polling places while narrowing down the options to switch to other modes of voting might lead to significant drops in turnout. What is more, the differing results also highlight that one should be cautious about drawing general conclusions from individual case studies. Therefore, it is useful to first get an impression about where closures occur on a national scale and not restrict observations to counties and states that appear suspicious.

Such large scale studies of poll closures are rare so far, due to lack of systematic administrative data. In the next section I will introduce two notable exceptions and show how my approach relates to these studies.

2.2.2 The Politics of Poll Closures

As mentioned above, the state of North Carolina is not only ideal to study the effects of poll closures on turnout, but also to study whether such closures occur strategically across a whole state. Reprecincting in North Carolina is in the hand of partisan election boards, there is variation between counties in terms of former coverage by the VRA and what is more, in 2013 a Republican won the office of Governor defeating the Democratic incumbent which creates the opportunity to study poll closures under different partisan regimes. Exploiting these circumstances Shepherd et al. (2021) utilizes the same data that Clinton et al. (2021) use to study turnout (in fact both papers are from the same group of researchers), but now focus on whether the probability that registered voters experience a change in polling place is contingent on partisanship and race of voters themselves. Due to extensive media coverage about vote-suppression in North Carolina (e.g., Jacobs, 2018), the highly contested nature of elections in North Carolina as well as the partisanship of election officials, the authors have strong theoretical priors about the expected results. However, they do not find evidence consistent with partisan manipulation for electoral gain during the three presidential elections they investigate (2008, 2012 and 2016). Although the authors "do find variation in the racial and partisan impact of precinct and polling place changes when we compare counties to one another" (Shepherd et al., 2021, p. 173), these variations are neither associated with targeting of opposition voters nor the result of removed coverage provided by the VRA and "are likely attributable to idiosyncratic voter differences rather than partisan targeting." (Shepherd et al., 2021, p. 174) This finding again highlights the danger, case-studies concerning single counties entail, if one wants to make inferences about widespread vote-suppression.

Furthermore, the authors emphasize that although they do not find evidence of strategic targeting in the years they study, this does not mean that such procedures were never attempted, they were just not implemented ef-

fectively in North Carolina. Therefore, the authors call for future research to collect further state-level data about precincts and polling places to study the phenomena in other electoral settings.

One approach to this is the hitherto unpublished dissertation by (Squires, 2021). Squires (2021) utilizes mere counts of polling places in counties obtained from the Election Assistance Commission's Election Administration and Voting Survey (EAVS). Thereby, Squires (2021) is neither able to study the effects of relocations if the number of polling places remains stable in a county nor if different neighborhoods *within* a county are affected differently contingent on their sociodemographic composition. Nevertheless, the study provides vital first scientific insights about where poll closures occur and whether this correlates with the sociodemographic composition of counties. By using the EAVS datasets of 2008, 2012 and 2016 Squires (2021) is also able to study the effect the repeal of the VRA in 2013 had on poll closures on the county level.

Due to the fact that local authorities are not required to fill out the EAVS, the final dataset of his analysis contains 40 states and 2,270 counties within them. Alaska, New York, Utah, Illinois, Wisconsin, New Mexico, and Virginia failed to complete the survey in at least one of the investigated years. The remaining three states of Colorado, Washington, and Oregon do not run traditional polling places anymore, as they switched completely to mail-in ballots.

In order to control for varying population size, Squires (2021) calculates the amount of polling places per eligible voter (voting age population) in each county. Using difference-in-difference models, he is therefore able to study the percentage change of voters per polling place between the elections of 2008, 2012 and 2016. Squires (2021) differentiates in his study between "(1) constitutionally valid reasons and (2) politically motivated reasons" (Squires, 2021, p. 52) to consolidate precincts and shutdown polling places. Constitutionally valid reasons include, according to Squires (2021), budgetary constraints

of jurisdictions as well as the implementation of alternative voting methods. Politically motivated reasons include the effect the repeal of the VRA had, racial composition of the county and selection method of local election officials. Squires (2021) argues that local election officials (LEOs) that are appointed rather than elected, are more likely to use poll closures for political reasons compared to elected LEOs. This is because elected LEOs are more concerned with appeasing the electorate they are responsible for, compared to appointed LEOs that are more interested in pleasing the partisan officials they were appointed by (Squires, 2021).

The results of his analysis are quite mixed. In line with expectations and anecdotal evidence by news outlets, he finds that counties which were formerly under preclearance closed polling places at a much higher rate compared to counties which were never covered by the VRA. Claims that especially counties with a high proportion of non-white inhabitants were affected by poll closures are not fully supported by Squires (2021). Neither in counties formerly covered by the VRA nor in non-preclearance counties reaches the effect of the share of Black or Hispanic inhabitants statistical significance. Only after including categorical indicators for counties with particularly high shares of black ($> 30\%$) and Hispanic ($> 15\%$) inhabitants, the results are in line with the expectations of Squires (2021). Between 2012 and 2016 former preclearance counties with high shares of non-white inhabitants closed polling places at a much higher rate compared to other counties. This effect does not occur prior to the 2012 election in the same counties, which suggests that the effect is related to the abolition of the VRA and not due to some longer trend in these counties.

With respect to the election method of local election officials as well as constitutionally valid reasons to close polling places, the results of Squires (2021) are not straightforward either. He finds that in 2016 counties with elected LEOs closed polling places at a higher rate compared to counties with appointed officials, which contradicts his expectations. One possible expla-

nation for this could be the unobserved partisanship of LEOs in his study. To code the election method of LEOs, Squires (2021) uses a data set by Kimball and Kropf (2006) which contains the selection method as well as the party affiliations for all local election officials in the United States in the year 2000. Over time, these party affiliations changed, and Squires (2021) therefore is only able to make use of the binary variable whether LEOs were appointed or elected. However, LEOs' party affiliation may have a far greater impact on poll closures than their method of selection. For example, research shows that Republican LEOs are more likely to resort to measures that make ballot access more difficult if their district is majority Democratic (Burden, Canon, Lavertu, Mayer, & Moynihan, 2013). In addition, Democratic LEOs in populous counties focus more on increasing the size of the electorate compared to Republican LEOs (Kimball & Baybeck, 2013). Therefore, only capturing the selection method of LEOs without controlling for partisanship in a model might be ill-advised.

Also, the operationalization of budgetary constraints employed by Squires (2021) appear to be far from perfect. As there is no national (or even statewide) database of local election budgets, Squires (2021) resorts to the median income of the county as a proxy for budgetary constraints. Although there might be some relation between income of inhabitants and budget available for administering elections, the measure is far from perfect. Squires (2021) acknowledges this and calls for future research and data collection in this domain. It is therefore not surprising that the effects of median household income fail to reach statistical significance for the 2012 election, and in 2016 even indicate that counties with higher income closed (slightly) more polling places.

Last, examining the effects of the availability of "convenient" voting methods, the results remain mixed. For the election in 2016 neither the availability of early voting, nor no-excuse absentee voting, nor whether voting centers exist in a state significantly affected the closure of polling places. No-excuse

absentee voting thereby means, that any voter can request to cast a ballot via mail without excuse or reason necessary. Currently twenty-six states and Washington, D.C. offer this form of voting.[3] States that operate voting centers on the other hand allow voters to cast their ballot in every voter center across the state and not only in their designated polling place. Eighteen states across the United States currently allow their counties to operate Voting Centers.[4]

The absence of effects of such convenient methods of voting might also be duo to the binary operationalization chosen by Squires (2021). Although, states might offer early voting or mail-in ballots, the processes vary widely among states in terms of the effort required to use these forms of voting. Binary variables do not capture such variation between states.

2.3 Deducing Hypotheses from previous Evidence

Taken together, both studies by (Squires, 2021) and (Shepherd et al., 2021) yield valuable insights for my thesis. First, the relationship between poll closures and racial composition of jurisdictions is not as clear-cut as reports from several news outlets and civil rights groups propose. Second, to foster future research, there is a need for fine grained data about polling place locations and voters on a national scale. On the part of polling places, I will address this need and supply a dataset with geocoded polling places for 35 states in the U.S., covering four general elections between 2012 and 2018. Unfortunately, concerning voter data, not all states in the U.S. provide as extensive and convenient access to their voter file data as North Carolina does. National voter file data such as the files provided by Catalist[5] are only accessible against a fee, and are therefore excluded from this thesis project with

[3] https://www.ncsl.org/research/elections-and-campaigns/vopp-table-1-states-with-no-excuse-absentee-voting.aspx

[4] https://www.ncsl.org/research/elections-and-campaigns/vote-centers.aspx

[5] https://catalist.us/

limited scope and budget. Another issue with voter file data, largely ignored by the listed papers above, concerns coverage error (e.g., Groves et al., 2011). While voter file data offers an intriguing alternative to self-reported survey data about voting behavior, by definition it only covers registered voters. Therefore, studies utilizing voter file data to study general voting behavior are prone to "differential registration bias" (Nyhan, Skovron, & Titiunik, 2017). Applied to poll closures this bias would come into play, if registered voters react differently to external cost-shocks compared to non/not yet-registered voters. Given the evidence of habitual voting (Plutzer, 2002; Gerber, Green, & Shachar, 2003) it is reasonable to assume that non-voters (non-registrants) are more sensitive to increased costs of voting and refrain from registering in the first place if costs go up. This would mean, that all evidence listed above which utilizes voter file data actually underestimates the effect on over-all turnout. Notably, this is only acknowledged by Tomkins et al. (2021) in their hitherto unpublished working paper.

Applied to my thesis, I conclude that the approach taken by Squires (2021) to study effects on an aggregate level of counties is justified. Especially as reprecincting is decided on the county level. Moreover, considering racial composition of inhabitants, counties still exhibit large between-group variation. I will provide descriptive details of this in the next section concerning data.

Although I follow the general approach of Squires (2021) my thesis and hypotheses deviate from his approach in certain points. First, I do not rely on self-reported counts of polling places from counties. The data I employ stems from an independent research project by the Center for Public Integrity[6] and missing data therefore is more likely due to chance than local election officials not wanting to reveal the extent of poll closures in their jurisdiction. Moreover, by geocoding the addresses of polling places, I am able to scrutinize not only whether the total amount of polling places changes between elections

[6] https://publicintegrity.org/about/

but also whether polling places move to a new location.

