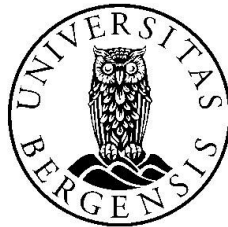


The Effect of Voter ID Laws on Turnout

A Counterfactual Analysis

Bendik Emil Basberg



Master's thesis

Spring 2021

Department of Comparative Politics
University of Bergen

Abstract

In the United States, laws requiring voters to show ID in order to vote are both a novel and highly controversial policy intervention. Disagreement over these laws largely centers on their expected effect on turnout, with opponents arguing that ID-requirements will deter voters and proponents arguing that they will not. However, there is reason to suspect that these laws are in part adopted strategically to gain electoral advantage through lower turnout among supporters of the opposing party – in turn, backlash against voter ID similarly represents the opposing party safeguarding their electoral interests. If so, voter ID-laws are unanimously expected to lower turnout. These expectations are supported by rational choice theory, in that ID-requirements represent an increase in the costs of voting and, accordingly, should make those lacking valid ID less likely to vote. Research in the field has proven inconclusive, though many studies find at least partial evidence of a negative effect. The public and academic disagreement on the issue presents an opportunity and a need for a more stringent causal research design.

Leveraging data on state-level turnout in US presidential elections between 1980 and 2020, I employ a synthetic control approach, using the matrix completion method to estimate turnout rates in voter ID-states over time in the counterfactual scenario in which they did not implement ID-requirements. Comparing actual turnout to this counterfactual turnout allows assessment of the causal impact of the intervention. Additionally, I investigate whether racial and ethnic minorities are affected more strongly than non-minority voters, using a difference-in-differences analysis of individual-level survey data. The analysis does not indicate that turnout rates in states with strict photographic ID-laws deviated significantly from what they otherwise would have been, neither overall nor among the demographic groups considered. The overall average treatment effect on the treated is estimated to lie between -2.86 and 5.03 percentage points change in turnout, though there is variation among individual states. Although the possibility of a small effect of voter ID-requirements on turnout cannot be ruled out, this study finds no significant evidence that one exists.

Acknowledgements

Som seg hør og bør vil jeg først og fremst takke mine foreldre. I tillegg til diverse hjelp og støtte gjennom studietiden ville jeg av åpenbare grunner neppe ha kunnet gjennomført en mastergrad – eller spesielt mye annet – om det ikke hadde vært for dem. Videre retter jeg en stor takk til min veileder, Tor Midtbø, for konstruktive kommentarer, tilgjengelighet for spørsmål og veiledning, og for å ved en rekke anledninger ha hjulpet meg tilbake på rett spor når arbeidet med oppgaven gikk i stå. Takk også til Michael Alvarez for all vennlighet og veiledning gjennom hele studieløpet ved SAMPOL, samt for hans store engasjement for sitt fag og sine studenter. En særlig takk til Ruben Mathisen for all hjelp og inspirasjon han har bidratt med – som regel over en gastronomisk utskeielse i lunsjpausen – og for godt vennskap både i og utenfor academia. Det er også på sin plass å rette en takk til min venn og selverklærte mentor, Vegard Kolnes. Foruten hans distingverte karrieropolitiske rådgivning hadde jeg ikke studert sammenliknende politikk i det hele tatt. Til slutt: tusen takk til alle små og store på Kronstad skole. Dere betyr mer enn dere aner.

Table of contents

ABSTRACT	II
ACKNOWLEDGEMENTS	III
TABLE OF CONTENTS	IV
LIST OF TABLES AND FIGURES	VI
1 INTRODUCTION	1
1.1 Why study voter ID-laws?	2
1.2 A novel contribution	3
1.3 Structure of the thesis	4
2 VOTER ID: CONTEXT, CONCEPT AND CONTROVERSY	6
2.1 Voting in the United States	6
2.2 Voter ID laws	10
2.3 The debate	13
2.3.1 The partisan angle	14
3 THEORY AND PREVIOUS RESEARCH	17
3.1 Voter ID and turnout – theoretical expectations	17
3.1.1 The calculus of voting	17
3.1.2 Voter ID-requirements and the costs of voting	19
3.1.3 Alternative theories of voting and counteracting forces	21
3.2 Hypotheses	24
3.3 Previous research on voter ID-laws	25
3.3.1 Empirical plausibility	25
3.3.2 Previous studies on the effect of voter ID-laws on turnout	27
4 METHODOLOGY	31
4.1 Causal effects and the problem of counterfactual outcomes	31
4.1.1 Dealing with the counterfactual problem	32
4.2 The synthetic control method	33
4.2.1 Generalized synthetic control	35
4.2.2 Matrix completion method	36
4.2.3 Assumptions	37
4.2.4 Implementation	39
4.3 Difference-in-differences	40
4.3.1 Implementation	42

5 DATA.....	44
5.1 Dataset – Synthetic control.....	44
5.1.1 Dependent variable	45
5.1.2 Treatment variable	46
5.2 Dataset – Difference-in-differences.....	47
5.2.1 Why not use the CPS data for synthetic control?.....	48
5.2.2 Dependent variable	49
5.2.3 Race/ethnicity variable.....	51
5.2.4 Treatment indicator	52
6 RESULTS.....	53
6.1 Do voter ID-requirements lead to lower turnout? Results from the synthetic control analysis	53
6.1.1 The effect of strict photographic voter ID-requirements on turnout.....	54
6.1.2 Robustness checks	59
6.1.3 Alternative analysis – all election years.....	62
6.2 Do voter ID-requirements disproportionately affect minorities? Results from the difference-in-differences analysis.....	64
6.2.1 Assumption check: parallel trends	66
6.2.2 Robustness checks - DID	67
7 DISCUSSION AND CONCLUSION.....	70
7.1 Evaluating the hypotheses and answering the research question	70
7.2 Why do we not find an effect of voter ID-requirements on turnout?	71
7.3 Implications	74
7.3.1 Implications for research.....	75
7.3.2 Implications for theory.....	77
7.3.3 Implications for voter ID.....	78
7.3.4 Implications for policy	79
7.4 Limitations.....	80
7.5 Further research	82
REFERENCES	84
APPENDIX	94
A.1 Factor & loading from generalized synthetic control	94
A.2 Robustness checks for analysis with all elections	96

List of tables and figures

Figure 2.1 <i>National turnout rates in presidential elections, 1789-2020</i>	8
Table 2.1 <i>Voter ID-laws in effect in 2020</i>	11
Table 2.2 <i>States with strict photo ID-requirements, by election</i>	12
Table 2.3 <i>Adoption of strict photo ID-laws and party control of state government</i>	15
Figure 5.1 <i>Treatment status by state and election</i>	47
Figure 6.1 <i>Turnout trends, treated versus untreated states</i>	54
Table 6.1 <i>State-level turnout rates</i>	54
Figure 6.2 <i>Treated and counterfactual average turnout over time</i>	55
Figure 6.3 <i>Treated-counterfactual turnout difference, with 95% confidence interval</i>	56
Table 6.2 <i>Estimated treatment effect by period</i>	57
Figure 6.4 <i>Treated and counterfactual turnout by state</i>	58
Table 6.3 <i>Estimated treatment effect (generalized synthetic control)</i>	59
Figure 6.5 <i>Treated-counterfactual turnout difference (generalized synthetic control)</i>	60
Figure 6.6 <i>Treatment status by state and election (restricted controls)</i>	61
Figure 6.7 <i>Treated-counterfactual turnout difference (restricted controls)</i>	61
Table 6.4 <i>Estimated treatment effect (restricted controls)</i>	62
Figure 6.8 <i>Treated and counterfactual average turnout over time (all elections)</i>	63
Figure 6.9 <i>Treated-counterfactual turnout difference (all elections)</i>	64
Table 6.5 <i>Estimated treatment effect (all elections)</i>	64
Table 6.6 <i>Difference-in-differences of turnout between treated and control, by group</i>	65
Figure 6.10 <i>Treated-control difference in turnout, 1980-2004</i>	66
Table 6.7 <i>Difference-in-differences (alternative specifications)</i>	67
Figure 6.11 <i>Group-level effect of strict photo ID-laws, with 95% confidence intervals</i>	69
Figure A.1 <i>Latent factor</i>	94
Figure A.2 <i>Factor loading</i>	95
Figure A.3 <i>Treated-counterfactual turnout difference (all elections – GSC)</i>	96
Table A.1 <i>Estimated treatment effect (all elections – GSC)</i>	96
Figure A.4 <i>Treated-counterfactual turnout difference (all elections – restricted controls)</i> ..	97
Table A.2 <i>Estimated treatment effect (all elections – restricted controls)</i>	97

1 Introduction

Voter ID-laws – requirements that voters provide documentary proof of their identity at the polling station – have proliferated in the United States in recent decades, with a majority of states now requiring some form of ID in order to vote. These laws have proven highly controversial, labelled by critics as “a part of an ongoing strategy to roll back decades of progress on voting rights” (American Civil Liberties Union 2017). Unsurprisingly, opposition has often been channeled through the courts. Despite continued efforts, such challenges have been largely unsuccessful. In upholding Indiana’s voter ID-law in a landmark 2008 decision in *Crawford v. Marion County Election Board*, the Supreme Court majority concluded that it had not been sufficiently proven that such laws would keep people from voting, and that they were therefore constitutional in the absence of evidence to the contrary.¹ Since then, numerous studies have sought to provide evidence one way or the other. The present study joins this growing field. Concretely, the aim of the thesis is to investigate how ID-requirements affect voting in the aggregate, i.e. turnout rates. I therefore consider the following research question:

Have voter ID-laws led to lower turnout?

Specifically, I study the most restrictive form of voter ID-laws: strict photographic requirements, which require voters to show government-issued, photographic ID in order to cast a vote on Election Day. Representing the strongest policy intervention, these laws should have the highest likelihood of affecting turnout. Using a counterfactual approach, I find that turnout in the states that have adopted strict photographic ID-requirements has not been significantly different from what it would have been if ID-requirements were never implemented. The overall average treatment effect on the treated is estimated to lie between -2.86 and 5.03 percentage points change in turnout. Additionally, I find no significant evidence that voters belonging to an ethnic or racial minority were affected differently than non-minority voters. In sum, this analysis suggests that voter ID-laws do not lead to lower turnout.

¹ The full decision is available at <https://www.supremecourt.gov/opinions/07pdf/07-21.pdf>.

1.1 Why study voter ID-laws?

Studying the effects of voter ID is important on three levels: normatively, electorally, and academically. First, the debate around voter ID-laws involves issues of strong normative importance. Proponents argue in favor of such requirements as a necessary protection of the election process against *voter fraud* (Fund 2008). Meanwhile, opponents argue against them out of concern that they will lead to *voter suppression* as otherwise eligible voters are turned away at the polls for lacking ID. Because both voter fraud and voter suppression threaten fundamental democratic principles and, as such, are undesirable, examining the degree to which efforts to eliminate the former serve to exacerbate the latter is a worthwhile pursuit. Moreover, as will be discussed further in the following chapter, voter ID-requirements relate particularly to the history of disenfranchisement and political inequality along racial lines that continues to shape American politics. (Keyssar 2000; Fraga 2018).

Voter ID also has potential implications for electoral politics. As will be discussed more thoroughly in the next chapter, there is reason to assume that opposing political elites are actually in agreement in expecting ID-requirements to negatively impact turnout, disagreeing only on the desirability of that outcome. Concretely, Republicans, who stand to gain from *lower* turnout, support the laws, while Democrats, who stand to gain from *higher* turnout, oppose them (Hansford & Gomez 2010; Highton 2017). Plausibly, the struggle over voter ID partly represents a struggle over electoral outcomes. Regardless of whether ID-requirements are intended to alter turnout rates, the possibility that they could, and the possibility that such a change could, in turn, affect the results of elections makes the issue worth investigating.

Academically, this is just the case, as the study of voter ID and turnout has grown to comprise a substantial body of research. However, scholars have reached conflicting conclusions. Some studies are in accordance with the prevailing view expecting a negative effect (Alvarez et al. 2011; Dropp 2013; GAO 2014; Hajnal et al. 2017; Pryor et al. 2019; Kuk et al. 2020; Grimmer & Yoder 2021), while others find varying, limited, or no evidence of altered turnout (Erikson & Minnite 2009; Fraga 2018; Grimmer et al. 2018; Heller et al. 2019; Cantoni & Pons 2021). Given the high salience of the subject, its contentious nature, and the normative and electoral

importance of the issues involved, the relative inconclusiveness of the existing research constitutes an obvious need for additional work. In the next section, I make the case for why my thesis, through a stringent methodological design, could provide a more definite answer as to the causal effect of voter ID-requirements on turnout.

1.2 A novel contribution

The main innovation of this paper is methodological. The fundamental problem when attempting to infer the effects of voter ID-laws concerns estimating the counterfactual: what would turnout levels be in voter ID-states had they *not* adopted such laws? This is, of course, unobservable. The problem is compounded by the fact that treatment assignment is nonrandom: states that adopt voter ID-laws are likely to differ systematically from those that do not, regarding both pre-intervention turnout levels and various other factors, in ways that may confound inferences if not controlled for. Simply observing that turnout is higher or lower after the passage of a voter ID-law is similarly insufficient to conclude that the new requirement *caused* this change, as it may have occurred regardless.