Thus, in line with Squires (2021) I examine the hypothesis whether counties with higher shares of Black or Hispanic eligible-voters are more likely to experience polling place closures. I expect that the effect of a high share of minorities is larger in counties which were covered by the VRA until 2013. To assure (partial) comparability with the results of Squires (2021) I limit the analysis to the presidential elections of 2012 and 2016. Concerning control variables, I deviate slightly from Squires (2021) investigation. I do not scrutinize whether the election method of LEOs influences the amount of poll closures or relocations. This is due to the mentioned unobserved partisanship of LEOs. Moreover, research has shown that although minority voters, for example, have to wait longer in line to vote, waiting time in front of polling places is not contingent on partisanship of counties or party affiliation of chief elections officers in a state (Grumbach, 2021). Racial discrimination might therefore not only be the product of partisan considerations. This is also fostered by studies which show that individuals with low socio-economic status are less likely to file complaints towards local administrations (Jones, Greenberg, Kaufman, & Drew, 1978; Mladenka, 1980). Thus, also LEOs with Democratic background might be closing polling places in majority minority precincts due to the expected low backlash from regions with low socio-economic status (Pettigrew, 2017).

I also refrain from including the measure of budgetary constraints of local election administers which Squires (2021) uses. This is due to the poor operationalization mentioned above. However, I follow Squires (2021) and take convenient voting methods such as vote-by-mail and early voting into account. Deviating from his work, I try to capture differences between states in terms of accessibility of such convenient voting methods by using percentages of votes which were casted by using such convenience methods in each state in 2016. Thereby, I expect that higher percentages of convenience votes indicate easier access to such methods and thus a constitutionally valid

reason for jurisdictions to close polling places.

Finally, it should be mentioned that I could not access the original data that Squires (2021) uses in his analysis, as the author is currently still preparing the code and the datasets for publication. As a big proportion of this thesis project was devoted to geocode the raw addresses collected by the Center of Public Integrity I will introduce the data employed in subsequent analysis in more detail in the next section.

3 Data

To examine how the racial composition of counties relates to the closure of polling places within these counties, I create a dataset which contains the addresses of polling places and sociodemographic information about the county they belong to. Therefore, I systematically geocode the address of each polling place contained in the database supplied by the Center for Public Integrity[7] in order to obtain the exact coordinates of each site. In total, the dataset on GitHub contains 440.542 observations and data about 35 states, covering four general elections between 2012 and 2018. A full overview of the states and elections included in the raw data is depicted in Appendix A. After geocoding and assigning each polling place a unique location, I perform spatial linkages to merge these coordinates with official county shapefiles provided by the United States Census Bureau. Such shapefiles contain the boundaries of counties as so-called multipolygon variables. Put simply, multipolygons describe a set of points (in this case coordinates) which can be connected by lines to form borders of objects (in this case county borders).[8] A spatial link between points of coordinates and county shapefiles therefore simply checks which pair of latitude and longitude points sits within the multipolygon boundaries of a county. In a last step of data preparation, I enrich the dataset of counties and polling places with sociodemographic information of counties obtained from the American Community Survey and further information about the accessibility of convenience voting and whether a county formerly was under pre-clearance of the Voting Rights Act.

[7] https://github.com/PublicI/us-polling-places

[8] for a more throughout explanation of multipolygons see: https://wiki.openstreetmap.org/wiki/Relation:multipolygon

3.1 Geocoding of Polling Places

At the core of the dataset which I provide for future research and utilize in my own analysis are the exact coordinates of polling places. The transformation of raw addresses to exact coordinates therefore also constituted the biggest effort during this thesis project. This is mostly due to drastically varying data quality of the raw addresses supplied by the Center for Public Integrity. Such caveats are also addressed by the research team which compiled the dataset for the Center for Public Integrity:"This dataset is only as good as the information that is provided by state and local officials."[9] Geocoding this data and subsequently scrutinizing it in analyses therefore also serves as a first test of the validity of this new data source. To show the extent to which the quality of the raw data varies, I provide a sample of observations from the original dataset in Figure 2.

state	jurisdiction	jurisdiction_type	precinct_name	name	address	county_name	tidy_address
ak	fairbanks north star	borough	aurora	agricultural museum - fairgrounds	1800 college road, fairbanks, ak 99101	NA	1800 college road, fairbanks, ak 99101
ma	arlington	town	NA	brackett school	66 eastern avenue (left @ water tower @ top of park avenue)	middlesex	brackett school, 66 eastern avenue, arlington, middlesex county
ms	NA	county	county precinct 2	claiborne county fire department	NA	claiborne	claiborne county fire department, claiborne county, ms
ms	NA	county	college hill	college hill community center	10 county road 130, oxford, ms 38655	lafayette	10 county road 130, oxford, ms 38655
pa	NA	county	cheltenham 4-3	wyncote elementary school	333 rices mill (& barker rd) rd wyncote pa 19095	NA	333 rices mill rd wyncote pa 19095
wv	NA	county	NA	junior elementary school	NA	barbour	junior elementary school, barbour county, wv

Figure 2: Sample of raw data

Looking at the entries for the state of Mississippi `(state = ms)` for example, it becomes evident that even within single states the amount of information and quality varies drastically. Thus, it was not feasible to use the raw addresses and supply them directly to geocoding services like Google's geocod-

[9] https://github.com/PublicI/us-polling-places

ing API[10] or the Nominatim API[11] from OpenStreetMap (OSM). Hence, I had to write generic functions for each of the 35 states in the raw data, in order to arrive at the cleaned addresses depicted in the green-highlighted column `tidy_address`. Like the complete analysis of this project, I carried out data cleaning using the statistical programming language `R`(R Core Team, 2021). Especially for data wrangling, I relied heavily on packages and paradigms from the `tidyverse` (Wickham et al., 2019). Although I tried to incorporate as many idiosyncrasies from each state in each generic data cleaning function, some entries just lacked any relevant information (beyond state and election) at all. As these missings did not occur systematically within states, I classified them *as-if* random and not a threat to the general validity of the dataset.

After creating tidy addresses for each state, I geocoded addresses in each state separately to directly assess whether the quality of geocoding responses varied between states. To actually derive latitude and longitude coordinates from cleaned addresses, I relied on two geocoding services, namely Google's geocoding API and the Nominatim API from OSM. This two-pronged approach is due to the limited monetary budget of a master's thesis. Although I would have preferred to rely solely on Google's commercial geocoding API, each request via this API is charged with 0.004 $ and supplying more than 400,000 addresses would therefore quickly exceed the free tier of 300$ which Google grants to new users of the API. Even in the best case scenario, if all polling places remained stable across all four elections, more than 100,000 individual addresses would need to be geocoded. Therefore, I partly relied on the free geocoding service provided by the Nominatim API which queries the database of the OpenStreetMap project. Switching completely to the Nominatim API was not an option, however, as the API only provides reliable results for complete addresses. For example, the last entry of the

[10] `https://developers.google.com/maps/documentation/geocoding/overview`
[11] `https://nominatim.org/release-docs/latest/api/Overview/`

sample data in Figure 2 (`state = wv`) would lead to an empty response by the Nominatim API. As such partial addresses occurred frequently, I employed a flexible approach, utilizing both Google's Geocoding API and the Nominatim API. First, I queried each address against the Nominatim API and whenever the response was empty (no exact coordinates were found) the address was automatically submitted to the Google API. In order to reduce server load for the free Nominatim API and to stay within the free tier of Google's API, I also ensured to only submit unique addresses. Communication with both APIs thereby was facilitated by the `tidygeocoder` package (Cambon, Hernangómez, Belanger, & Possenriede, 2021).

As a measure of caution, I also implemented a sanity check for the retrieved coordinates. For each state I checked whether the retrieved coordinates fell within the boundaries of said state. Each observation which failed this sanity check was subsequently manually researched by myself and recoded if I could find sufficient information online. I applied the same logic to observations which could not be geocoded to any coordinate. The summary table in Appendix B gives a complete overview of the number of missing values after geocoding, broken down by states. Moreover, the amount of observations which were geocoded by OSM or Google is depicted. In short, the vast majority of addresses could be geocoded successfully by either OSM or Google. Out of 440,542 observations just 2509 entries (0.57%) did not yield plausible coordinates. Of these 2509, I was able to retrieve 1948 via manual research online. Thus, only 0.127% of observations from the raw data could not be matched to coordinates. To give a better overview about how missing values distribute over the single states in my sample, I depict the percentage of missing values on the map of Figure 3.

As my analysis focuses on the change of polling places between the two presidential elections of 2012 and 2016, I reduce the geocoded dataset for further analysis. The full dataset covering all general elections is provided

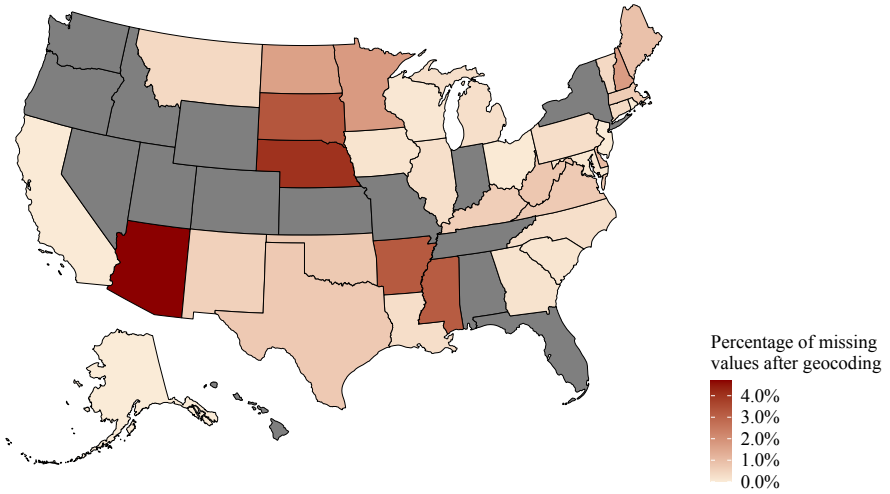


Figure 3: Percentage of missing Addresses after Geocoding by State

Note: Grey states were not included in the database

alongside all replication material on GitHub.[12]

Restricting the dataset to the elections of 2012 and 2016 reduces the dataset to 30 states and a total of 1,822 counties within them. These counties represent the final sample utilized for my analysis and are depicted in Figure 4.

Figure 4 confirms that there is no systematic undercoverage on the level of counties within states in the dataset provide by the Center for Public Integrity. To further asses the validity of the generated dataset I compare it with the survey data from the EAVS in Figure 5.