This study falls within the tradition of causal empiricism, which focuses on accurate causal inference through “careful use of an identification strategy research design and interpretation of the specificity of the results” (Samii 2016, 949). My research design pays explicit attention to the counterfactual scenario and – by utilizing a synthetic control approach – applies a novel and sophisticated method towards its estimation. Conceptualizing the cause of interest (in this case, voter ID-requirements) as an intervention analogous to a medical treatment, a time-series of the dependent variable (turnout) is generated for each unit (state) under treatment to simulate the counterfactual scenario in which the unit did not receive treatment, based on data from untreated units. This is then compared to the actual (treated) time-series. Any deviations are, conditional on some identifying assumptions, evidence of a causal effect of the treatment.

A key issue is to ensure similarity between treated and (counterfactual) control. The synthetic control method can capture and control for both constant and time-varying differences between states, even ones that are unknown and unobserved. Many previous studies have relied on tools

that can only control for static or observed variation between states; synthetic control has no such limitation. It is therefore particularly well suited to addressing the issue of non-comparability across treated and untreated states.

An additional advantage relative to much of the previous literature concerns the amount of data available for analysis. The passage of time means more data is now available, as both additional states adopting voter ID-laws and new elections being held have yielded new observations since most previous studies. This allows me to focus on the strictest form of ID-requirements, which are also the most recent and for which data has therefore until recently been limited. Because these are the types of laws most likely to have an effect on turnout, and because my analysis incorporates all states and presidential elections in which strict photographic ID-requirements have been in effect, including the election of 2020, this study is positioned to summarize what overall effect voter ID-laws have had on turnout, if any.

1.3 Structure of the thesis

The thesis is structured as follows. Chapter 2 introduces the concept of voter ID along with contextual information on the history of voting and turnout in the United States. I also present a way to classify different types of voter ID-laws and examine what has made them so controversial.

In chapter 3, I present a theoretical framework with which to formalize expectations of the effect voter ID-requirements have on turnout. Drawing mainly on rational choice theory, I also highlight alternative explanations of why people vote and the degree to which they lead to alternative expectations of the causal relationship under examination. From this discussion, I specify two hypotheses for testing. I then review relevant literature, both descriptive research regarding the empirical plausibility of the hypotheses and previous studies on the causal effect of voter ID-laws on turnout.

In chapter 4, I detail the methods used to investigate the research question. Following an elaboration of the counterfactual conceptualization of causality, I introduce the synthetic control method as a powerful tool for exploring causal effects within this framework. I also present difference-in-differences designs as a useful substitute for when synthetic control is unfeasible.

In Chapter 5, I describe the data used and discuss issues related to measurement and data availability. I also detail further the data limitations necessitating the switch in methodology from synthetic control to difference-in-differences analysis.

In chapter 6, I present the results of the analysis. Several alternative model specifications are also considered.

In chapter 7, I evaluate the hypotheses and answer the research question. I then discuss how to explain the results, and consider their implications for the issue of voter ID and theories of voting. I also compare my findings to those of previous studies and relate them to policy. Finally, I highlight some drawbacks to the analysis, before concluding the thesis by offering some suggestions for future research.

2 Voter ID: context, concept and controversy

In this chapter, I frame the analysis by providing some key background information. Beginning broadly, I give an overview of the history of voting and turnout in the United States. I then introduce the phenomenon of study: voter ID-laws. After reviewing the brief history of ID-requirements in American elections, I provide a typology of voter ID-laws wherein states are classified by the nature and strictness of their ID-requirement. Finally, I introduce the controversy that surrounds these laws and in turn provides the impetus for this paper, discussing both the argument in favor of ID-requirements and the suspicions with which they have met. The two sides of the debate largely focus their attention on two different threats to democracy: voter fraud and voter suppression.

2.1 Voting in the United States

This section serves to contextualize the dependent variable (turnout) and the independent variable (voter ID-requirements). I begin with the latter, giving a historical account of the right to vote and related regulations and restrictions of the franchise in the United States, before summarizing how turnout rates have developed over time.

In his tellingly titled book, *The Right To Vote : The Contested History of Democracy in the United States*, Keyssar (2000) challenges what he terms “the progressive presumption” of the American franchise as a linear progression towards ever fewer restrictions on the right to vote. Rather, he charts the history of suffrage in the United States as a turbulent oscillation between efforts of expansion and contraction.

Following the colonial period, wherein voting was mainly a privilege of the property-holding class, the first half of the 19th century saw an increasingly democratic national sentiment and a

corresponding gradual expansion of the franchise among adult white men. In 1870, the 15th Amendment formally granted African-American men the right to vote. Momentum then shifted in the time between the Civil War and the First World War, “when faith in democracy was challenged by doubts about the ability of ordinary people to exercise the vote intelligently, and class, ethnic, and racist prejudices gave rise to new restrictions on the franchise” (Briffault 2002, 1513). One notable exception was the extension of the suffrage to women in 1920 with the adoption of the 19th Amendment.

After a period of stability in the interwar years, the tide turned once more in the post-war period, this time towards inclusiveness, in a process that would culminate in the near-universal suffrage of today. Several landmark moments punctuate this era of enfranchisement (Briffault 2002, 1520-1522). The 1965 Voting Rights Act curtailed states’ ability to enact discriminatory electoral policy, while the 26th Amendment lowered the voting age to 18 in 1971. Additionally, several Supreme Court decisions struck down restrictions like poll taxes and requirements of property, tax payment, and residency of more than 50 days. More recently, there is the National Voter Registration Act of 1993, which streamlined the process of voter registration, including the possibility of registering at a Department of Motor Vehicles and simultaneously registering to vote when applying for a driver’s license – hence its nickname of the Motor Voter Bill (Keyssar 2000, 314). Finally, following the controversial 2000 presidential election, the Help America Vote Act of 2002 created new minimum standards for election administration, including requiring first-time voters who register by mail to present ID when voting (Congress.gov 2002)

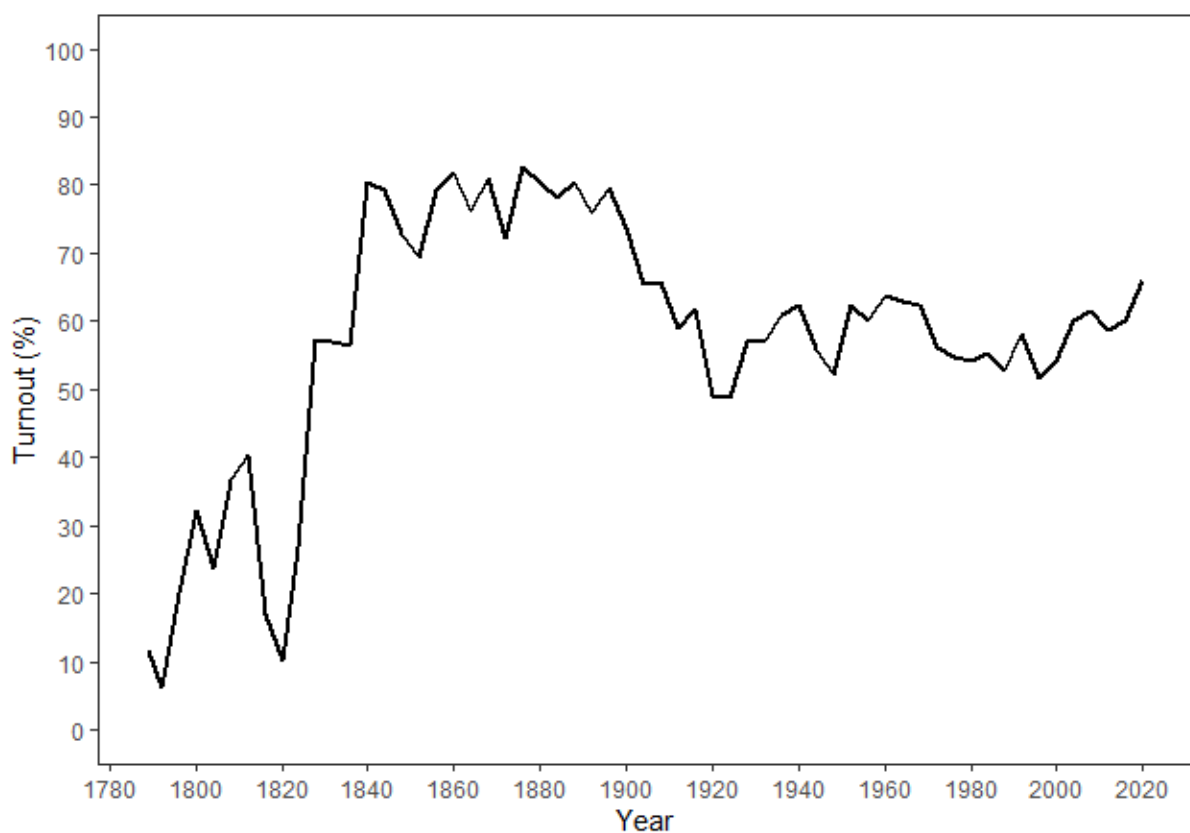
The enfranchisement of the American population, then, has been a contested process, with numerous examples of backsliding at various times and in various places. The instruments used to restrict the right to vote have been varied, too: From outright bans on voting among women, slaves, and the property-less, via more indirect tactics like poll taxes, literacy tests, and white-only primaries, to outright violence and intimidation (Briffault 2002, 1520). Concluding his account with “a partially happy ending”, Keyssar (2000, 316) gives a status report at the turn of the millennium:

What once was a long list of restrictions on the franchise has been whittled down to a small set of constraints. Economic, gender-based, and racial qualifications have been

abolished; literacy tests are gone, If not forgotten; residency requirements have been reduced to a matter of weeks; the age of political maturity has been lowered; and the burden of registration has been rendered less onerous.

Despite this largely positive picture, Briffault (2002, 1523-1527) points out that a few exceptions to the universal right of suffrage do remain, most notably concerning convicted felons, non-residents, and non-citizens.

Figure 2.1 National turnout rates in presidential elections, 1789-2020



Source: "U.S. VEP Turnout 1789-Present", collected from McDonald (2021a).

Echoing the United States' history of disenfranchisement is the low levels of participation in contemporary American elections. However, this has not always been the case, as is evident from figure 2.1, which displays the development of national turnout rates for US presidential elections.² In a manner similar to Keyssar's account of the development of the American

² Of course, the turnout rates in any given year are calculated from the portion of the population actually eligible to vote; thus, the high turnout rates for much of the 19th century are somewhat misleading given the exclusion of a majority of the populace from the electorate.

electorate, McDonald (2010) divides the history of turnout in the United States into several distinct eras.

In the Founding Era (1789-1824), elections were both low in salience and difficult to participate in due to the poor infrastructure and low population density, leading to the lowest turnout rates in American history. This pattern reversed in the Party Machine Era (1828-1896), when strong party organizations designed to deliver votes drove turnout to record highs. Participation declined in the Segregation Era (1900-1948), particularly in the Southern states, where African-Americans were *de facto* disenfranchised. Generally, the dismantling of the party machines reduced mobilization efforts, and “voting rates dropped as voting costs previously born by the political parties’ organizations were shifted to individuals” (McDonald 2010, 135). Finally, the Nationalization Era (1952-present) has seen electoral barriers disappear as the national government committed to overseeing election administration. As a result, Southern turnout has climbed back towards that of the rest of the country; the national average for presidential elections, meanwhile, has remained relatively stable in the 50- and 60-percent range (McDonald 2010, 128).

While the deterioration in turnout in later years observed by some scholars is largely an artefact of inaccurate measurement due to an increase in the number of residents not eligible to vote – accounted for in figure 2.1 – there is no denying that current US turnout rates are, relatively speaking, conspicuously low (McDonald 2010, 139; McDonald 2021e; Leighley & Nagler 2013).³ Comparing the election of 2016 to other OECD nations, the United States ranks 30th out of 35 countries for which data on turnout among the voting age population is available (Desilver 2020). Finding that differences in individual-level characteristics are unable to explain this pattern, Martinez (2010) suggests as an explanation the comparatively low salience of American elections, owing to the uncompetitive nature of many districts, opaque accountability resulting from the separation of powers, and the weak party-group linkage of the two-party system.

³ See chapter 5 for a discussion on how best to measure turnout.

Lastly, there is considerable variation in participation rates at the group-level, with highly educated and wealthy individuals being considerably more likely to vote (Leighley & Nagler 2013). Additionally, Fraga (2018) identifies a growing “turnout gap” between white voters and racial and ethnic minorities, robust to controlling for prominent sociodemographic factors – a finding which is of particular relevance to the study of voter ID-laws.

2.2 Voter ID laws

Because the administration of elections in the US is largely decentralized, there is a great deal of heterogeneity among states’ voting rules (McDonald 2010, 128). One notable exception came by way of the previously mentioned Voting Rights Act of 1965, which included a provision known as “preclearance”, whereby states with a history of discrimination at the ballot were required to seek approval from the federal government before making changes to their electoral rules (Fraga 2018, 32). In 2013, the Supreme Court decision in *Shelby County v. Holder* abolished that requirement, effectively restoring to the covered states their autonomy over election administration (United States Department of Justice 2020).

Enter voter ID-laws. Historically, requiring citizens to show an identity document in order to vote is not the norm in the United States, but in the latter half of the 20th century, some states began passing laws requiring voters to present some form of ID at the polls (NCSL 2017). Since their introduction, an increasing number of states have adopted voter ID-requirements – in later years, the list has come to include several states previously restricted by the preclearance requirement. A milestone occurred in 2005, when Indiana and Georgia became the first states to pass a new, stricter form of requirement whereby voters must identify themselves using photographic ID. Though challenged in court, the Supreme Court upheld the constitutionality of Indiana’s measure in the 2008 decision of *Crawford v. Marion County Election Board*, which paved the way for additional states adopting similar voter ID-laws (Highton 2017, 151). As of 2020, 16 states, as well as the District of Columbia, still require no document in order to vote – instead, voters are asked to state their name and occasionally address or date of birth, or provide a signature (NCSL 2021).

The National Conference of State Legislatures [NCSL] classifies states with voter ID-laws based on two dimensions: first, whether they require photographic or non-photographic ID; second, whether procedures are in place for voters without proper ID to still cast a vote on Election Day (non-strict) or whether such voters may only cast a provisional ballot, upon which they must later return to the polling place or an election office and show valid ID for their vote to be counted (strict) (NCSL 2020). These categories are not uniform – for instance, while all states with photographic ID-requirements accept driver’s licenses and passport, some may also accept student IDs and firearms licenses, among other types (Highton 2017, 150). Still, within-group variation is largely overshadowed by between-group differences.

Table 2.1 *Voter ID-laws in effect in 2020*

	Photographic ID	Non-photographic ID
Strict	Alabama Georgia Indiana Kansas Mississippi Tennessee Wisconsin	Arizona North Dakota Ohio
Non-strict	Arkansas Florida Hawaii Idaho Louisiana Michigan North Carolina Rhode Island South Carolina South Dakota Texas	Alaska Colorado Connecticut Delaware Iowa Kentucky Missouri Montana New Hampshire Oklahoma Utah Virginia Washington West Virginia
<i>Source: NCSL (2020).</i>		

Table 2.1 categorizes states with ID-laws in effect in 2020 according to the NCSL-typology. The NCSL classifies Alabama as a non-strict photo ID state but notes that “some might call Alabama’s law a strict photo identification law” because the only way voters can avoid returning to an election office to provide the required identification is to have “two election

officials [...] sign sworn statements saying they know the voter” (NCSL 2020). I agree with Highton (2017, 152) in placing Alabama’s law in the category of strict photographic ID-requirements, as this exemption is so narrow as to be virtually non-existent for most voters.

My thesis (and much of the existing research) focuses on states with strict photographic ID-requirements, as these represent the strongest policy intervention and thus are the most likely to affect turnout (Highton 2017, 151). Because of this, these laws are also the most controversial – discussed further in the next section. In total, 14 states have passed bills adopting strict photographic ID-requirements, nine have implemented them for at least one election, and seven states had such requirements in effect for the 2020 election (Highton 2017, 153; NCSL 2017; NCSL 2020). Table 2.2 provides a visual overview of the distribution of the treatment of interest across states and over time among the states that have adopted strict photo ID-requirements. Note that the year adopted refers to the year a law was passed, not the year it was implemented. Because some laws were blocked in court prior to implementation, not all of the states have actually undergone treatment.

Table 2.2 *States with strict photo ID-requirements, by election*

State	Law adopted	In effect for election							
		2006	2008	2010	2012	2014	2016	2018	2020
Indiana	2005		✓	✓	✓	✓	✓	✓	✓
Georgia	2005		✓	✓	✓	✓	✓	✓	✓
Missouri	2006								
Alabama	2011					✓	✓	✓	✓
Kansas	2011				✓	✓	✓	✓	✓
Tennessee*	2011				✓	✓	✓	✓	✓
Texas	2011					✓			
Wisconsin	2011						✓	✓	✓
Mississippi	2012					✓	✓	✓	✓
Pennsylvania	2012								
Arkansas	2013								
North Carolina	2013								
Virginia	2013					✓	✓		
North Dakota	2015								

* Tennessee made its law stricter in 2013 by reducing the types of photo IDs registrants are allowed to use to verify their identities.

Source: NCSL (2017; 2020), Brennan Center for Justice (2016), various news articles.

2.3 The debate

To the international observer, the controversy surrounding voter ID-requirements in the United States may seem puzzling – for instance, requiring photographic ID to cast a ballot is established practice in Norway and numerous other democracies (Norwegian Directorate of Elections 2021; Shaffer & Wang 2009). However, examining Indiana’s pioneering strict photo ID-law in an international perspective, Schaffer & Wang (2009) find it to be an outlier in terms of its strictness, the types of documents accepted, the ease of acquiring these, the consequences of lacking ID, and the exceptions granted. More generally, another key difference lies in the American context and the highly partisan nature of the debate over voter ID.

Proponents of voter ID argue that they are necessary in order to prevent *voter fraud* (Fund 2008). Voter fraud – “the intentional corruption of the electoral process by voters” – is distinguishable from the wider concept of *election fraud* in that it is perpetrated by voters, and not election workers, parties, or other organizations, while the requirement of intent separates it from accidental errors (Minnite 2010, 36). Voter fraud can take several forms; the most relevant to the topic of voter ID is impersonation fraud, whereby a fraudster casts a vote in someone else’s name (Hasen 2012, 61). Highlighting the need to ensure public confidence in American elections in the face of such threats, supporters of voter ID-laws maintain that “requiring individuals to authenticate their identity at the polls is a fundamental and necessary component of ensuring the integrity of the election process” (Von Spakovsky 2011, 1).

Opponents of voter ID point out that voter fraud in the United States – particularly impersonation fraud – is so rare as to be virtually nonexistent, and that the true threat comes from the new laws themselves, in the form of *voter suppression* (Hasen 2012). As the argument goes, requiring ID at the polling station makes it more difficult to vote for eligible voters lacking valid ID, which could in turn deter them from voting. In this view, voter ID-laws are at best a flawed solution to an imaginary problem, and at worst a deliberate attempt at voter suppression – yet another entry in the history of disenfranchisement outlined by Keyssar (2000).

Regardless of whether one accepts the charges of voter suppression, the issue has a historical parallel in the introduction of the secret ballot and voter registration in the late 19th century. Like voter ID-requirements, these measures intuitively offer an intrinsic benefit in terms of safeguarding against voter fraud – indeed, to the modern observer secret ballots are commonsensical to the point where their *absence* is likely to be considered a threat to electoral integrity. However, in the era of Jim Crow, “the secret ballot also operated as a de facto literacy test for illiterate voters who now had to vote on their own” (Briffault 2002, 1518), while requiring voters to register in advance “kept large numbers (probably millions) of eligible voters from the polls” (Keyssar 2000, 158). They thus offer an early example of the tension between efforts to combat voter fraud and the ideal of electoral inclusiveness. Note that Keyssar (158-162) cautions that we cannot know the motives of the legislators passing these measures, and suggests that electoral self-interest did plausibly combine with genuine desires to safeguard the democratic process.

2.3.1 The partisan angle

The final key to the debate over voter ID is its highly partisan nature, which is closely tied to the hypothesized effect such requirements will have on turnout. Generally, higher turnout is thought to advantage Democrats while lower turnout favors Republicans, as non-voters are slightly left-leaning (Martinez & Gill 2005; Hansford & Gomez 2010; Leighley & Nagler 2013, 159). In the case of voter ID-requirements, a central part of the backlash against them is the claim that they will disproportionately burden certain particularly vulnerable groups: poor voters, minorities, the elderly, women, and the disabled (Sobel & Smith 2009, 107). Crucially, most of these groups tend to vote Democratic (CNN 2021). For this reason, opposition to and support of voter ID-laws generally forms around party lines (Gronke et al. 2019).

Several studies examining the determinants of the adoption of ID-requirements have found that voter ID-laws are most likely to be adopted when Republicans have a legislative majority but elections are closely competitive – in other words, when Republican lawmakers have both the means and the motive to do so (Hicks et al. 2015; Biggers & Hanmer 2017). Bentele & O’Brien (2013) find that a similar pattern extends to the passage of restrictive voting policies more generally. Furthermore, this partisan cleavage is not just an elite phenomenon: when examining popular support at the individual-level, Stewart et al. (2016, 1455) find that “partisan identity

[is a] powerful variable in predicting both support for voter identification laws and beliefs in the prevalence of voter fraud.”

Perhaps the strongest indicator of the partisan nature of voter ID-laws lies in table 2.3. Adapted from Highton (2017, 153), it shows the states that have adopted strict photographic ID-requirements and which party controlled the respective branches of state government at the time the law was passed. In all but two cases, all three branches of government were under Republican control. In Arkansas, the Democratic governor dutifully vetoed the law but was overruled by the legislature, while the Virginia state senate split down the middle as all Republican lawmakers voted in favor of the bill and all Democrats voted against – the stalemate was broken when the Republican lieutenant governor voted in favor.

Table 2.3 *Adoption of strict photo ID-laws and party control of state government*

State	Law adopted	Party control of lower house/upper house/governorship
Indiana	2005	Rep/Rep/Rep
Georgia	2005	Rep/Rep/Rep
Missouri	2006	Rep/Rep/Rep
Alabama	2011	Rep/Rep/Rep
Kansas	2011	Rep/Rep/Rep
Tennessee	2011	Rep/Rep/Rep
Texas	2011	Rep/Rep/Rep
Wisconsin	2011	Rep/Rep/Rep
Mississippi	2012	Rep/Rep/Rep
Pennsylvania	2012	Rep/Rep/Rep
Arkansas	2013	Rep/Rep/Dem
North Carolina	2013	Rep/Rep/Rep
Virginia	2013	Rep/Even/Rep
North Dakota	2015	Rep/Rep/Rep
<i>Source: Highton (2017).</i>		

Taken as a whole, it seems plausible that “partisan debates about voter identification laws reflect party competition over election outcomes” (Highton 2017, 154). In this thesis, I make no judgment as to whether a lower turnout resulting from voter ID-requirements is an acceptable cost to ensure against voter fraud, or indeed if conceptualizing the issue as a dichotomous trade-off is valid in the first place. Similarly, I do not attempt to prove or disprove that voter ID-laws are adopted strategically to manipulate turnout. Rather, I simply seek to determine whether they do.

3 Theory and previous research

In this chapter, I discuss the reasons why requiring voters to show photographic ID may be expected to negatively impact turnout rates, as well as the countervailing forces that could potentially offset this dampening effect. The discussion culminates in the specification of two hypotheses for testing. I also review past research on the effects of voter ID-laws.

3.1 Voter ID and turnout – theoretical expectations

My research question is motivated by the high degree of public debate surrounding the issue of voter ID and its potential real-world implications for democratic fairness. Claims of how and why voter ID-laws will lower turnout are plentiful in the public discourse; below, I show that these claims implicitly rest on a logic of individual rationality and cost/benefit-analysis. Explicating this underlying reasoning and anchoring the analysis in theory is useful in order to systematize these claims, evaluate their merit, and render them testable. My objective in this section is not to provide a comprehensive theoretical framework of the determinants of turnout. Rather, I introduce a simple rational choice model of voting as a tool with which to more rigorously explore the intuition behind the allegation that voter ID-laws will lead to lower turnout. I also discuss alternative theoretical explanations of why people vote, along with scholarly work offering a different perspective on the effect of voter ID-laws on turnout.

3.1.1 The calculus of voting

For the purposes of studying the effect of voter ID-laws, it is useful to think of voting in terms of costs and benefits. I therefore adopt a rational choice approach, using the so-called *calculus of voting* first introduced by Downs (1957). According to Downs, an individual's decision to vote or abstain can be summarized in the following simple equation,

$$(1) \quad R = PB - C,$$

wherein R represents the individual's utility-gain from voting, which is determined by: B , the benefit of having his preferred candidate or party win (relative to the alternative); P , the

probability that his vote will matter (in terms of producing the desired outcome); and C , the costs associated with the act of voting (Riker & Ordeshook 1968, 25). Whether or not the individual votes depends on whether or not R is positive or negative. Put simply: “[E]very rational man decides whether to vote just as he makes all other decisions: if the returns outweigh the costs, he votes; if not, he abstains” (Downs 1957, 260).

A significant criticism levelled against the calculus of voting is that equation (1) seems to predict that virtually all actors should abstain when voting is not entirely costless. In a national election with a large electorate, the probability of any single vote being decisive is exceedingly small. For the term PB to outweigh even a small C , then, the differential benefit of an individual having his preferred candidate win would have to be enormously – unrealistically – large (Riker & Ordeshook 1968, 26). The above model thus leads to the obviously false prediction that no rational actors will vote.