From Figure 7, we can see that the two data sources match reasonably well. If both datasets would perfectly match, regarding polling places per county, all points would fall on the angle bisector colored in red. As most points are below this line, the geocoded data from the Center of Public Integrity tends to underestimate the amount of polling places in a county, compared to the self reported numbers from the EAVS. Apart from ran-

[12] https://github.com/PhilippMartinBosch/Master_thesis

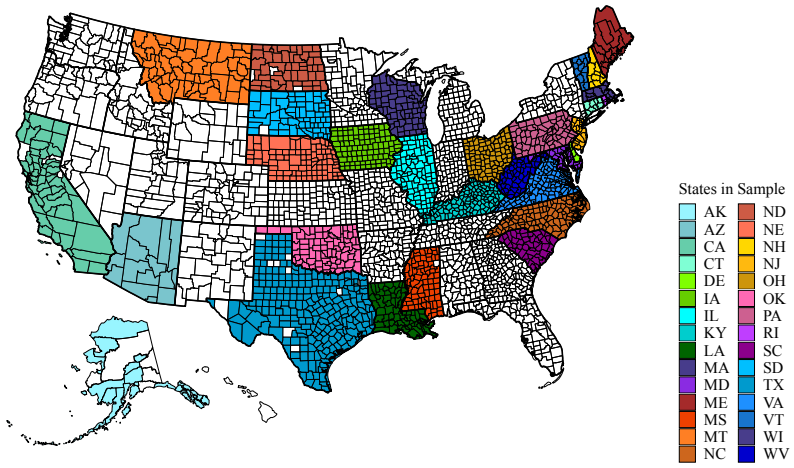


Figure 4: Counties with Polling Places for 2012 and 2016 Election in Sample

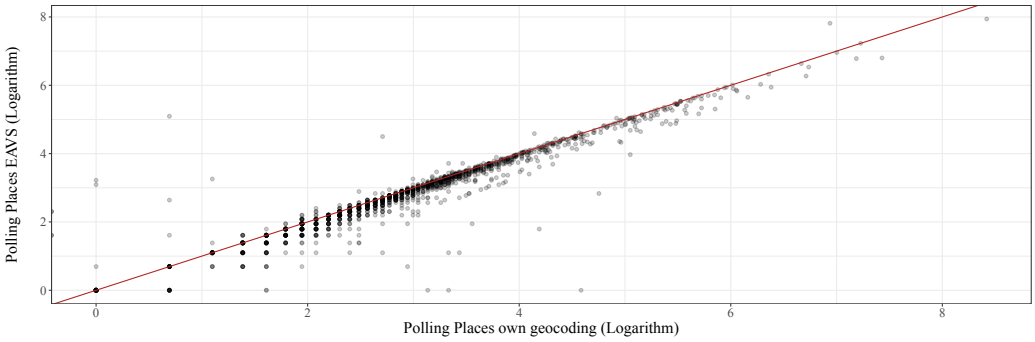


Figure 5: Comparison between the amount of Polling Places per County found in own data and EAVS data

Note: Both on logarithmic scale for better overview

dom errors in data collection, this might be due to how polling places are operationalized in the EAVS. According to the questionnaire of the EAVS, jurisdictions should provide the "number of physical polling places". If polling places of different precincts in the same county are situated in one building (for example in the biggest city of a rural county) but in different tracts or wings of the building, counties may count these polling places as sepa-

rate entities. However, during geocoding these polling places would only be counted as one entity, as they share the same address and coordinates. In general this does not pose a problem for my analysis, as I am not interested in the amount of polling places for a given year *but* the change between two elections. Nonetheless, investigating cases where both datasets diverge extensively would be fruitful for future research to learn more about the caveats of both datasets and how the data is generated. I will discuss such outliers in more detail in section 4, where I derive my empirical approach and dependent variable.

3.2 Further Variables and Final Dataset

To construct the final dataset which I utilize in my analysis, I enrich the data about polling places and counties with sociodemographic information from the American Community Survey as well as information about former VRA coverage and accessibility of convenience voting methods. To gather sociodemographic information about the citizen voting age population (CVAP), as well as the share of Black and Hispanic citizens among this CVAP, I consult the 5-year estimates of the American Community Survey. Thereby, the citizen voting age is defined as the population of a county which is over 18 years old. The American Community Survey (ACS) gets published each year by the U.S. Census Bureau, and the 5-year estimates include geographic information on all geographic ares in the United States.[13] I rely on 5-year data from the ACS, as some counties in my sample have very few housing units and the ACS does not provide reliable single-year estimates for such counties. Moreover, this ensures comparability with the data of Squires (2021) as he relies on 5-year estimates as well.

To gather information about former coverage of the Voting Rights Act I

[13] county data with shapefiles can be acessed under: https://www2.census.gov/geo/tiger/TIGER_DP/

consulted the website of the Department of Justice of the United States.[14] To operationalize access to convenience voting methods I relied on data provided by ballotpedia.org which in turn draw on data from the EAVS.[15] For the sake of clarity, I do not present summaries for each of the 1,822 individual counties in my final sample. Nevertheless, two summaries are of relevance for the following analysis. Figure 6 presents the share of Black voters in each county, broken down by state in the year 2016. This is to illustrate that counties, although they represent some form of aggregation, still yield enough variance in the independent variables of my model. A similar Figure for the Hispanic population is supplied in the appendix. In short, on average 9% of the voting age population are black across countries in my sample. Thereby, the share ranges from 0% to a maximum of 84.9% between counties.

Each boxplot in Figure 6 shows the median value of the share of Black voters for each state, together with the 25th and 75th percentiles corresponding to the lower and upper hinges. Lower and upper whiskers stretch to the largest and lowest value, but no further than 1.5× the interquartile range. More extreme outliers are depicted as points. This form of boxplot is based on the work of McGill, Tukey, and Larsen (1978).

As I suspect in my hypotheses that counties which were formerly covered by the Voting Rights Act, are more likely to close polling places, I visualized these formerly covered counties in my sample in Figure 7. Altogether, 652 counties out of 1,822 in the final sample were formerly covered by the VRA.

Having introduced the key variables to put my hypothesis to test with the novel data set, I will introduce the statistical model which I utilize to study the change of polling places between 2012 and 2016 in the next section.

[14] https://www.justice.gov/crt/jurisdictions-previously-covered-section-5
[15] https://ballotpedia.org/Analysis_of_absentee/mail-in_voting,_2016-2018

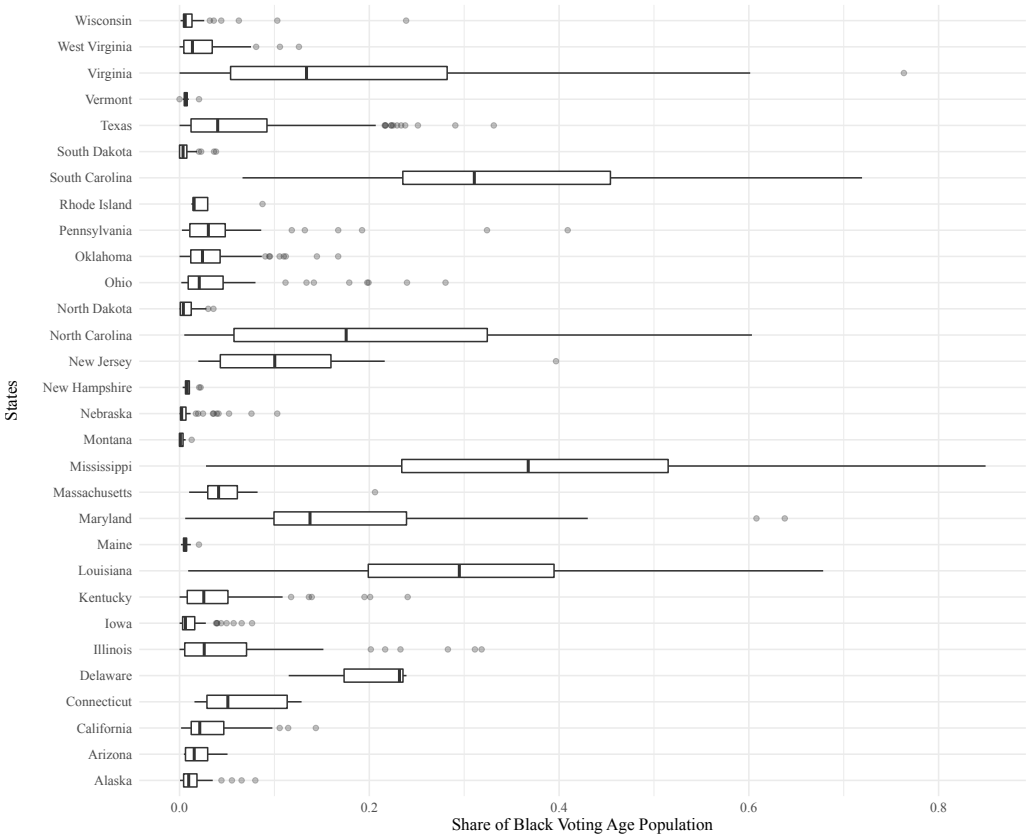


Figure 6: Distribution of the Share of Black Voters by State

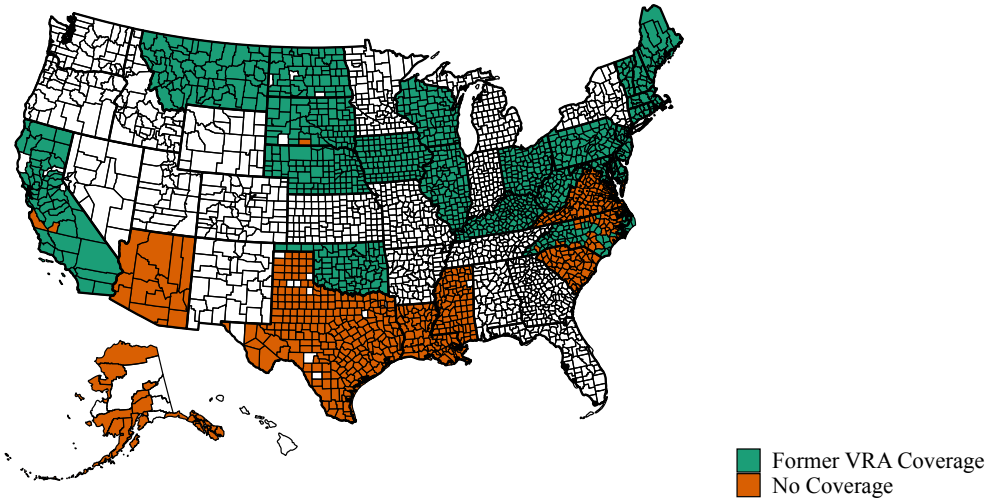


Figure 7: Map of Counties which were covered by the Voting Rights Act until 2013

County	CVAP 2012[1]	Polling Places 2012	CVAP 2016[1]	Polling Places 2016	PCVPP
Muhlenberg County, KY	24627	20	24623	10	99.97%
Hunterdon County, NJ	98115	41	99269	49	−15.34%

[1] Citizen Voting Age Population

Figure 8: PCVPP Values for two Counties from the Sample

point on the far right of California in Figure 9 corresponds to Yolo County. For this specific county there was only one polling place in the database of the Center for Public Integrity in the year 2016. However, in the year 2012 the database contained 84 polling places for a total of 164,218 citizens of voting age in Yolo County.