Several changes to the model have been suggested to explain why voting nevertheless occurs (Blais 2000, 3; Evans 2004, 83-87). Famously, Riker & Ordeshook (1968) amend the model by introducing a new term, D :

$$(2) \quad R = PB - C + D.$$

While B represents the *instrumental* value of voting and is thus conditional on achieving the desired outcome (represented by P), D represents the *intrinsic* value associated with the act of voting itself – regardless of the outcome. Elements of D include one’s sense of civic duty, the desire to support the political system, expression of one’s partisan preferences, and the satisfaction derived from informing oneself and participating in politics (Riker & Ordeshook 1968, 28). Though the contents of this term are not exhaustively defined, embodied within it are “any additional benefits that an individual receives from the act of voting” (Rolje 2012, 8). It therefore serves to explain why rational actors may still decide to bear the costs of voting, despite the low probability of individual instrumental efficacy.

The key prediction of the calculus of voting and the rational choice approach to turnout is that the decision to vote or abstain is likely a marginal one. Aldrich (1993) conceptualizes voting as both a low-cost and a low-benefit action, and argues that as such, even slight changes in the

costs of voting may alter turnout. He emphasizes how this creates an opportunity for elites to influence turnout through their strategic actions – recall from the previous chapter that voter ID-laws are likely once such example (Aldrich 1993, 274; Highton 2017, 156). Given this, it appears plausible that, *ceteris paribus*, increasing the cost of voting is likely to lead to lower turnout rates, if only slightly (Aldrich 1993, 250).

3.1.2 Voter ID-requirements and the costs of voting

What, then, are the costs of voting? Since the abolishment of poll taxes, voting is rarely costly in the monetary sense – at least not directly. However, this is not to say it is costless. Downs (1957, 265) argues that the main price voters pay is one of time: “[T]ime is the principal cost of voting: time to register, to discover what parties are running, to deliberate, to go to the polls, and to mark the ballot. Since time is a scarce resource, voting is inherently costly.” Blais et al. (2019) further distinguish between the direct costs of the act of voting itself, and indirect costs associated with informing oneself and making a decision.

The direct costs of voting are both inescapable⁴ and fairly straightforward – time spent going to, from, and waiting at, the polling place – though, as Blais (2000, 84) points out, ultimately “voting is not a very demanding activity.” Information- and decision-costs, meanwhile, are more nebulous. According to Downs (1957, 210) voters must gather information, analyze it, and evaluate how it relates to their preferences in order to make a decision. Because this all takes time, voters employ a cost/benefit-analysis and strive only to become minimally informed (Downs 1957, 207, 219). Still, even these minimal costs are avoidable if the individual decides at the outset of the campaign that he will abstain (Blais 2000, 84).

Additionally, voters may also be required to bear costs before the election in order to be eligible to vote. In the American context, the prime example of this is the fact that prospective voters must register as such. In this respect, voter ID represents a similar cost to voter registration. Essentially, there are two channels through which ID-requirements may cause fewer people to vote: first, there is a mechanical effect, whereby voters without ID are turned away at the polls;

⁴ With the exception of absentee and mail-in ballots, which historically have constituted a small but increasing share of votes (Stewart 2021). In the election of 2020, nearly half of all voters voted by mail, though circumstances were exceptional owing to the COVID-19 pandemic.

second, there is a deterrent effect, whereby voters without ID fail to turn out at all (Grimmer & Yoder 2021).

For voters lacking valid identification under the new law, voter ID-laws add an additional cost to voting in that they must now obtain ID prior to Election Day. Again, this is mainly a time-cost: time spent gathering documentation, travel time to the nearest issuing office, and time spent waiting in line (Shapiro & Moran 2019). Compounding this is the fact that many voters lack access to transportation⁵ and live far from an ID-issuing office, which may in turn have limited opening hours (Gaskins & Iyer 2012). Though small, this cost may cause some to choose abstention where they would otherwise have voted. Recall that the decision to turn out is likely marginal: if the equation is already perilously balanced, the time- and resource-expenditure required to acquire ID may tip the scales in favor of abstaining.

For some voters, the barrier might be higher still: while some might simply find it inconvenient to obtain valid identification, for others it may be practically impossible due to prohibitively long travel time, unavailability of transportation, an inability to take time off from work during opening hours, or other factors. In this scenario, voters are not simply discouraged from voting, but effectively barred from it.

Finally, much of the backlash against voter ID-laws centers on the assumption that the costs of voter ID-laws are not evenly distributed, but rather that some groups will be affected more strongly than others. Mainly, this concerns groups that tend to vote Democratic – see the discussion on the partisan nature of voter ID in the previous chapter. The prediction of a differential impact is mainly a function of supposed differences in rates of ID-holding: If members of certain groups are less likely to have ID then these groups are more likely to have turnout rates be negatively affected by ID-requirements, as more of their members perceive their costs of voting to have increased. Furthermore, socio-economic differences between groups mean the costs of acquiring ID could be comparatively harder to bear for some. The effect could compound if such cleavages overlap with differences in ID-holding. This paper focuses on racial and ethnic minorities, as they are the group that has received the most attention

⁵ Voters without a driver's license are, for obvious reasons, somewhat restricted in this regard.

both in the literature and in the public eye. See the section on previous studies below for more on the empirical plausibility of the above claims as they pertain to minority voters.

3.1.3 Alternative theories of voting and counteracting forces

The calculus of voting is not the only suggested explanation of why people vote. Turning the question on its head, Brady et al. (1995, 271) highlight three alternative reasons why people choose *not* to participate: “because they can’t, because they don’t want to, or because nobody asked.”

The first reason – “They can’t” – is commonly referred to as the *resource model*, and focuses on the resources available to the individual: time, money, and civic skills (Brady et al. 1995, 273). The more of these an individual has, the more able he is to bear the costs of voting and thus the more likely he is to participate (Blais 2000, 12). In contrast to the calculus of voting, however, this approach pays little attention to the benefits of voting – they are simply assumed. The resource model highlights how the impact of cost-increases are likely conditional on individual socio-economic status and thus lends credence to the hypothesized differential impact of ID-requirements.

Meanwhile, “They don’t want to” corresponds to the *psychological engagement* theory of voting, which focuses precisely on the motivation for voting: “Bluntly put, it asserts that the more interested a person is in politics, the more likely she is to participate in general and to vote” (Blais 2000, 13). However, this risks triviality, as one must then explain why people take an interest in politics in the first place (Blais 2007, 631).

While both the resource model and the psychological engagement theory have largely individualistic views on voting, “Nobody asked” is the answer associated with *mobilization theory*, which emphasizes contextual causes. Rosenstone & Hansen (1993) argue that people vote because they are induced to do so by external actors: informal social networks, political parties, and group networks like churches, voluntary associations and, trade unions (Schulz-Herzenberg 2019, 142). Mobilization occurs through social pressure that effectively raises the

cost of abstaining or, as Blais (2000, 13) points out, through efforts that reduce the cost of voting whereby “[actors] drive people to the poll on election day and provide cheap information about the issues.”

None of these traditions are necessarily incompatible with rational choice theory; however, other theories of voting more directly challenge this approach. The inclusion of a *D*-term in the calculus of voting has garnered criticism for including non-rational elements in the model and thus rendering it tautological and the rational choice approach useless in terms of predictive power (Blais 2000, 4-5; Rolfe 2012, 8). Meanwhile, what Blais (2000, 14) calls the *sociological interpretation* sees these elements of the *D*-term as the main drivers of voting. Here, voters are understood as collective actors, acting not based on their own interest but that of the community.

In an effort that draws on both the sociological tradition and mobilization theory, Rolfe (2012) develops a social theory of voting in which an individual’s decision to vote is not a function of an internal cost/benefit-analysis, but rather is driven by the behavior of those in his social network. This “conditional choice” approach “puts social cognition and social interaction – not individual preferences – at the center of individual decision-making.” (Rolfe 2012, 5). Thus, in deciding to vote or not, voters do not seek to maximize their own benefits so much as act in accordance with *group* goals (Rolfe 2012, 6). People turn out to vote when they perceive voting to be important to members of their reference group – the larger the social network the more likely that they will be mobilized (Rolfe 2012, 98-101). Put simply: voting is contagious.

For present purposes, the key takeaway from Rolfe’s theory and others in the sociological tradition is that they are theories in which individual benefits and – more importantly for the discussion of voter ID – costs are not the key factors determining whether people vote. In this view, we may expect less of an effect on turnout resulting from the increased cost of new ID-requirements than the calculus of voting suggests.

Relatedly, in addition to the idea that costs and benefits may not be integral to the voting decision, several other factors could potentially counteract the turnout-depressing effects of voter ID-requirements (Highton 2017, 157). Mobilization theory suggests one avenue, in that

actors may respond by attempting to neutralize the perceived negative effect through deliberate mobilization efforts. While the low-cost, low-benefit nature of voting provides strategic politicians with the opportunity to suppress turnout, it also enables them to augment it. Political elites who see their position threatened by lower turnout resulting from the new laws have every incentive to try to counteract them:

[G]iven the belief that strict voter identification laws advantage the Republican Party, the Democratic Party has a strong incentive to mobilize Democratic voters with proper identification and to help those who do not already have proper identification to obtain it.” (Highton 2017, 157)

There is also the possibility of voter ID-laws leading to increased psychological engagement, as people outraged by the perceived attempt at voter suppression rally at the polling station.⁶ Valentino & Neuner (2017) demonstrate how media frames emphasizing the controversy surrounding voter ID makes voters angry and increase their likelihood of participation. Moreover, this outrage-effect is stronger for Democratic voter groups, who are mobilized by exposure to both frames of voter fraud and voter suppression (Valentino & Neuner 2017, 347). This individual mobilizing effect may thus specifically offset the differential impact of voter ID-laws outlined previously.

Finally, Vercellotti & Andersen (2009) argue that the negative effects of ID-requirements are likely to be strongest immediately after implementation. Even voters who own or would acquire ID may initially fail to comply simply because they are unaware of the new requirements.⁷ Vercellotti & Andersen suggest a learning curve in which the frequency of such occurrences decline with time as more people become aware of the new law and acquire valid identification. However, this temporal weakening of the demobilizing effect may apply equally to the counteracting forces outlined above: as the issue of voter ID becomes less salient, both party mobilization efforts and voter outrage are likely to decrease. The expected long-term effect is

⁶ See Biggers & Smith (2020) for a study on the mobilizing effect of disenfranchisement in a different area of electoral policy.

⁷ Stewart et al. (2015, 1482) find that only 57% of respondents living in states with strict photo ID-laws at the time of their study were aware of the requirement.

thus theoretically unclear. A key strength of my research design is precisely that it explicates how the treatment effect develops over time and therefore allows investigation of this issue.

3.2 Hypotheses

The controversy surrounding voter ID-laws suggests that requiring voters to show photographic ID could cause fewer people to vote. For the portion of the electorate that lack valid ID, voter ID-requirements represent an increase in the costs associated with voting. For some, this increase may be sufficiently large to cause them to abstain. We can therefore expect ID-laws to negatively affect turnout:

H1: *Implementation of strict photographic voter ID-requirements lowers a state's turnout rate, relative to what it would otherwise be.*

However, while this prediction builds on a rational choice approach to voting, alternative perspectives on participation suggest that increasing the cost of voting may not severely influence turnout. Additionally, the contentious nature of the issue might spur mobilization – both top-down and at grassroots level – which could moderate, or even nullify, any demobilizing effects in the aggregate. There is thus reason to expect only a weak or no effect.

Finally, much of the backlash against voter ID-laws centers on an assumption that the dampening effect on turnout will be unevenly distributed. Minority voters are thought to be particularly vulnerable. I therefore consider a second hypothesis to investigate the possible differential impact of voter ID-laws:

H2: *Strict photographic voter ID-requirements lower minority turnout more than non-minority turnout.*

3.3 Previous research on voter ID-laws

In this section, I examine the empirical literature relating to the study of voter ID-laws. I begin by investigating some empirical patterns on which the plausibility of my hypotheses depends, drawing on descriptive studies of ID-holding and turnout. I then review studies directly estimating the effect of voter ID-laws on turnout and summarize their findings.

3.3.1 Empirical plausibility

How exactly do voter ID-laws increase the costs of voting? Highton (2017, 156) concisely lays out the micro-level prerequisites for a negative treatment effect:

If some people (a) who would otherwise vote (b) do not have one of the required forms of identification and (c) are not sufficiently interested and motivated or lack the resources to obtain the necessary identification in advance of the election, then turnout will be lower as a result of a voter identification law.

The macro-level effect on turnout thus depends on the proportion of the electorate who exhibit (a), (b), and (c). If everyone either has valid ID or is willing and able to bear to cost to obtain it, or no one without ID would vote anyway, voter ID-laws are unlikely to alter turnout significantly. The previous discussion leads us to assume (c) to hold for at least part of the electorate. The plausibility of a negative effect of voter ID-laws on turnout thus hinges on some empirical patterns, which warrant examination.

Logically, voter ID-requirements mainly represent an increase in costs to the portion of the electorate who lack valid ID; for those who already have the requisite documents, no new actions are required and thus the costs are unchanged. Stewart (2013) conducts a nationally representative study on rates of ID-holding among registered voters. Though many forms of ID exist and there is variation even among states with strict photographic ID-requirements regarding which types are accepted, driver's licenses and passports are – in addition to being the most commonly held forms of photographic ID – accepted in all these states (Stewart 2013, 38; NCSL 2020). Of the two, driver's licenses are by far the most common: 91% of respondents report having a driver's license while only 41% own a passport (Stewart 2013, 36). Nine percent

of voters thus lack any form of driver's license. However, some states additionally require licenses to be unexpired and match the address and exact name under which the voter is registered. Fully 20% of voters fail to meet these criteria (Stewart 2013, 40). It therefore appears plausible that for at least some voters, voter ID-laws of the kind considered here represent an additional obstacle to voting.

Of course, if voters lacking valid ID under the new law would not have voted anyway, turnout will remain unaffected. As far as could be ascertained, no comprehensive national study exists comparing turnout and ID-holding. However, in a study of Georgia, Hood & Bullock (2012) find that registered voters without a driver's license were considerably less likely to vote than those with licenses in the two elections preceding implementation of ID-requirements, with a gap in turnout rates in excess of 30 percentage points. Thus, while we can reasonably assume that the turnout rate among the ID-less portion of the electorate is not zero, the relatively low turnout among this group suggests a limited potential for voter ID-laws to impact overall turnout rates.

Regarding the hypothesized differential impact, Stewart (2013, 41) finds that 93% of white voters have a driver's license, compared to 90% of Hispanics and 79% of African-Americans. However, when controlling for the more stringent criteria mentioned above, fully 37% of blacks and 27% of Hispanics lack valid ID, compared to 16% of whites. Black and Hispanic voters, then, are less likely to hold valid ID, especially under the strictest requirements. Barreto et al. (2019) also find that various minority groups are less likely than whites to hold unexpired, government issued ID. Furthermore, they show that this pattern holds even when controlling for key socio-economic factors like income and education, suggesting that the gap is distinctly racial. Additionally, because minority voters are on average poorer than non-minorities, the cost of acquiring identification is arguably harder to bear for this group (Kochhar & Cilluffo 2018). These observations lend credence to the claim that minorities will be disproportionately affected. On the other hand, they are balanced by the fact that minority voters are on average more likely to abstain relative to white voters, meaning the proportion of potentially excluded voters is smaller (Fraga 2018).

3.3.2 Previous studies on the effect of voter ID-laws on turnout

In an article titled “Voter Identification Laws and Turnout in the United States” published in the *Annual Review of Political Science*, Highton (2017) conducts a metastudy of existing research investigating the effects of ID-requirements on turnout. He acknowledges the self-selection and resulting nonrandom assignment of treatment as a major obstacle to causal inference, and therefore gives precedence to studies that he deems to sufficiently control for cross-state differences, of which he finds four.

Using a difference-in-differences design, Erikson & Minnite (2009) find restrictive ID-laws to be associated with a slight drop in turnout, but the effect fails to reach conventional significance levels. Meanwhile, Alvarez et al. (2011) employ a Bayesian shrinkage estimator to show that turnout decreases slightly as ID-requirements become more restrictive, but the effect is again small – at most 2 points difference in turnout.

While neither of the first two studies analyze elections after 2006, Dropp (2013)⁸ leverages more recent data. He finds a noticeable drop in turnout of treated states in half of the four election-pairs under study. Lastly, the US Government Accountability Office [GAO] (2014) conduct a study comparing two states implementing strict photographic ID-requirements to four similar control states. The analysis shows an average decrease in turnout of 2.6 percentage points between 2008 and 2012. Regarding the group-level differential impact, GAO (2014) finds a stronger effect among African-Americans in both treated states, while Dropp (2013) finds similar evidence only for some elections.

Emphasizing that none of these studies found a turnout-effect larger than 4 percentage points – neither overall nor at the group-level – Highton (2017, 163) concludes that “the claim that voter identification laws depress turnout to a substantial degree is difficult to sustain based on existing

⁸ This study is cited by Highton as an unpublished manuscript, with a hyperlink to a now expired domain. As I have been unable to retrieve it elsewhere, I have not included it in my references. The full citation as it appears in Highton is as follows:

Dropp KA. 2013. *Voter identification laws and voter turnout*. Unpublished manuscript.
<http://kyledropp.weebly.com/current-research.html>

evidence.” However, the most recent article considered by Highton was published in 2015. Since then, some additional studies have been published.

Hajnal et al. (2017) find evidence of a clear negative effect. Focusing on the differential impact of strict ID-requirements (both photographic and non-photographic), they use validated survey data for congressional elections in all states between 2006 and 2014 to investigate the impact on turnout. The analysis shows that turnout rates of ethnic and racial minorities are disproportionately negatively affected, particularly in primary elections. Consistent with claims of the strategic nature of voter ID-laws, Hajnal et al. also find that left-leaning voters are more likely to experience turnout declines than those on the political right. Grimmer et al. (2018) argue that these findings are due to methodological errors in the study that, when corrected for, could yield positive, negative or null estimates of the effect. When replicating the study using data from the Current Population Survey, Pryor et al. (2019) find a negative overall effect on turnout, but also – curiously – that white voters are more strongly affected than non-whites.

Fraga (2018) integrates and employs a multi-year DID on so-called voter file data to assess the potential differential impact of photo ID-laws (both strict and non-strict) among minority voters. His results display a high degree of variability depending on election, state, exact ethnic group, and type of ID-requirement. While many permutations of the above indicate a significant negative effect, Fraga also finds equally strong evidence of a *positive* effect for others. Most estimates range between -4 and +4 percentage points change in turnout. He concludes that “we do not yet have consistent evidence of a differential, negative effect of voter identification laws on minority turnout as compared to non-Hispanic White voter turnout” (Fraga 2018, 184).

Meanwhile, using data on state-level turnout rates over 8 elections, Heller et al. (2019) find no significant evidence of a decrease in neither overall turnout rates nor those of black voters. They do find a small (2.6 - 5.4 percentage point) decrease among Hispanics, but this effect disappears when using fixed effects to control for cross-state heterogeneities.

Comparing the 2012 and 2016 presidential elections using a difference-in-differences design, Kuk et al. show that in the four states which implemented strict photo ID-requirements in this

period, turnout declined more in racially diverse counties (relative to predominantly white ones) than in other states. They find that turnout in counties with a 75% non-white population fell by 1.5 - 2.6 percentage points more in treated states than in the control group (Kuk et al. 2020, 5).

Grimmer & Yoder (2021) focus on the group theory predicts to be most strongly affected: those without ID. They exploit the experience of North Carolina in 2016, where a strict photo ID-law was in effect for the primary election but was struck down before the general election. They argue that the deterrent effect of ID-laws persists even after the law itself is repealed if voters are not sufficiently informed of the change. The results support this, but the estimated effect is not very large: among the 3% of registered voters lacking ID there was a 0.7 percentage point drop in turnout for the 2016 primaries (relative to those with ID), and a 2.6 point decrease in the general election. Additionally, because minorities and Democrats were less likely to hold ID, they were disproportionately affected. Note that the causal estimate here is the differential effect on those without ID relative to ID-holders, not the overall effect on turnout (Grimmer & Yoder 2021, 10).

Most recently, Cantoni & Pons (2021) apply a difference-in-differences design to investigate the effect of strict ID-requirements (both photographic and non-photographic). Utilizing voter file data compiled by a private data vendor, their dataset contains information on virtually all American voters over a 10-year period, for a total of 1.6 billion (!) observations. The analysis reveals “no negative effect on registration or turnout, overall or for any group defined by race, gender, age, or party affiliation” (Cantoni & Pons 2021, 1). The overall change in turnout resulting from strict ID-laws is estimated to lie between -3.0 and 2.8 percentage points. In exploring the cause of the null findings, they find no evidence of a counteracting force due to backlash, though there is some evidence of party mobilization.

In summary, the picture is mixed. Some studies find no significant impact of voter ID-requirements on turnout. Many, however, seem to indicate a negative effect, particularly among minorities. Taken together, the literature seems to suggest that to the degree to which an effect exists, it is likely to be small in magnitude. In any case, researchers seem unanimous on one important point: correctly identifying the effect of voter ID-laws is difficult, and no study is

likely to do it perfectly. In the next chapter, I elaborate on the challenges to causal inference and detail how they are addressed in this paper through a synthetic control approach based on a counterfactual conceptualization of causality.⁹

⁹ Highton (2017), in fact, identifies synthetic control as a fruitful avenue for future research.

4 Methodology

In this chapter, I will account for the strategies employed to estimate the effect of voter ID-laws on turnout: the synthetic control method and difference-in-differences analysis. I begin more broadly, however, by introducing the potential outcomes framework as a useful way of conceptualizing causality and discussing the feasibility of making causal inferences regarding social science phenomena.

4.1 Causal effects and the problem of counterfactual outcomes

Gerring (2012, 199) defines causality as follows: “[T]o say that a factor, X , is a cause of an outcome, Y , is to say that a change in X generates a change in Y relative to what Y would otherwise be.” If X is a binary treatment to which units of interest are either exposed or not, the causal effect of X under the above definition is the difference between Y when the treatment is in effect and Y when it is not. This can be formalized using the *potential outcomes* framework – also referred to as the Rubin causal model or the counterfactual model of causality (Rubin 1974; Morgan & Winship 2007). Conceptualizing the cause X as a binary treatment, we must consider two conceivable outcomes Y for any unit i . $Y_i(1)$ is the outcome when the unit receives treatment ($X = 1$) and $Y_i(0)$ is the outcome without treatment ($X = 0$). Estimating the causal effect of the treatment on any unit i is simply a matter of subtracting the “baseline” outcome from the treatment outcome: $Y_i(1) - Y_i(0)$.

The problem, as the name of the framework suggests, is that these outcomes are *potential*. Because they are mutually exclusive in the sense that no unit can simultaneously receive and not receive treatment, at any given time only one outcome is realized and observable. To put it – somewhat absurdly – in the language of the present application: no state simultaneously does and does not require voters to show photographic ID. If unit i receives treatment, we can observe its outcome under treatment $Y_i(1)$, but not its outcome absent treatment $Y_i(0)$. Vice versa, for a unit that did not receive treatment, we can observe $Y_i(0)$ but not $Y_i(1)$. This is what is known as the fundamental problem of causal inference (Imai 2017; Morgan & Winship 2007;

Gerring 2012). Because counterfactual outcomes are unobservable the above arithmetic becomes impossible to perform, and thus we *cannot know* the effect of a treatment on an individual unit of analysis. Writes Gerring (2012, 218):

One can never know with absolute certainty whether some factor caused an outcome to occur, because one cannot go back in time to re-play events exactly as they happened, changing only the factor of interest and observing the outcome under this altered condition. The causal counterfactual can never be directly observed for there are no time-machines.

4.1.1 Dealing with the counterfactual problem

Absent Gerring's hypothetical time machine, causal inference centers on how best to estimate the unknowable counterfactual outcome. The strength of a causal research design therefore revolves around the choice of identification strategy – or approach to identifying a substitute for the counterfactual. Effectively, this involves making an assumption “about what *observed* quantity is a good counterfactual for the treated units [emphasis added]” (Keele & Minozzi 2013, 194).

The randomized experiment is commonly referred to as the gold standard for causal inference (Imai 2017, 49). Two features allow experimental designs to effectively circumvent the counterfactual problem. First, they shift the focus from individual-level effects to estimating group-level *average effects*, by dividing units of observation into a minimum of two groups, where units in one group receive the treatment while those in the other group do not. Second, units are *randomly assigned* to the treatment and control groups. When the number of units is sufficiently large, this randomization means that we can reasonably assume that the two groups are – in the aggregate – identical in all aspects save for the treatment (Imai 2017, 50). Put differently, there are no *systematic* differences across groups, only random ones. We can therefore calculate the average treatment effect (ATE) as the difference between mean treatment and control group outcomes. Because of this ability to ensure comparability of treated and control units, the experiment functions as an ideal type for causal research; alternative methods are judged on the basis of how well they can replicate the features of an experiment. The closer the approximation of experimental conditions, the more valid are the causal inferences (Gerring 2012, 257).

Conducting an experiment, however, requires that the researcher be able to manipulate the treatment – to control its assignment to ensure that it is random. Given the phenomena of study, this is rarely possible in the social sciences – for instance, no researcher can dictate which states do and do not adopt voter ID-requirements. Rather, social scientists have to make due with observational data, where the treatment is out of their control and therefore likely nonrandomly assigned or even self-selected. (Morgan & Winship 2007, 41). This means treatment and control groups are likely to differ in ways other than the treatment, which in turn makes control group outcomes invalid as a counterfactual, because between-group differences in outcomes are likely to be confounded if these additional dissimilarities also affect the dependent variable (Imai 2017, 57-58). Simple comparison of group means is therefore inappropriate, as it will generate biased estimates of the causal effect.

For this reason, social science research often requires more sophisticated methods to solve the counterfactual problem. When the treatment of interest is in the form of a government policy change – as is the case with this paper – one should be particularly aware of endogeneity issues whereby treatment assignment is correlated with the outcome variable, as units essentially self-select into treatment and control groups. Besley & Case (2000, 693) write: “There is little doubt that policy choice is purposeful action and can rarely be treated as experimental data. The real issue is how to deal with this.” To this question I now turn.