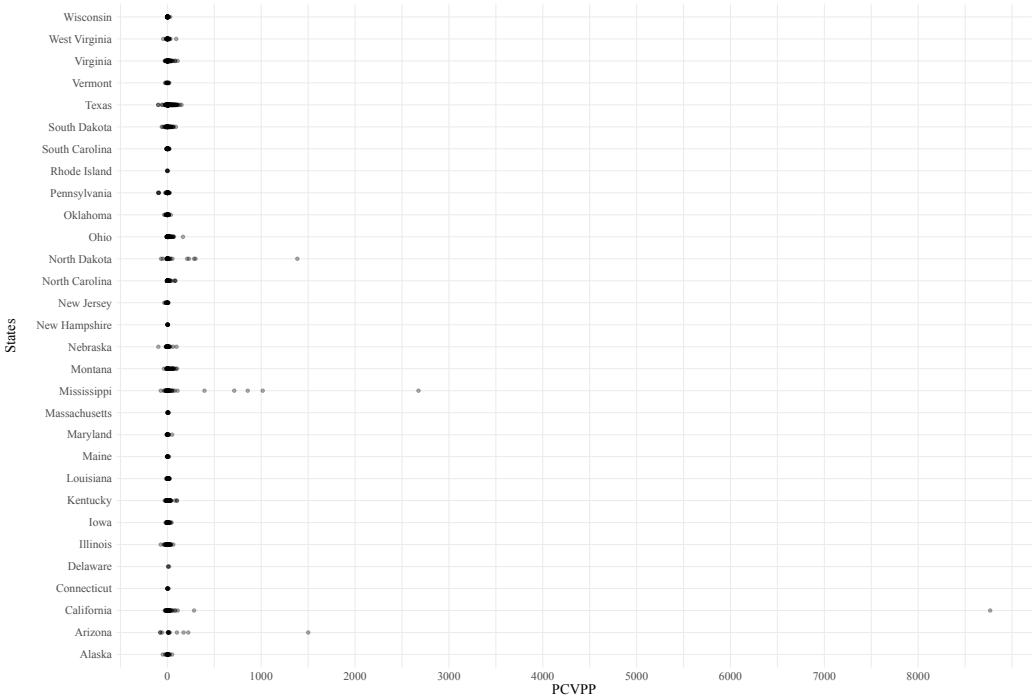


Figure 9: Distribution of PCVPP Values by State

Unfortunately, I was not able to trace down the actual reasons for such extreme discrepancies between years. Most likely, such outliers are due to

poor data quality supplied by local authorities or coding errors on the side of the Center for Public Integrity. However, as these extreme outliers in all probability are errors and therefore not actually present extreme closing of polling places, I exclude them from my sample for further analysis. After deleting 0.1% observations (18 counties) with the most extreme PCVPP values, the final sample for my analysis consists of 1,804 counties which PCVPP values range from -71.41% for Schuyler County, Illinois to 284.63% for Eddy County, North Dakota. The mean of PCVPP over all counties in my sample is 7.29%.

As the discrepancies do not speak in favor of the overall validity of the data from the Center for Public Integrity, the results of my analysis should be interpreted with caution. Nevertheless, future research, with a larger scope than a thesis project, should investigate the discrepancies which I found and use this knowledge to enhance the data collection of the Center for Public Integrity.

Still, the results of the models which I introduce below, help to advance the ongoing debate whether minority voters are disproportionately affected by poll closures.

4.2 Modeling Closures between Elections

In principle, I follow the work of Squires (2021) as I employ a linear regression model to predict the values of PCVPP across counties. Deviating from Squires (2021) however, I incorporate the fact that counties are nested in states and what is more, rely on a Bayesian approach instead of the frequentistic approach Squires (2021) chooses. Thus, I set up a Bayesian multilevel model. Using a multilevel model is reasonable as such models allow clustering and include shared group level effects (Gelman & Hill, 2009; Cheah, 2009). Especially for the data at had, counties nested in states, hierarchical models are a natural choice. In particular considering the decentralized nature of the U.S., which gives states ample influence over the way elections are

administered. Hence, the key assumption of OLS, that observations are independent and identically distributed is most likely violated. Squires (2021) incorporates such nesting structures by using clustered standard errors at the county level. Research has shown however, that clustering approaches still lead to biased standard errors, compared to multilevel models (Cheah, 2009).

I choose the Bayesian approach mostly due to the appealing and more natural interpretation of Bayesian credible intervals and the possibility to incorporate prior information in future research, when more data about polling places and elections becomes available. For this first rather explorative approach, I will rely on uninformative priors.

In formal notation the hierarchical model takes the following form:

$$\text{PCVPP}_i \sim N\left(\mu, \sigma^2\right)$$

$$\begin{aligned}\mu = \alpha_{j[i]} &+ \beta_{1j[i]}(\text{prop\_black}) + \beta_2(\text{VRA}) + \beta_{3j[i]}(\text{prop\_hisp}) \\ &+ \beta_4(\text{prop\_hisp} \times \text{VRA}) + \beta_5(\text{prop\_black} \times \text{VRA}) \\ &+ \beta_6(\text{convenience\_voting})\end{aligned}$$

Thus, I model the percentage change of voters per polling place for county i as a function of the proportion of Black voters in the county, Hispanic Voters in a county, whether a county was formerly covered by the VRA and the accessibility of convenience voting methods. Moreover, to investigate whether the effect of racial composition is stronger in formerly covered counties, I include the interaction effects β_4 and β_5. To allow for variation between states, I also include a random intercept and random slopes for β_1 and β_2 on the level of states. This is to study, whether the relationship between the closure of polling places and racial composition differs between states. Subsequently, building on Squires (2021), I scrutinize whether counties with particularly

large non-white populations exhibit more poll closures. Therefore, I substitute β_1 to β_5 with a categorical variable capturing four groups of counties. Large Black and Hispanic population under former VRA coverage, former coverage without a large share of non-white voters and counties which were never covered by the VRA.

To fit the final model and derive the posterior distributions of the model coefficients, I use the R package `brms` (Bürkner, 2017). As mentioned above, for priors I stick with the default uninformative priors of `brms`. These default priors are flat for all population level effects and use a student-t distribution with three degrees of freedom, $\mu = 2.1$ and $\sigma = 6.1$ for the intercept. The full `Stan` code generated by `brms` is supplied on GitHub. In terms of model diagnostics, I relied on trace plots to assess convergence of the MCMC algorithm, which `Stan` utilizes to derive at posterior distributions. Visual inspection as well as `brms` diagnostics output confirmed convergence of both models.

5 Results

For the sake of clarity, I present the results of both my models graphically. Figure 10 and 11 show the posterior distributions of the population level coefficients of my base model. I separate the coefficients of the racial composition variables from the coefficients associated with convenience voting methods and whether counties were formerly covered by the VRA. This is due to the varying size of coefficients and fosters readability.

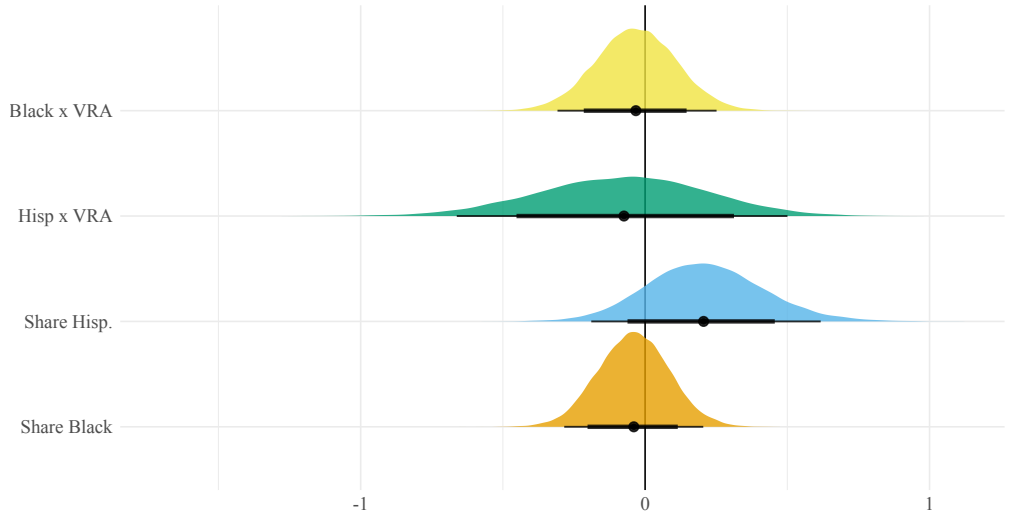


Figure 10: Posterior Distribution of Coefficients

Note: 80% and 95% credible intervals shown in black

Figure 10 shows that all coefficients associated with racial composition are distributed around zero and therefore constitute null findings. This does not support the hypothesized relationship between racial composition and poll closures. However, as Figure 11 shows, the effects of former coverage of the VRA and the accessibility of convenience voting methods are more aligned with my theoretical expectations. This finding indicates that counties which were formerly covered by the Voting Rights Act close polling places at a

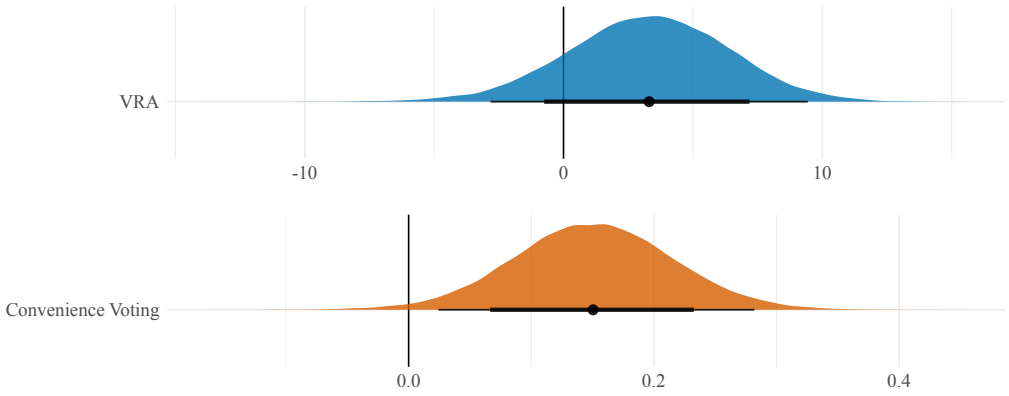


Figure 11: Posterior Distribution of Coefficients

Note: 80% and 95% credible intervals shown in black

higher rate compared to counties which were never covered. Moreover, the more accessible other forms of voting are in a state, the more likely counties close polling places, according to the findings in Figure 11. Therefore, both constitutionally valid reasons and potential discriminatory reasons seem to influence polling place closures on the county level.