4.2 The synthetic control method

A common strategy for inferring counterfactual outcomes is through the comparative method. In a most-similar-systems design, the researcher compares a case exposed to the cause of interest to one or more cases that are not, using the latter as a substitute for the counterfactual scenario (Landman 2008, 70). Crucially, this hinges on choosing control cases that are similar to the treated case. Pioneered by Abadie & Gardeazabal (2003), the synthetic control method is fundamentally a form of comparison in which a single, synthetic comparison unit is created as a weighted average of several control cases, under the assumption that “a combination of units often provides a better comparison for the unit exposed to the intervention than any single unit

alone” (Abadie et al. 2010, 494).¹⁰ If the trend in the dependent variable (e.g. turnout) of the synthetic control closely matches the treatment unit in the period prior to treatment, there is reason to believe that it would have matched in the post-treatment period as well if no intervention occurred (Abadie et al. 2015, 496-498). In other words: to the degree to which the pre-treatment fit is good, the synthetic control is assumed to be a valid counterfactual for the treated unit post-treatment, and any post-intervention discrepancy between the observed and counterfactual outcomes can be attributed to the intervention itself.

The synthetic control method requires time-series data. To formalize how it deals with the counterfactual problem, it is therefore necessary to introduce a time dimension to the potential outcomes framework outline above. Consider a panel dataset with N units observed over T periods, one or more of which is at some point exposed to the treatment. The potential outcomes for unit i at time t are therefore $Y_{i,t}(1)$ and $Y_{i,t}(0)$, denoting the value with and without the treatment, respectively (Doudchenko & Imbens 2016, 2). All observations can be summarized as part of one of four matrices depending on whether the unit is treated or a control and whether the observation is made before or after the treatment occurred: $\mathbf{Y}_{c,pre}^{obs}$, $\mathbf{Y}_{c,post}^{obs}$, $\mathbf{Y}_{t,pre}^{obs}$, and $\mathbf{Y}_{t,post}^{obs}$, the latter of which represents the treated outcomes, so that:

$$\mathbf{Y}^{obs} = \begin{pmatrix} \mathbf{Y}_{t,post}^{obs} & \mathbf{Y}_{c,post}^{obs} \\ \mathbf{Y}_{t,pre}^{obs} & \mathbf{Y}_{c,pre}^{obs} \end{pmatrix} = \begin{pmatrix} \mathbf{Y}_{t,post}(1) & \mathbf{Y}_{c,post}(0) \\ \mathbf{Y}_{t,pre}(0) & \mathbf{Y}_{c,pre}(0) \end{pmatrix}$$

Recall that the causal effect for a given unit in a given period can be calculated as $Y_{i,t}(1) - Y_{i,t}(0)$. To estimate the effect of the treatment we therefore need to compare the matrices $\mathbf{Y}_{t,post}(1)$ and $\mathbf{Y}_{t,post}(0)$. However, as we only observe the former, the problem of inference turns into one of forecasting missing data (Xu 2017, 61).

A fundamental identification assumption of the synthetic control method is that the counterfactual outcomes for treated units – the matrix $\mathbf{Y}_{t,post}(0)$ – can be predicted as a linear combination of control unit outcomes (Doudchenko & Imbens 2016, 8). Weights are selected for each control unit so that the difference between the synthetic control and the treated unit is minimized over the pre-treatment period. One approach is to specify certain variables thought

¹⁰ For a Bayesian alternative to causal impact evaluation in a counterfactual framework, see Brodersen et al. (2015).

to have predictive power over the outcome. By matching on these, one can increase the probability that the similarities between treated and control units extend beyond the surface-level to the underlying causal processes (Abadie et al. 2015, 498). Alternatively, the researcher can match solely on the pre-treatment trajectory of the outcome. Abadie et al. (2015, 498) suggest that this may help better control for *unobserved* confounders. I discuss this issue more in depth in the section on implementation below. In any case, to be a useful counterfactual, the trend in the dependent variable of the synthetic control should closely resemble that of the treated unit in the pre-treatment period.

Relative to a qualitative comparative design, the synthetic control method has several advantages. As mentioned above, a key issue in comparative case studies concerns the selection of comparison units. When studying cases drawn from a relatively small population – as is often the case when studying aggregate units like states or countries – a suitably similar comparison may not exist (Abadie et al. 2015, 500). By combining data from several control units the synthetic control method solves this problem. Additionally, it provides a systematic and data-driven way of selecting these units in a way that maximizes comparability, which enhances transparency and reliability (Abadie et al. 2015, 496). Finally, by plotting the predicted counterfactual alongside the treated unit in the pre-treatment period, synthetic control allows intuitive interpretation of its predictive power and – consequently – the validity of the results obtained (Samartsidis et al. 2019, 494).

4.2.1 Generalized synthetic control

While the original synthetic control method was designed as a form of case-study with only one treated unit, newer variations allow for both multiple treated units and variable treatment timing – a necessity for the present application, given that several states have adopted voter ID-requirements and have done so at different times. One approach is the generalized synthetic control method (GSC) developed by Xu (2017), which simulates the counterfactual scenario using an interactive fixed effects (IFE) model interacting time-varying coefficients with unit-specific intercepts.

The GSC imputes treated counterfactuals in three steps (Xu 2017, 63). First, the control group data is used to estimate an IFE model and obtain the time-varying coefficients (factors). Using a cross-validation procedure, the algorithm automatically selects the number of latent factors. Second, unit-specific intercepts (factor loadings) are estimated for each treated unit so that the mean square prediction error (MSPE) – or the average difference between the actual and predicted outcome – is minimized over the pre-treatment periods. This is analogous to the original synthetic control estimator seeking to match the treated unit as closely as possible prior to intervention. Finally, the obtained factors and factor loadings are used to calculate counterfactual outcomes for the treated units in the post-treatment periods. This latent factor-approach allows the GSC to capture the effect of unobserved confounders if they can be decomposed into interactions between time-varying common shocks and unit-specific intercepts (Xu 2017, 4; Liu et al. 2020, 2). This includes time-varying confounders and is a key advantage relative to the difference-in-differences approach detailed below (Samartsidis et al. 2019, 489).

There are a number of other advantages to the GSC (Xu 2017, 59). While the original synthetic control method could technically accommodate several treated units by individually creating a synthetic control for each, the GSC does this in a single run. Additionally, it provides a simple and automated way of specifying the IFE model through the cross-validation procedure mentioned above. Finally, the GSC produces frequentist uncertainty estimates that make the validity of inferences straightforward to assess. This is achieved through a nonparametric bootstrap procedure when the number of treated units is large. When N_{tr} is small, a parametric bootstrap is employed, resampling the entire time series of residuals to preserve the serial correlation within units, which in turn avoids underestimation of the standard errors (Xu (2017, 64). Through a number of simulation tests, Xu (2017, 67-68) show that the GSC produces relatively unbiased estimates and performs favorably compared to alternative estimators, including the original synthetic control estimator.

4.2.2 Matrix completion method

The matrix completion method (MCM) offers an alternative, yet similar way to simulate counterfactual outcomes for treated units/periods. Again treating the counterfactual problem as one of missing data, the MCM adopts a latent factor approach like that of the GSC, though it

does not explicitly estimate factors and loadings (Liu et al 2020, 11). Instead, the missing values are imputed directly in a matrix under the assumption that: “(i) the complete matrix is the sum of a low rank matrix plus noise and (ii), the missingness is completely at random” (Athey et al. 2018, 2). Athey et al (2018) generalize this approach to situations where the missingness is not random but rather has a systematic time series structure, as is the case with the present application. Similar to the GSC, the MCM uses a cross-validation procedure to select an estimator that minimizes pre-intervention MSPE, with certain restrictions (Athey et al. 2018, 10-12). Like the GSC, the MCM can also accommodate both multiple treated units and variable treatment timing, as well as capture the effect of unobserved confounders (Athey et al. 2018, 2; Liu et al. 2020). Below, I discuss the practical differences between the two estimators and their relative merit for this study.

4.2.3 Assumptions

Synthetic control – using both GSC and MCM – relies on a number of assumptions being satisfied in order to select the correct model and provide valid inferences. Because the estimator will still impute values even if certain assumptions are violated, it is not made immediately clear to the analyst when the results are inaccurate (Xu 2017, 59). Given this, it is crucial to make these assumptions explicit.

The main assumption necessary for causal identification is that of strict exogeneity – that the error terms are independent of treatment assignment as well as both observed and unobserved covariates (Xu 2017, 61; Athey et al. 2018, 10, 16). Also known as the unconfoundedness or ignorability assumption, this is to say that, conditional on the model, there are no additional covariates confounding the relationship between treatment and outcome (Imai 2017, 371-372). Treatment assignment is then as-if random, without systematic differences between treated and control groups, as thus ignorable (Keele & Minozzi 2013, 195). Note that this assumption is more plausible for the GSC and MCM than many other methods, due to their ability to control for even unobserved and time-varying confounders (Xu 2017, 62).

Additionally, we assume only a weak serial dependence of the error terms, whereby error terms are independent both across units and over time. Xu (2017, 62) also specifies moment

conditions that ensure the convergence of the estimator. When using a parametric bootstrap procedure to generate uncertainty estimates, a further assumption is homoscedasticity of the error terms across units – though not necessarily across time (Xu 2017, 62).

A general assumption when using time-series data is that of stationarity: that the mean and variance are constant over time, and that the covariance between values at any two different points in time is a function only of the distance between them and not time (Dougherty 2011, 465). Testing the dependent variable for stationarity shows this assumption to hold.¹¹

Finally, it is necessary to explicate and discuss a fundamental assumption for most counterfactual analyses: the stable unit treatment value assumption (SUTVA). Put simply, SUTVA states that the effect of the treatment on a given unit is the same regardless of the treatment status of other units, and that there are no different or varying versions of the treatment assigned (Imbens & Rubin 2015, 10). SUTVA is what allows us to conceptualize potential outcomes: if units' potential outcomes depend on the treatment status of others, or the treatment is not fixed, there are no single $Y_{i,t}(1)$ and $Y_{i,t}(0)$ with which to calculate the causal effect (Morgan & Winship 2007, 38). We must therefore consider whether SUTVA will hold in the present application. First, non-interference appears a reasonable assumption: even though ID-less voters in a state with ID-requirements may be turned away at the polls, those in a different, non-ID state will be allowed to vote all the same. One possible exception concerns the deterrent effect, if voters in other states are confused as to whether the new requirements apply in their state as well and therefore decide to stay home. Though not impossible, chances of this occurring on a large scale seem slim. Second, there is the issue of non-variation of the treatment, which is particularly delicate when evaluating policy interventions, as they usually consist of numerous components (Keele 2015, 317). Because each state's voter ID-law is a unique piece of legislation, there is considerable variation between them, arguably threatening this assumption. However, focusing the analysis on the sub-category of strict photo ID-laws should reduce variation and help broadly satisfy this part of SUTVA.

¹¹ $p < 0.01$ in an augmented Dickey-Fuller test.

Xu (2017) advocates conducting diagnostic tests to safeguard against misspecification, specifically plotting the raw data and the predicted counterfactuals, as well as plotting the factor loadings of both treated and control units if using the GSC (Xu 2017, 73). When treated and control units share common support the model can estimate by interpolation rather than potentially erroneous extrapolations (King & Zeng 2006). As mentioned above, visualizing the pre-intervention fit of the synthetic control also provides a useful way to assess the predictive power of the model. The above diagnostics, along with robustness checks, are presented and discussed in chapter 6.

4.2.4 Implementation

First, a note regarding the causal estimand of the analysis. Because we are using observational data and treatment assignment therefore is nonrandom we are not estimating the ATE. Rather, we are only estimating the effect of the treatment on those units that did actually receive it – the average treatment effect on the treated (ATT) (Gerring 2012, 221; Xu 2017, 61). This is because we are only imputing untreated counterfactuals for the treated units. Technically we could also impute treated counterfactuals for the untreated units, but this requires a substantial number of treated units/periods or additional assumptions (Athey et al. 2018, 23). Doudchenko & Imbens (2017, 3) point out that in many applications it can be difficult to even conceptualize a treated state for control units. Even when this is not the case, the ATT can be just as meaningful as the ATE. The question of how turnout levels would change if *all* states adopted voter ID-requirements is interesting, but given the relative implausibility of this scenario it is arguably equally relevant to investigate what the impact has been in the states that have actually done so – which is the purpose of this thesis.

Synthetic control is fairly demanding in terms of data requirements. According to Xu (2017, 73), the GSC requires at least 10 pre-treatment periods and 40 control units to reliably produce unbiased estimates. Meanwhile, Athey et al. (2018, 21) demonstrate that the MCM performs well when datasets are asymmetrical with either $N \ll T$ or $N \gg T$. Given that my dataset is thin ($N = 49$, $T = 11$) with 42 control units and 7-9 pre-treatment periods, I will employ the MCM for my main analysis. As a robustness check, I also estimate the effect of voter ID-laws using the GSC. Given the relatively small number of treated units ($N_{tr} = 7$), a parametric bootstrap procedure is used to generate uncertainty estimates.