Although, the coefficients regarding racial composition of counties are in contrast to my hypothesis as well as claims in news outlets about excessive closures of polling places in minority counties, they are in line with the findings of Squires (2021) and Shepherd et al. (2021). As I include both varying slopes and varying intercepts on the level of states, the effect of the share of Black and Hispanic voters can differ between states and states themselves can deviate from the population grand mean. Figure 12 shows such a random effects for the state of Arizona, where the coefficient for the share of Hispanics is distinguishable from zero. In all other states, the random slope for both Black and Hispanic population does not differ from zero.

Moreover, Figure 13 shows the random intercepts for the states of Texas, Ohio, North Dakota and Kentucky. The coefficient of the intercept for these states is different from zero and differs from the grand population mean

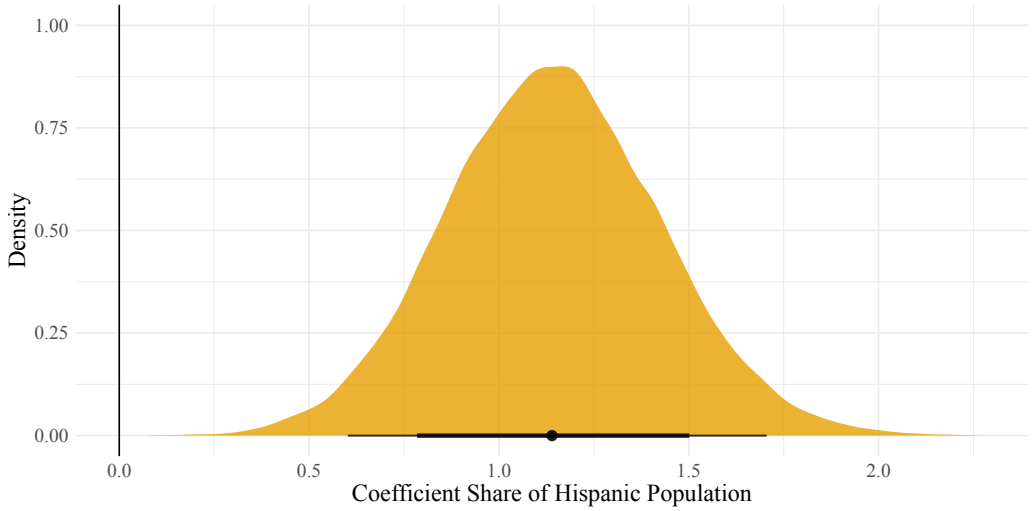


Figure 12: Posterior Distribution of Random Slope for Share of Hispanic Population in Arizona

Note: 80% and 95% credible intervals shown in black

(highlighted by the black vertical line), indicating that these states close polling places at a higher rate compared to other states in the sample. The full set of random effects for all states can be found in the appendix.

Figure 12 and 13 therefore reveal interesting patterns in the data. Although, widespread vote-suppression by means of poll closures can't be supported by my findings, some states do deviate from this pattern. What is more, it is evident that although the struggle for equal access to the polls in the past was mostly a struggle of the Black population int U.S., Figure 12 shows that the strongest effects can be found in a swing-state like Arizona for Hispanic voters. The four states that appear to deviate the most from the population intercept are not a major surprise either. Each of them enacted strict voting laws in the past, which restricted access to the polls (e.g., Weinberg, 2021; Ortiz, 2020; Smyth, 2021).

Lastly, to zoom in on counties with a high share of minorities, I fitted

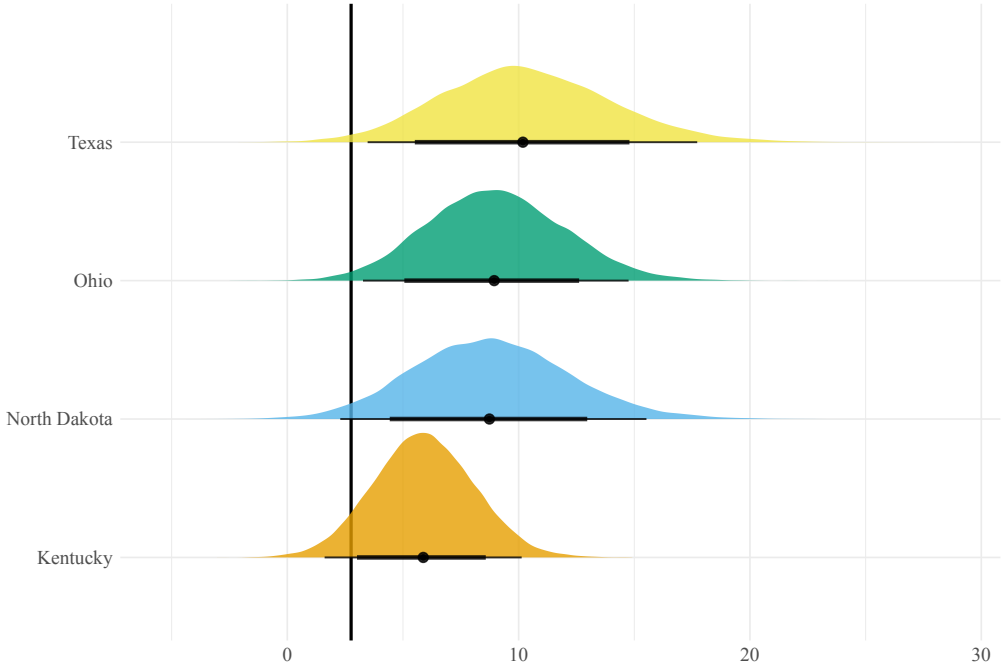


Figure 13: Posterior Distribution of Random Intercept

Note: Black vertical line corresponds to population intercept

a model which substitutes the continuous variables of minority shares with one categorical variable, capturing counties with a high share of black or Hispanic voters and whether these were covered by the VRA. This follows the rational that local election officials might target counties with exceptional high shares of minorities more extensively, and therefore the relationship is not adequately modelled with a continuous variable. Figure 14 shows the results of this model. The coefficients are in line with my previous findings, but do differ from the findings Squires (2021) generates in his thesis. Most notably, I do not find an effect for counties with a high share of Black voters. Again, the strongest effects appear in counties with a high share of Hispanic voters. The variable is coded as a factor in R. Hence, coefficients have to be interpreted in relation to the base category. This base category are counties

without former VRA coverage. The coefficient for large Hispanic counties therefore can be interpreted as follows: Compared to counties without former coverage of the VRA, counties with a large share of Hispanic voting age population have closed about 10% more polling places in relation to their total voting population. Put simply, voters in counties with a high share of Hispanic voters had to share their polling place with more people in the 2016 election.

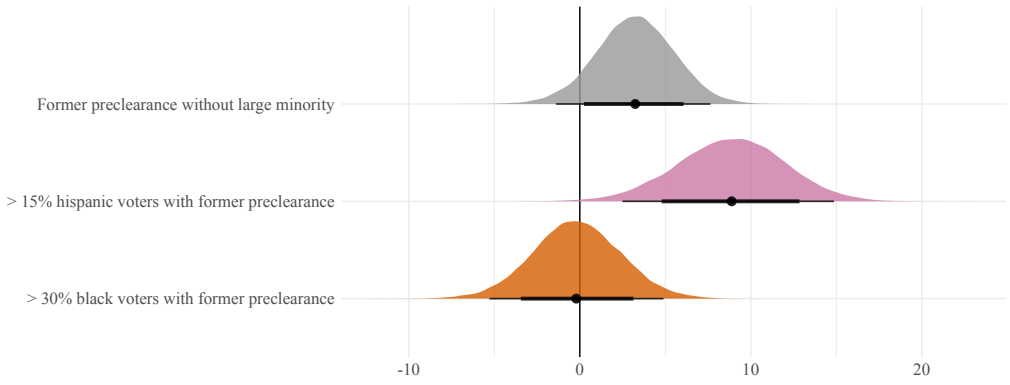


Figure 14: Posterior Distribution of Coefficients for Categorical Model

Note: 80% and 95% credible intervals shown in black

As a final remark to my results, I want to emphasize again that these findings should be treated with caution. On the one hand, my sample does not include highly debated states such as Florida and on the other, I had to exclude some implausible values during data cleaning. This fact does not speak in favor of the overall validity of the data set. Although, I did implement several validity checks during geocoding and data cleaning, I am subject to the same limitations as the researchers at the Center for Public Integrity and have to conclude that the results are just as good as the data supplied by local authorities.

6 Discussion

This thesis project had two main goals. First, I supply a novel dataset of geocoded polling places in the U.S. and second, I scrutinize whether counties with a high share of minorities are exposed to poll closures using this very dataset. By extensively engaging with current reports in news outlets and scientific evidence about poll closures, turnout and vote-suppression, I show that the debate about access to the polls in the U.S. is far from settled. Against the backdrop of this debate, I highlight the benefits such a fine-grained, large scale dataset of polling places brings to research.

However, these benefits can only be realized if such a new data source provides a valid and true picture of the reality. Data cleaning and subsequently measuring polling place changes in relation to population changes, cast some doubt on the validity of the database provided by the Center for Public Integrity. Future research therefore should engage with these doubts and scrutinize how these errors arise and whether they pose the risk to systematically bias results. Therefore, I see this thesis as a valuable first step into the direction of a valid and growing database of polling places in the United States. Moreover, it also serves as a cautionary example to engage with the data generating process when employing novel and not-yet validated sources of data for substantial research.

Due to these caveats, inference regarding the research question should be treated with caution. However, my findings align with prior research about targeting of minority voters measured across counties and states. Such research reveals that for single states such as North Carolina, no strategies of targeting can be found (Shepherd et al., 2021). This is despite ample anecdotal evidence of news outlets, claiming the contrary. On a national scale, Squires (2021) does find evidence of targeting of Black and Hispanic voters, but only in counties which were formerly covered by the VRA. My findings deviate from Squires (2021) as I do not find evidence of the targeting of Black voters but strong evidence for Hispanics. Moreover, the hierarchical

approach allows me to differentiate between states and highlights significant differences between clusters.