Like the original synthetic control estimator, both GSC and MCM can incorporate additional covariates to improve model fit (Xu 2017, 8; Athey et al. 2018, 22). I do not include any covariates, however, which warrants explanation. The specification of covariates is somewhat controversial in the literature. Writing about the original synthetic control estimator, Ferman et al. (2020) argue that the lack of consensus on which covariates to include allows for specification searches on the part of the researcher, which threatens the key advantage of synthetic control as a data-driven process. They therefore recommend that control units used to generate the synthetic control be chosen purely based on similarity on the outcome variable, and then choosing the model that exhibits the lowest pre-treatment MSPE (Ferman et al. 2020, 522-523). Additionally, Doudchenko & Imbens (2017, 20) note that predictor variables tend to play a minor role in practice: “In terms of predictive power the lagged outcomes tend to be substantially more important, and as a result the decision how to treat these other pre-treatment variables need not be a a [*sic*] very important one”. Finally, poorly chosen predictor variables with little predictive power over the outcome would actually lead to a worse model, as donor units used to generate the synthetic control would be selected in part based on these causally irrelevant criteria rather than their similarity in outcomes. Estimating the counterfactual based solely on the pre-treatment trend in the dependent variable eliminates this danger and – by obviating the need for researcher discretion in specifying which covariates to include – arguably makes the analysis less vulnerable to misspecification and model dependence.

I conduct the analysis in RStudio using the package “gsynth”, through which both MCM and GSC can be implemented (Xu & Liu 2018). The script is available upon request.

4.3 Difference-in-differences

For reasons of data availability explained in the next chapter, investigating the hypothesized differential impact of voter ID-laws among minority voters requires a change in empirical strategy. For this secondary analysis, I employ a difference-in-differences (DID) design.

Consider two groups – one treated and one control – and two periods – one pre- and one post-intervention. Comparing the average treated and control outcomes in the post-period would yield biased estimates if the two groups had different initial values prior to treatment, due to some between-group difference resulting from non-random treatment assignment. Meanwhile, comparing treatment group outcomes before and after intervention ignores that the outcome might have changed over time even in the control group, due to some common factor. The difference-in-differences approach combines these two methods of comparison – cross-sectional and longitudinal – to reduce the aforementioned threats to inference (Leighley & Nagler 2013, 98). Rather than compare single values, the researcher compares the difference between the average outcomes of the treated group before and after treatment ($Y_{t,post} - Y_{t,pre}$) with the corresponding difference for the control group ($Y_{c,post} - Y_{c,pre}$). The difference between these differences represents the effect of the treatment (Imai 2017, 62). Concretely, the identification strategy here is to use the change in turnout of the control group as a substitute for the counterfactual turnout-change of the treated group absent treatment (Highton 2017, 159). Meanwhile, allowing for the two groups having different absolute values eliminates any fixed (time-invariant) confounders between the two groups (Angrist & Pischke 2014, 203).

The fundamental assumption of DID designs is that of parallel trends: that the development of the outcomes of the treated and control groups would have moved in tandem over time had no intervention occurred (Angrist & Pischke 2014, 204). This is a strong assumption that is likely to be violated if time-varying confounders are present. Furthermore, it cannot be tested directly. However, an informal diagnostic is to plot outcomes of the treated and control groups in the pre-treatment period to check whether the assumption holds prior to intervention (Imai 2017, 63). While finding the assumption to hold here does not guarantee that it would have held post-treatment, deviation from parallel trends in the pre-treatment period casts serious doubt on its plausibility.

Difference-in-differences designs are similar to synthetic control approaches in that they combine cross-sectional and time-series data to combat the counterfactual problem. However, the reliance of DID on the parallel trends assumption is a key weakness not shared by the synthetic control method. DID controls for unobserved confounders but assumes that the effects of these are constant over time; furthermore, it can only capture temporal trends which affect

both treated and control units (Samartsidis et al. 2019, 489). GSC and MCM do not require such assumptions, as they can accommodate time-varying as well as time-invariant confounders (Xu 2017, 57). For this reason, they are arguably stronger tools for causal inference. However, insufficient data availability means we cannot estimate state-level turnout rates by demographic group to generate a panel dataset. This makes synthetic control unfeasible for the secondary analysis of this paper, something I will expand upon in the next chapter. For this reason, I employ a DID design to investigate the hypothesis regarding the differential impact of voter ID-laws on minority voters.

4.3.1 Implementation

I employ a two-group, two-period design. While some applications of DID average data over multiple years for each period, this is unfeasible here (Bertrand 2004). This is due to the variability of treatment timing, which means there is no common pre- and post-treatment period. Rather, I use data from two election years, one – the pre-period – being the last presidential election before any states implemented strict photographic ID-requirements (2004) and another – the post-period – being the most recent presidential election for which data is available (2016). Choosing this and not some earlier election as the post-period maximizes the number of states under treatment.

To answer hypothesis H2, I must compare turnout across racial/ethnic categories. I describe the dataset used in more detail in the next chapter; for present purposes, it suffices to say that it consists of survey data on American individuals from the Current Population Survey (CPS). Respondents living in a state which had voter ID-laws in effect for the 2016 election constitute the treated group, while those living in other states constitute the control. Note that respondents are not the same between the two years; however, a key advantage to the DID design is that it does not require panel data, since the analysis is conducted at the group-level. The surveys are designed to be nationally representative in the year they are conducted. Essentially, we are comparing a representative sample of the US population in 2004 with a 2016 equivalent. The treated group is thus assumed to be representative of all Americans living in treated states, while the control group is assumed to be representative for the remainder of the population living in untreated states.

To assess the hypothesis, I calculate the DID using only data from respondents belonging to an ethnic/racial minority. I then estimate a separate DID for non-minority respondents. Comparing the two estimates reveals whether one group was affected more strongly by the implementation of voter ID-requirements than the other.

Note that the unit of analysis here is not states and their turnout rates, but individuals and whether they voted. The treatment effect estimated concerns the proportion of people voting in each group in the sample, i.e. the effect on the turnout rate of minority voters living in treated states versus the effect on the turnout rate of non-minority voters in treated states. Because the sample is nationally representative, we can infer to the broader American population. Observe also that we are again estimating the ATT, not the ATE; we cannot infer what the effect of voter ID-laws would be if *all* states were to adopt them.

While the data are collected at the individual-level, treatment assignment (adoption of voter ID-laws) occurs at the state level. Since the analysis pools respondents from all states into one of two groups, to avoid underestimating the standard errors due to within-cluster correlations among respondents, I therefore follow the recommendations of Abadie et al. (2017) and use cluster-robust standard errors to generate uncertainty estimates, clustering on states.

The data are weighted to account for the sampling process using accompanying weights from the CPS. To account for overreporting in the CPS sample, I also adopt the weighting scheme suggested by Hur & Achen (2013) – see the next chapter for details.

I implement the DID in RStudio using the “`lm.cluster`” function from the package “`miceadds`”, which enables cluster robust standard errors (Robitzsch & Gruns 2020). The script is available upon request.

5 Data

In this chapter, I present the data used in the analysis. I begin by introducing the dataset and variables of the primary analysis with the synthetic control method. I then move on to the data used in the secondary analysis on differential impact, explaining how poor data availability necessitates a change in methodological approach, as well as discussing the associated drawbacks and how they may be alleviated.

5.1 Dataset – Synthetic control

For my main analysis, I require data on state-level turnout over time. Unhelpfully, the federal US government neither comprehensively collects nor publishes such data itself. Instead, I use data from the United States Elections Project, compiled from official state government records by Michael P. McDonald. Combining four datasets yields a complete panel of turnout rates in all US states plus the District of Columbia in every national election year from 1980 to 2020 (McDonald 2021a).¹²

While including all national elections – which occur every other year – maximizes the number of observations, for reasons explained below I restrict the analysis to presidential election years – which occur every four years.¹³ Additionally, I remove Texas and Virginia from the dataset.¹⁴ Texas had strict photo ID-requirements in effect for the 2014 election, and Virginia in 2014 and 2016, but both have since had their laws blocked or overturned so that they were not implemented in subsequent election years. Because the synthetic control method assumes that units remain in the treated group once treated, I exclude these states from the analysis.¹⁵ The final dataset thus has $N = 49$ and $T = 11$.

¹² Datasets (all collectable from McDonald 2021a): “1980-2014 November General Election”, “2016 November General Election”, “2018 November General Election”, “2020 November General Election”.

¹³ 1980, 1984, 1988, 1992, 1996, 2000, 2004, 2008, 2012, 2016, and 2020.

¹⁴ The results are substantively unchanged when retaining these states in the analysis.

¹⁵ For work on identifying the effects of treatments that switch on and off repeatedly, see “Matching Methods for Causal Inference with Time-Series Cross-Sectional Data”, a working paper by Kosuke Imai, In Song Kim & Erik Wang.

5.1.1 Dependent variable

The dependent variable of the analysis is state-level turnout rates. The key measurement issue is therefore determining which components to use in the calculation thereof. Beginning with the numerator, McDonald reports the total votes cast for the highest office in a given election, or, alternatively, the office for which the most votes were cast (McDonald 2021d).¹⁶ In a presidential election year, this is almost always the presidential race. In midterm election years, however, votes are reported for the statewide office with the highest vote total – governor or US senator – or, if no statewide office is on the ballot, the sum of congressional races. Thus in a given midterm election-year some states’ reported turnout rates are those for congressional seats, others’ those for a senatorial seat and yet others’ those for a governorship. This threatens the comparability of turnout rates across states if elections for one office are more salient to voters than another – for instance if a gubernatorial election is viewed as more important by voters – and thus draws more of them to the polling station than a congressional race. This idiosyncratic variation in turnout could bias the estimation of the treatment-effect. If turnout for congressional races only was reported separately for both presidential and midterm election years, one could compare apples to apples both across states and over time. However, such data is not available. For this reason, I restrict the analysis to presidential election years to enhance comparability.

Still, another benefit to restricting the analysis to presidential elections – even if comparable data on congressional elections were available – is that it holds constant across states a number of election-specific factors that may affect turnout, for instance relating to the candidates running for office. While state-level turnout for a congressional election is effectively an aggregation of multiple, unique local races, presidential elections are – by virtue of consisting of a single race for a single national office – highly similar across states. Comparing turnout for presidential elections only thus has an intrinsic benefit of comparability.

Finally, the denominator of the turnout calculation presents a particular puzzle in the American context. As there is no official national population register in the US, researchers must specify

¹⁶ McDonald also collects data on “Total ballots counted”, which includes blank votes and votes cast for multiple candidates where only one is acceptable. Though he considers this the best measure of the total number of people turning out to vote, it is not used here as it is unavailable for a large number of states/years.

the population from which to calculate the percentage of voters (Leighley & Nagler 2013, 22). A common approach is to construct an estimate of the voting age population (VAP) using Census data, i.e. people over the age of 18. The VAP, however, does not necessarily reflect the number of people actually eligible to vote, because it counts both non-citizen residents and disenfranchised felons. Furthermore, both felony disenfranchisement laws and the number of non-citizens vary across states and over time; for instance, turnout falling over time is at least partially explained by an increase in the prison population rather than a true decline in participation (McDonald 2021e). For this reason, McDonald estimates a new denominator measuring the voting *eligible* population (VEP), which adjusts the VAP for the number of non-citizens and disenfranchised felons, to avoid skewing the turnout estimate. The estimate is imperfect, but it is likely more accurate than the unadjusted VAP denominator (McDonald 2021c). I therefore estimate my model using the VEP turnout rate for the highest office.¹⁷