My conclusion and advice for further research therefore is to revisit the geocoded polling places which I supply and manually research cases with implausible values. My measure of polling place changes between elections serves as a first indicator where to look for inconsistencies. After clearing these last doubts about the validity of the data, many analyses and research questions are imaginable that go far beyond my first rather explorative approach.

For example, keeping the focus on a national level, the amount of relocations could be measured by evaluation which coordinates of polling places remain stable between elections. The dependent variable in such a model would be the proportion of polling places that move between elections. In statistical terms, this relationship could be modelled with a zero-inflated beta regression model. Whereby at first a logistic regression model is fit to assess whether there are any relocations between elections and in a second step, a beta regression is used to model the share of relocations. Although the data at hand already would allow for such a model, I refrained from implementing it due to the limit scope of a thesis and the mentioned doubts about the validity of the data.

Another approach how to utilize the full sample of polling places would be to measure the accessibility of each polling place. Therefore, not even fine-grained individual voter file data nor precinct borders would be necessary. In a first step, one could calculate the geographic distance from each polling place to the nearest public transportation stop. This would add to the debate whether pure distance to the polling place (Brady & McNulty, 2011; Gimpel & Schuknecht, 2003) lowers turnout for individual voters, or if this relationship is moderated by the availability of public transport (Barreto, Cohen-Marks, & Woods, 2009).

Finally, to leverage the full potential such a fine-grained dataset possesses,

it would be fruitful to enrich it with voter file data from different states. This would allow researchers to study the effects of poll closures over time on individual voting behavior, and thus avoid the risk of ecological fallacy. Furthermore, picking up a point both Squires (2021) and I stress out in our work, more information about the partisanship of local election officials is needed to study the interplay of partisan targeting and poll closures. Future research should aim to create a new edition of the dataset that (Kimball & Kropf, 2006) provide.

To conclude, I can only echo the words of Charles Stewart III, professor of Political Science at the MIT, about the new database of polling places:"Knowing where these polling places are will actually be really valuable" (Levine, Rebala, & Vasilogambros, 2020). The emphasis hereby is on *will*. My thesis offers a promising start and glimpse into this new data, but more steps need to be taken to establish this data source. All the more, as COVID-19 completely changed the way elections are administered. Vote by mail and early voting are on the rise in some states, while other states might use the pandemic as an excuse to further close polling places while tightening the rules for convenience voting (Curiel & Clark, 2021). As of the writing of these words, new data has been released by the Center for Public Integrity to cover polling places in the 2020 election. Equipped with the knowledge provided by this thesis about the caveats and possibilities of this data, valuable insights about the interplay of voting behavior and the location of polling places should be gained.

References

Abrams, S., Anderson, C., Kruse, K. M., Richardson, H. C., & Thompson, H. A. (2020). *Voter Suppression in US Elections*. University of Georgia Press.

Akee, R., Copeland, W., Costello, E. J., Holbein, J., & Simeonova, E. (2018). *Family Income and the Intergenerational Transmission of Voting Behavior: Evidence from an Income Intervention* (Working Paper No. 24770).

Amos, B., Smith, D. A., & Ste. Claire, C. (2017). Reprecincting and Voting Behavior. *Political Behavior*, *39*(1), 133–156. doi: 10.1007/s11109-016-9350-z

Anderson, C. (2018). *One person, no vote: How voter suppression is destroying our democracy*. Bloomsbury publishing USA.

Astor, M. (2018). *Seven Ways Alabama Has Made It Harder to Vote* . The New York Times. Retrieved 2022-02-24, from https://www.nytimes.com/2018/06/23/us/politics/voting-rights-alabama.html

Barreto, M. A., Cohen-Marks, M., & Woods, N. D. (2009). Are all precincts created equal?: The prevalence of low-quality precincts in Low-Income and Minority Communities. *Political Research Quarterly*, *62*(3), 445–458. doi: 10.1177/1065912908319572

Bishop, B. (2009). *The big sort: Why the clustering of like-minded America is tearing us apart*. Houghton Mifflin Harcourt.

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. doi: 10.1017/S0003055410000596

Brady, H. E., Verba, S., & Schlozman, K. L. (1995). Beyond SES: A Resource Model of Political Participation. *American Political Science Review*, *89*(2), 271–294. doi: 10.2307/2082425

Burden, B. C., Canon, D. T., Lavertu, S., Mayer, K. R., & Moynihan, D. P. (2013). Selection Method, Partisanship, and the Administration of

Elections. *American Politics Research*, *41*(6), 903–936. doi: 10.1177/1532673X12472365

Burden, B. C., Fletcher, J. M., Herd, P., Jones, B. M., & Moynihan, D. P. (2017). How different forms of health matter to political participation. *Journal of Politics*, *79*(1), 166–178.

Bürkner, P.-C. (2017). brms: An r package for bayesian multilevel models using stan. *Journal of Statistical Software*, *80*(1), 1–28. Retrieved from https://www.jstatsoft.org/index.php/jss/article/view/v080i01 doi: 10.18637/jss.v080.i01

Cambon, J., Hernangómez, D., Belanger, C., & Possenriede, D. (2021). tidygeocoder: An r package for geocoding. *Journal of Open Source Software*, *6*(65), 3544. Retrieved from https://doi.org/10.21105/joss.03544 (R package version 1.0.5) doi: 10.21105/joss.03544

Campbell, A., Converse, P. E., Miller, W. E., & Stokes, D. E. (1960). The American Voter.

Cantoni, E. (2020). A precinct too far: Turnout and voting costs. *American Economic Journal: Applied Economics*, *12*(1), 61–85. doi: 10.1257/app.20180306

Cheah, B. C. (2009). Clustering standard errors or modeling multilevel data.

Chen, M. K., Haggag, K., Pope, D. G., & Rohla, R. (2020). Racial Disparities in Voting Wait Times: Evidence from Smartphone Data. *The Review of Economics and Statistics*, 1–27. doi: 10.1162/rest_a_01012

Clinton, J. D., Eubank, N., Fresh, A., & Shepherd, M. E. (2021). Polling place changes and political participation: Evidence from North Carolina presidential elections, 2008-2016. *Political Science Research and Methods*, *9*(4), 800–817. doi: 10.1017/psrm.2020.43

Curiel, J. A., & Clark, J. T. (2021). Disparities in Poll Closures in the Age of COVID-19: A Case Study of Wisconsin. *Election Law Journal: Rules, Politics, and Policy*, *20*(4), 345–363. doi: 10.1089/elj.2020.0669

Davidson, C. (1992). The voting rights act: A brief history. In B. Grofman

& C. Davidson (Eds.), *Controversies in Minority Voting* (pp. 7–34). Washington, DC: Brookings Institution Washington, DC.

Deford, D. R., Eubank, N., & Rodden, J. (2021). Partisan Dislocation: A Precinct-Level Measure of Representation and Gerrymandering. *Political Analysis*. doi: 10.1017/pan.2021.13

DellaVigna, S., & Kaplan, E. (2007). The Fox News Effect: Media Bias and Voting. *The Quarterly Journal of Economics*, *122*(3), 1187–1234. doi: 10.1162/qjec.122.3.1187

DellaVigna, S., List, J. A., Malmendier, U., & Rao, G. (2017). Voting to tell others. *Review of Economic Studies*, *84*(1), 143–181. doi: 10.1093/restud/rdw056

Downs, A. (1957). *An Economic Theory of Democracy*. Harper & Row.

Fausset, R. (2014). *Mistrust in North Carolina Over Plan to Reduce Precincts* . The New York Times. Retrieved 2022-03-02, from https://www.nytimes.com/2014/07/08/us/08northcarolina.html

Fiorina, M. P. (1976). The Voting Decision: Instrumental and Expressive Aspects. *The Journal of Politics*, *38*(2), 390–413. doi: 10.2307/2129541

Fraga, B. L. (2018). *The turnout gap: Race, ethnicity, and political inequality in a diversifying america*. Cambridge University Press.

Fresh, A. (2018). The Effect of the Voting Rights Act on Enfranchisement: Evidence from North Carolina. *Journal of Politics*, *80*(2), 713–718. doi: 10.1086/695852

Gelman, A., & Hill, J. (2009). *Data analysis using regression and multilevel/hierarchical models* (Repr. with corr., 11. print. ed.). Cambridge [u.a.]: Cambridge Univ. Press.

Gerber, A. S., Green, D. P., & Larimer, C. W. (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review*, *102*(1), 33–48. doi: 10.1017/S000305540808009X

Gerber, A. S., Green, D. P., & Shachar, R. (2003). Voting may be

habit-forming: Evidence from a randomized field experiment. *American Journal of Political Science*, *47*(3), 540–550. doi: 10.1111/1540-5907.00038

Gimpel, J. G., & Schuknecht, J. E. (2003). Political participation and the accessibility of the ballot box. *Political Geography*, *22*(5), 471–488. doi: 10.1016/S0962-6298(03)00029-5

Greenblatt, A. (2013). *The Racial History Of The 'Grandfather Clause'.* National Public Radio. Retrieved 2022-03-09, from https://www.npr.org/sections/codeswitch/2013/10/21/239081586/the-racial-history-of-the-grandfather-clause

Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2011). *Survey methodology*. John Wiley & Sons.

Grumbach, J. M. (2021). *Laboratories of Democratic Backsliding* (Working Paper).

Haspel, M., & Knotts, G. (2005). Location, location, location: Precinct placement and the costs of voting. *Journal of Politics*, *67*(2), 560–573. doi: 10.1111/j.1468-2508.2005.00329.x

Holbein, J. B. (2016). Left Behind? Citizen Responsiveness to Government Performance Information. *American Political Science Review*, *110*(2), 353–368. doi: 10.1017/S0003055416000071

Holbein, J. B., & Schafer, J. P. (2017). *Time Zones, Tiredness, and Turnout: A Natural Experiment on How Time Constraints Influence Elections* (Working Paper No. 2881452).

Ingraham, C. (2016). *The 'smoking gun' proving North Carolina Republicans tried to disenfranchise black voters .* The Washington Post. Retrieved 2022-02-23, from https://www.washingtonpost.com/news/wonk/wp/2016/07/29/the-smoking-gun-proving-north-carolina-republicans-tried-to-disenfranchise-black-voters/

Jacobs, J. (2018). *In North Carolina, Voting Controversies Are Common. Here's the Recent History.* The New York Times. Retrieved 2022-

03-05, from https://www.nytimes.com/2018/12/13/us/politics/north-carolina-voting-rights.html

Jones, B. D., Greenberg, S. R., Kaufman, C., & Drew, J. (1978). Service delivery rules and the distribution of local government services: Three detroit bureaucracies. *The Journal of Politics*, *40*(2), 332–368.

Keyssar, A. (2009). *The right to vote: The contested history of democracy in the United States*. Basic Books.

Kimball, D. C., & Baybeck, B. (2013). Are All Jurisdictions Equal? Size Disparity in Election Administration. *Election Law Journal: Rules, Politics, and Policy*, *12*(2), 130–145. doi: 10.1089/elj.2012.0174

Kimball, D. C., & Kropf, M. (2006). The street-level bureaucrats of elections: Selection methods for local election officials. *Review of Policy Research*, *23*(6), 1257–1268. doi: 10.1111/j.1541-1338.2006.00258.x

King, D. (2017). Forceful federalism against American racial inequality. *Government and Opposition*, *52*(2), 356–382. doi: 10.1017/gov.2016.52

Krogstad, J. M., & Lopez, M. H. (2017). *Black voter turnout fell in 2016, even as a record number of Americans cast ballots.* Pew Research Center. Retrieved 2022-02-24, from https://www.pewresearch.org/fact-tank/2017/05/12/black-voter-turnout-fell-in-2016-even-as-a-record-number-of-americans-cast-ballots/

Lassen, D. D. (2005). The Effect of Information on Voter Turnout: Evidence from a Natural Experiment. *American Journal of Political Science*, *49*(1), 103–118. doi: 10.1111/J.0092-5853.2005.00113.X

Leighley, J. E., & Nagler, J. (2013). *Who votes now?* Princeton University Press.

Levendusky, M. S. (2009). *The Partisan Sort: How Liberals Became Democrats and Conservatives Became Republicans*. University of Chicago Press. Retrieved from https://doi.org/10.7208/9780226473673 doi: doi:10.7208/9780226473673

Levine, C., Rebala, P., & Vasilogambros, M. (2020). *National data release sheds light on past polling place changes.* The Center for Public Integrity. Retrieved 2022-03-14, from https://publicintegrity.org/politics/elections/ballotboxbarriers/data-release-sheds-light-on-past-polling-place-changes/

Levitsky, S., & Ziblatt, D. (2018). *How democracies die* (First ed.). New York, NY: Crown.

Lublin, D. (1997). *The Paradox of Representation: Racial Gerrymandering and Minority Interests in Congress*. Princeton University Press. doi: 10.1515/9780691221397

Marshall, T. (1957). The Rise and Collapse of the "White Democratic Primary". *Journal of Negro Education*, *26*(3), 249–254.

May, G. (2013). *Bending toward justice: the Voting Rights Act and the transformation of American democracy*. Basic Books.

McGill, R., Tukey, J. W., & Larsen, W. A. (1978). Variations of box plots. *American Statistician*, *32*(1), 12–16. doi: 10.1080/00031305.1978.10479236

Mickey, R. (2015). *Paths out of dixie: The democratization of authoritarian enclaves in america's deep south, 1944-1972*. Princeton University Press. doi: doi:10.1515/9781400838783

Mladenka, K. R. (1980). The urban bureaucracy and the chicago political machine: Who gets what and the limits to political control. *American Political Science Review*, *74*(4), 991–998.

Nall, C. (2015). The political consequences of spatial policies: How interstate highways facilitated geographic polarization. *Journal of Politics*, *77*(2), 394–406. doi: 10.1086/679597

Nyhan, B., Skovron, C., & Titiunik, R. (2017). Differential Registration Bias in Voter File Data: A Sensitivity Analysis Approach. *American Journal of Political Science*, *61*(3), 744–760. doi: 10.1111/ajps.12288

Ortiz, E. (2020). *North Dakota, native tribes agree to settle voter ID lawsuit*

to combat voter suppression. NBC News. Retrieved 2022-03-14, from https://www.nbcnews.com/news/us-news/north-dakota-native-tribes-settle-voter-id-lawsuit-combat-voter-n1137141

Pettigrew, S. (2017). The racial gap in wait times: Why minority precincts are underserved by local election officials. *Political Science Quarterly*, *132*(3), 527–547. doi: 10.1002/polq.12657

Pettigrew, S. (2021). The downstream consequences of long waits: How lines at the precinct depress future turnout. *Electoral Studies*, *71*(July 2020), 102188. doi: 10.1016/j.electstud.2020.102188

Plutzer, E. (2002). Becoming a habitual voter: Inertia, resources, and growth in young adulthood. *American Political Science Review*, *96*(1), 41–56. doi: 10.1017/S0003055402004227

R Core Team. (2021). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from https://www.R-project.org/

Rebala, P., Levine, C., Hernández, K., Kaneya, R., Ellerbeck, A., Johnston, T., ... Henderson, T. (2022). *U.S. Polling Places (2012-2020).* https://github.com/PublicI/us-polling-places. The Center for Public Integrity.

Riker, W. H., & Ordeshook, P. C. (1968). A Theory of the Calculus of Voting. *American Political Science Review*, *62*(1), 25–42. doi: 10.2307/1953324

Riker, W. H., & Ordeshook, P. C. (1973). *An introduction to positive political theory* (Vol. 387). Prentice-Hall Englewood Cliffs, NJ.

Roth, Z. (2015). *Study: North Carolina Polling Site Changes Hurt Blacks* . NBC News. Retrieved 2022-02-10, from https://www.nbcnews.com/news/nbcblk/study-north-carolina-polling-site-changes-hurt-blacks-n468251

Rothstein, R. (2017). *The color of law: A forgotten history of how our government segregated america.* Liveright Publishing.

Schlozman, K. L., Brady, H. E., & Verba, S. (2018). *Unequal and Unrepresented: Political Inequality and the People's Voice in the New Gilded Age*. Princeton University Press. doi: doi:10.23943/9781400890361

Schuit, S., & Rogowski, J. C. (2017). Race, Representation, and the Voting Rights Act. *American Journal of Political Science*, *61*(3), 513–526. doi: 10.1111/ajps.12284

Shepherd, M. E., Fresh, A., Eubank, N., & Clinton, J. D. (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. doi: 10.1089/elj.2019.0602

Simpson, S. (2016). The Great Poll Closure. *The Leadership Conference Education Fund*(November). Retrieved from http://civilrightsdocs.info/pdf/reports/2016/poll-closure-report-web.pdf

Smyth, J. C. (2021). *Ohio is latest state to see GOP-backed voting law rewrite.* AP News. Retrieved 2022-03-14, from https://apnews.com/article/ohio-voter-registration-voting-voting-rights-laws-49d0138b74397f74b2dc0853bb09f540

Sondheimer, R. M., & Green, D. P. (2010). Using Experiments to Estimate the Effects of Education on Voter Turnout. *American Journal of Political Science*, *54*(1), 174–189. doi: 10.1111/J.1540-5907.2009.00425.X

Squires, J. M. (2021). Shutting the Door on Voting: The Effects of "The Great Poll Closure". *Graduate Theses, Dissertations, and Problem Reports*, *8294*.

Tomkins, S., Yao, K., Gaebler, J., Konitzer, T., Rothschild, D., Meredith, M., & Goel, S. (2021). *Blocks as geographic discontinuities: The effect of polling place assignment on voting.*

Valelly, R. M. (2004). *The Two Reconstructions: The Struggle for Black Enfranchisement*. Chicago: University of Chicago Press. doi: 10.7208/chicago/9780226845272.001.0001

Vasilogambros, M. (2018). *Polling Places Remain a Target Ahead of*

November Elections . Pew Trusts. Retrieved 2022-02-10, from https://www.pewtrusts.org/en/research-and-analysis/blogs/stateline/2018/09/04/polling-places-remain-a-target-ahead-of-november-elections

Verba, S., Schlozman, K. L., Brady, H., & Nie, N. H. (1993). Race, Ethnicity and Political Resources: Participation in the United States. *British Journal of Political Science*, *23*(4), 453–497. doi: 10.1017/S0007123400006694

Verba, S., Schlozman, K. L., & Brady, H. E. (1995). *Voice and equality: Civic voluntarism in American politics*. Harvard University Press.

Waldman, M. (2016). *The fight to vote*. Simon and Schuster.

Washington, E. (2006). How black candidates affect voter turnout. *Quarterly Journal of Economics*, *121*(3), 973–998. doi: 10.1162/qjec.121.3.973

Weinberg, D. (2021). *Texas's Proposed Voter Suppression Law.* Brennan Center for Justice. Retrieved 2022-03-14, from https://www.brennancenter.org/our-work/analysis-opinion/texass-proposed-voter-suppression-law

Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, *4*(43), 1686. doi: 10.21105/joss.01686

Wilson, R. (2017). *Voter turnout dipped in 2016, led by decline among blacks.* The Hill. Retrieved 2022-02-24, from https://thehill.com/homenews/campaign/332970-voter-turnout-dipped-in-2016-led-by-decline-among-blacks

Yoder, J. (2018). How Polling Place Changes Reduce Turnout: Evidence from Administrative Data in North Carolina. *SSRN Electronic Journal*. doi: 10.2139/ssrn.3178184

Appendices

A Covered Elections in full Data

State	2012	2014	2016	2018
Alabama	✗	✗	✗	✗
Alaska	✓	✓	✓	✓
Arizona	✓	✓	✓	✓
Arkansas	✓	✓	✗	✓
California	✓	✓	✓	✓
Colorado	✗	✗	✗	✗
Connecticut	✓	✓	✓	✓
Delaware	✓	✓	✓	✓
Florida	✗	✗	✗	✗
Georgia	✗	✗	✓	✓
Hawaii	✗	✗	✗	✗
Idaho	✗	✗	✗	✗
Illinois	✓	✓	✓	✓
Indiana	✗	✗	✗	✗
Iowa	✓	✓	✓	✓
Kansas	✗	✗	✗	✗
Kentucky	✓	✓	✓	✓
Louisiana	✓	✓	✓	✓
Maine	✓	✓	✓	✓
Maryland	✓	✓	✓	✓
Massachusetts	✓	✓	✓	✓
Michigan	✓	✓	✗	✗
Minnesota	✗	✓	✓	✓
Mississippi	✓	✓	✓	✓
Missouri	✗	✗	✗	✗

Figure 15: Overview of covered Elections and States

Note: Continued on the next page

State	2012	2014	2016	2018
Montana	✔	✔	✔	✔
Nebraska	✔	✔	✔	✔
Nevada	✖	✖	✖	✖
New Hampshire	✔	✔	✔	✔
New Jersey	✔	✔	✔	✔
New Mexico	✖	✔	✔	✔
New York	✖	✖	✖	✖
North Carolina	✔	✔	✔	✔
North Dakota	✔	✔	✔	✔
Ohio	✔	✔	✔	✔
Oklahoma	✔	✔	✔	✔
Oregon	✖	✖	✖	✖
Pennsylvania	✔	✔	✔	✔
Rhode Island	✔	✔	✔	✔
South Carolina	✔	✔	✔	✔
South Dakota	✔	✔	✔	✔
Tennessee	✖	✖	✖	✖
Texas	✔	✔	✔	✔
Utah	✖	✖	✖	✖
Vermont	✔	✔	✔	✔
Virginia	✔	✔	✔	✔
Washington	✖	✖	✖	✖
West Virginia	✔	✔	✔	✔
Wisconsin	✔	✔	✔	✔
Wyoming	✖	✖	✖	✖

Figure 16: Overview of covered Elections and States

B Summary Tables of Missing and Manually coded Data

State	Total Observations	OSM	Google	Missing after geocoding	Percentage missing	Missing after handcoding
Alaska	1584	802	782	0	0.00	0
Arkansas	14546	0[1]	14546	468	3.22	80
Arizona	5377	4296	1081	254	4.72	52
California	58979	48799	10180	10	0.02	1
Connecticut	3135	2857	278	11	0.35	0
Delaware	1726	1124	602	18	1.04	0
Georgia	5351	2820	2531	9	0.17	0
Iowa	6731	5510	1221	10	0.15	0
Illinois	42220	0[1]	42220	108	0.26	0
Kentucky	14799	8910	5889	92	0.62	37
Louisiana	16105	0[1]	16105	43	0.27	4
Massachusetts	8699	0[1]	8699	59	0.68	4
Maryland	8208	6270	1938	9	0.11	9
Maine	2311	0[1]	2311	21	0.91	3
Michigan	9700	4853	4847	25	0.26	0
Minnesota	13979	8986	4993	247	1.77	21
Mississippi	7029	2417	4612	223	3.17	102

[1] Note: Sending addresses to OSM failed multiple times due to connection problems, these states were therefore geocoded solely using Google

Figure 17: Summary Table of Geocoding Process

Note: Continued on the next page

State	Total Observations	OSM	Google	Missing after geocoding	Percentage missing	Missing after handcoding
Montana	2845	909	1936	11	0.39	0
North Carolina	10892	7155	3737	27	0.25	0
North Dakota	2027	1259	768	33	1.63	27
Nebraska	4744	2605	2139	192	4.05	160
New Hampshire	1281	936	345	22	1.72	0
New Jersey	31634	26278	5356	0	0.00	0
New Mexico	1439	677	762	8	0.56	0
Ohio	36266	23786	12480	10	0.03	0
Oklahoma	7796	3857	3939	58	0.74	16
Pennsylvania	36428	16093	20335	99	0.27	7
Rhode Island	1911	1847	64	0	0.00	0
South Carolina	8909	6009	2900	12	0.13	0
South Dakota	2704	1296	1408	89	3.29	28
Texas	26006	12954	13052	189	0.73	3
Virginia	9694	4627	5067	69	0.71	0
Vermont	1057	362	695	4	0.38	0
Wisconsin	27218	17765	9453	22	0.08	0
West Virginia	7212	3641	3571	57	0.79	7

Figure 18: Summary Table of Geocoding Process

C Distribution of Share of Hispanic Voting Age Population by State

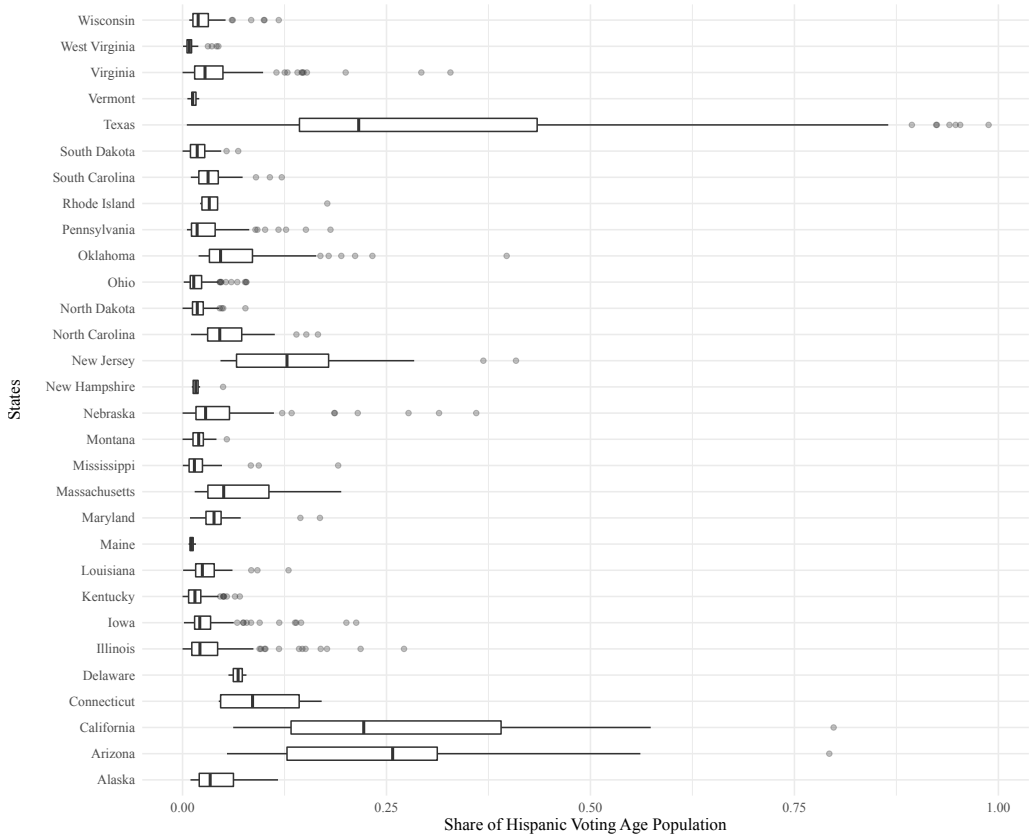


Figure 19: Distribution of Share of Hispanic Voting Age Population

D Posterior Distribution for base Model - all States

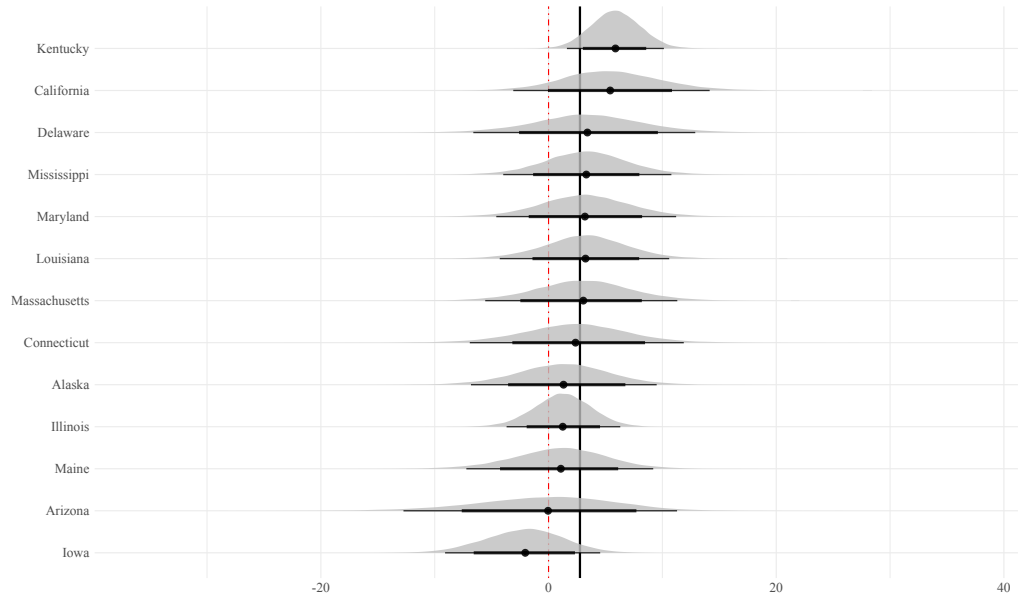


Figure 20: Posterior Distribution of Random Intercept

Note: 80% and 95% credible intervals shown in black. Black vertical line corresponds to population intercept and red dashed line indicates 0

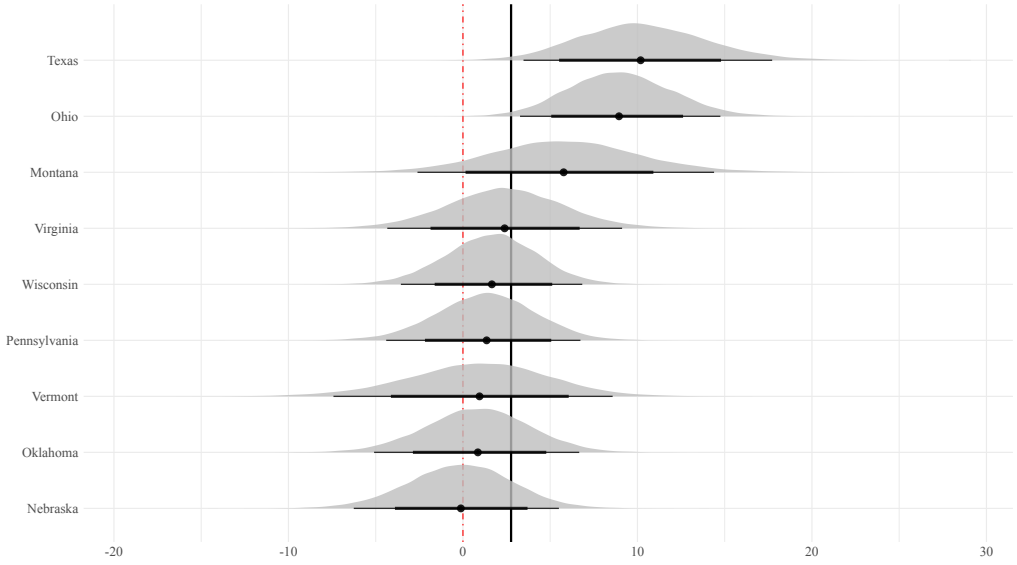


Figure 21: Posterior Distribution of Random Intercept

Note: 80% and 95% credible intervals shown in black. Black vertical line corresponds to population intercept and red dashed line indicates 0