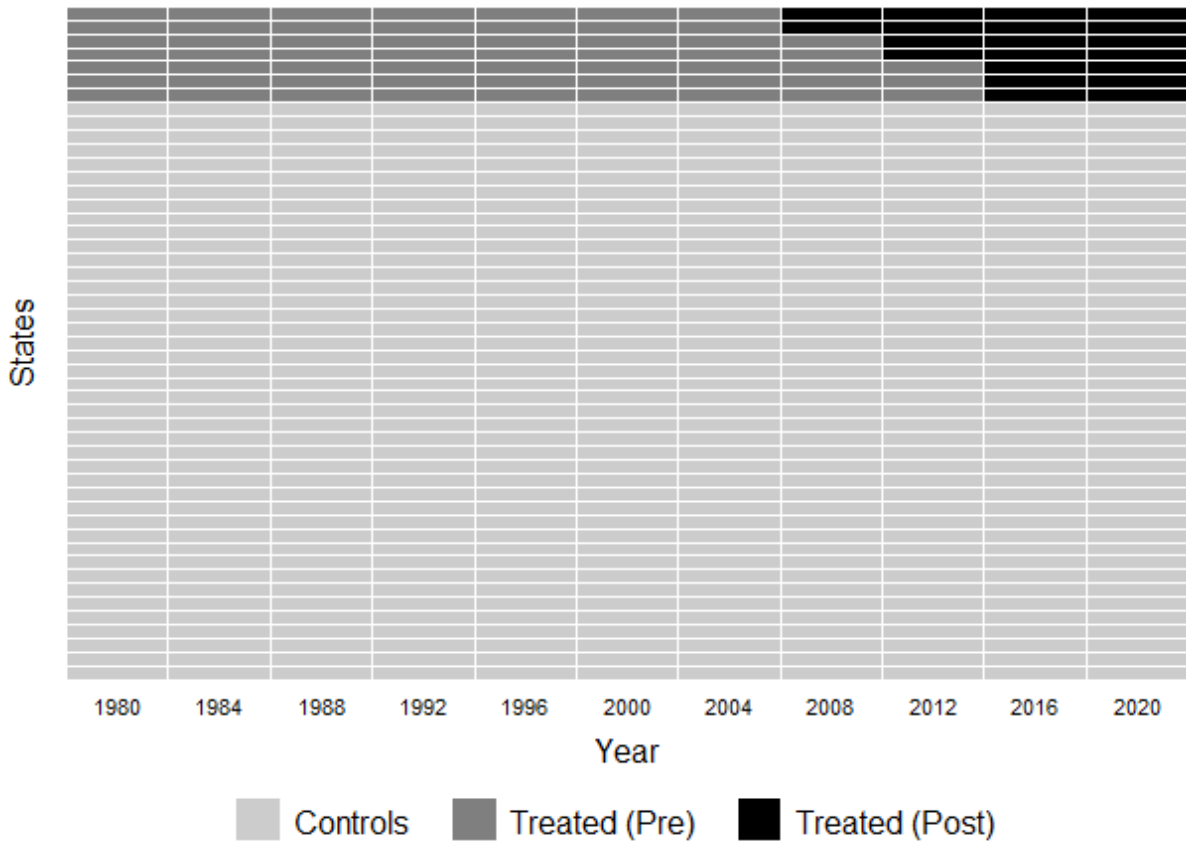
5.1.2 Treatment variable

The treatment variable indicates whether a state had strict photographic voter ID-requirements in effect for a given election – see table 2.2 in chapter 2 for more. The variable is coded dichotomously, with a value of 1 if a state is treated and 0 if it is not. The total $N_{tr} = 7$.¹⁸ Note that states are only coded as treated if they actually implement their ID-requirements. Several states have passed bills adopting strict photo ID-requirements only to have them blocked or overturned by court challenges. These states are coded as untreated. Similarly, treated states are only coded as such from the first election in which ID-requirements are implemented, rather than from the year of adoption. Figure 5.1 visualizes the distribution of treated and untreated states/years. Each row represents a state, and each panel an election in that state. The first seven rows are the treated states – i.e. states that have implemented strict photo ID-laws – with dark grey panels representing election-years prior to treatment and black panels representing elections during which the treatment was in effect. Meanwhile, the light grey rows represent the control states from which the synthetic counterfactual is generated – these are never under treatment.

¹⁷ The results are substantively unchanged when using the VAP turnout rate for highest office.

¹⁸ Georgia, Indiana, Alabama, Kansas, Tennessee, Wisconsin, and Mississippi.

Figure 5.1 *Treatment status by state and election*



5.2 Dataset – Difference-in-differences

Investigating the hypothesis of differential impact on minority voters requires comparing turnout data for different demographic groups. This, however, proves problematic. Because information on individual voter characteristics like race is not collected at the polling station, no official data on turnout among such groups exists. Rather, researchers must rely on aggregating micro-level survey data to estimate group-level turnout rates. I use data from the Current Population Survey (CPS), a commonly used survey when analyzing US turnout.

Conducted by the Census Bureau and the Bureau of Labor Statistics, the CPS surveys over 65,000 US households monthly on various labor force and demographic questions (IPUMS CPS 2021a). In addition to this basic monthly survey, additional question batteries are added for certain months. Among these is the Voting and Registration Supplement, included

biennially in November of every election year, wherein respondents are polled on various questions related to voting behavior, including whether they voted or not in the November election (IPUMS CPS 2021b). The data are retrieved from the IPUMS CPS project, which integrates all CPS surveys from 1962 onwards while providing quality-of-life improvements like harmonized variables and a data-extraction tool (IPUMS CPS 2021a).

In addition to a relatively low non-response rate, the main benefit of the CPS is its uniquely large sample size (Leighley & Nagler 2013, 19). In each household, the respondent is asked to provide information on behalf of other household members, which increases the sample size further. Besides the lower sampling error, a large N is a necessity for analysis of group-level differences (Leighley & Nagler 2013, 18). Note how the household-level reporting could lead to misreporting if other household-members misinform the respondent. Below, I discuss the problem of overreporting, which is of particular relevance here.

I merge data from the 2004 and 2016 November surveys (Flood et al. 2020). I then remove units from Texas, for reasons explained above.¹⁹ I also remove those units classified as “Not in universe” (see below), as well as those without a clear Yes/No response to the voting question: those who refused to answer, did not know, or were never asked the question at all. Some – including the Census Bureau – retain these latter groups when calculating turnout, counting them all as having not voted, likely leading to an overestimation of turnout (Hur & Achen 2013). Says McDonald (2021b): “To underscore how poor of an [sic] practice this is, the Census Bureau counts persons *who were never asked the voting question* as having not voted [emphasis in original].” He therefore recommends removing these units. After cleaning, the final dataset has $N = 159.737$.

5.2.1 Why not use the CPS data for synthetic control?

Despite its large N, the CPS data are insufficient to conduct a synthetic control analysis, because – as demonstrated in the previous chapter – synthetic control requires a panel dataset with long time-series of turnout rates by state, like the one used for the main analysis. The CPS is designed

¹⁹ Units from Virginia remain in the analysis, as Virginia had not yet left the treated group in 2016.

to be representative at the state level (United States Census Bureau 2019). However, despite the large N, separating between demographic groups *within* states – as hypothesis H2 requires – will likely not yield reliable estimates of group-level turnout due to the low N for certain races/ethnicities (Fraga 2018, 99-100). Thus, it is not feasible, for instance, to aggregate CPS data into a panel-dataset of minority turnout by state, as one would have to in order to implement the synthetic control method. It is for this reason that I rather employ a DID design, in which respondents are pooled into a single treated group and a single control group, without differentiating between individual states. As discussed in the previous chapter, this is not optimal relative to the synthetic control method, but it offers a passable replacement.²⁰

5.2.2 Dependent variable

The dependent variable of the analysis is a binary measure of whether the respondent reported to have voted in the November elections of that year, constructed from the CPS variable *VOTED*. Units are recoded as 100 if they voted and 0 if they did not. This is simply a practical choice to have the results return as percentages rather than decimals, which makes the results more intuitively readable as turnout rates.

When it comes to calculating turnout, we must again confront the issue of specifying a numerator and denominator discussed previously. I begin with the denominator, which in this instance is the number of respondents in the sample from each group under analysis. Data from the Voting and Registration Supplement of the CPS is restricted to citizens above the age of 18 (IPUMS CPS 2021b). Furthermore, the CPS – by virtue of being a survey of private households – naturally excludes the incarcerated population. It does, however, include non-institutionalized disenfranchised felons (Leighley & Nagler 2013, 20). There are a number of respondents who are citizens above the age of 18 yet are labelled “Not in universe”, though the reason why they are ineligible to vote is unknown, and thus the degree to which this group overlaps with McDonald’s measure of disenfranchised felons is unclear (IPUMS CPS 2021b). As mentioned,

²⁰ Another dataset commonly used when studying US turnout, the American National Election Survey, has a considerably smaller sample size than the CPS, while a third, the Cooperative Congressional Election Study, is only available from 2006 onwards. As far as I am aware, no publicly available dataset exists that offers (a sample size large enough to generate) accurate state-level estimates of minority/non-minority turnout over time of the kind that would enable a synthetic control approach.

these individuals are removed from the analysis. The denominator thus approximates the VEP, albeit imperfectly.

The turnout-numerator here is the number of respondents who report to have voted. This, too, is not without issues. Self-reported data on voting is vulnerable to misreporting where respondents provide inaccurate answers, likely due to “an eagerness on the part of respondents to report socially desirable behavior; in this case political participation” (Fraga 2018, 100).²¹ The CPS is widely acknowledged to be overestimating turnout (Fraga 2018; McDonald 2021b; Leighley & Nagler 2013). For the purposes of the present application, two patterns are worth considering. First, some states have higher overreport biases than others, in ways that are stable over time, in that “states with high rates of overreporting in one year tend to have high rates in other years, while those with low rates tend to stay low” (Bernstein et al. 2003, 368). Second, some demographic groups have systematically higher rates of over-reporting than others. Using data from six states in which race is self-reported as part of the voter registration process as a comparison, Ansolabehere et al. (2021) show that the CPS overestimates Black and Hispanic turnout relative to non-Hispanic whites.

The CPS data, then, has varying rates of over-report bias both between demographic groups and between states. The difference-in-differences design, however, helps address both these issues by focusing on *changes* in turnout rates. First, we are not directly comparing turnout rates between minorities and non-minorities. Rather, we compare minority voters in treated states with minority voters in control states. Second, while comparing turnout rates within groups directly would still yield biased estimates if treated states have systematically different rates of over-reporting than the control states, the DID design means we are comparing the *trends* in turnout rather than the absolute values. Recall from chapter 4 that allowing permanent differences in the outcome is a key advantage of this method. We must therefore make two assumptions: one, that the differences in misreporting rates between racial groups are stable across states; and two, that the overall differences in misreporting between states are stable across time (Leighley & Nagler 2013, 20). These are strong assumptions, to be sure, but more plausible than assuming – erroneously – that there are no differences at all.

²¹ See also DeBell et al. (2020).

As an additional measure, I implement the weighting scheme suggested by Hur & Achen (2013) to account for the overreport bias and correct overall state-level turnout rates to official VEP rates as reported by McDonald (2021b). Essentially, voters among the CPS respondents are weighted down, while non-voters are weighted up (Hur & Achen 2013, 991). Hur & Achen (2013, 992) write:

Roughly speaking, this procedure will be statistically successful if, after weighting, the self-reported voters in a particular state who responded to the Voting Supplement are, in expectation, representative of all actual voters in that state. A parallel requirement holds for self-reported nonvoters.

They acknowledge that these assumptions are unlikely to hold completely, but argue that theirs is the current best available adjustment until the Census Bureau itself provides more sophisticated weights. This scheme combines with the CPS sampling weight *WTFINL*.

5.2.3 Race/ethnicity variable

To divide respondents by demographic groups, I create a variable for whether they belong to a racial minority. The CPS allows respondents to self-select one or more of a number of racial categories (CPS variable: *RACE*). Additionally, a separate question is included for whether they are ethnically Hispanic, and if so, to which country they trace their lineage (CPS variable: *HISPAN*). However, as mentioned, some categories have a very low number of respondents, particularly in certain states. To avoid truncating the sample to the point of threatening inference, I therefore collapse the various CPS measures of race and ethnicity into a single, dichotomous variable indicating whether a respondent belongs to a minority or not. This distinction remains substantially meaningful as well, given that blacks and Hispanics (by far the two largest minority groups in the sample) both exhibit lower rates of ID holding than whites, and therefore are expected to be negatively affected by voter ID-requirements as per hypothesis H2 (Highton 2017, 160-161). Respondents are coded with the value 1 if they are a non-Hispanic white (non-minority), and 0 if they belong to any other racial/ethnic group (minority).

5.2.4 Treatment indicator

To estimate the difference-in-differences, we must separate respondents into treated and control groups. Respondents living in states with strict photographic voter ID-requirements in effect for the 2016 elections are pooled together and designated as the treated group.²² These are coded as 1. Respondents living in the remaining, untreated states make up the control group, coded as 0. Note that due to the variation in timing regarding when states implemented voter ID-laws, respondents from some states have been exposed to the treatment for longer than others in the post-treatment observation period. If, as Vercellotti & Anderson (2009) propose, the magnitude of the treatment-effect declines over time, the DID will underestimate the effect for the states that implemented ID-requirements early relative to those with more recent laws. This is a genuine limitation on unit comparability – resulting from the two-period nature of the research design – and is worth keeping in mind when interpreting the results.

²² Georgia, Indiana, Alabama, Kansas, Tennessee, Wisconsin, Mississippi, and Virginia.

6 Results

In this chapter, I present my findings, beginning with the synthetic control method before moving on to the difference-in-differences. Neither analysis finds evidence of a significant effect of ID-requirements on turnout. In addition to the main results, I conduct various robustness checks using alternative model specifications, which yield results similar to those of the main analysis.

6.1 Do voter ID-requirements lead to lower turnout? Results from the synthetic control analysis

First, some preliminary descriptive statistics. Figure 6.1 shows the trend in turnout rates over time for both treated and untreated states. Two observations are immediately available. First, the treated states seem to exhibit broadly similar trajectories to the untreated states – at least superficially. Second, the treated outcomes are all within the convex hull of the control outcomes in the pre-treatment period, meaning the synthetic control can likely be estimated without excessive extrapolation (Abadie et al. 2010, 502; King & Zeng 2006).

Table 6.1 displays the mean turnout rates, as well as the standard deviation and the lowest and highest single values, by group. Note that because there is no single time of treatment there is no control group equivalent to the pre- and post-treatment averages for the treatment group. When averaged over all periods, the treated group turnout is 3 percentage points lower than that of the control group. Post-treatment, however, the mean treated turnout is .5 percentage points higher than the control average. Relative to the pre-treatment treated mean, post-treatment turnout is 4.7 points higher. This cursory examination thus suggests a potential *positive* effect of implementing strict photo-ID requirements. Naturally, this simplistic analysis is insufficient for causal inference.

Figure 6.1 Turnout trends, treated versus untreated states

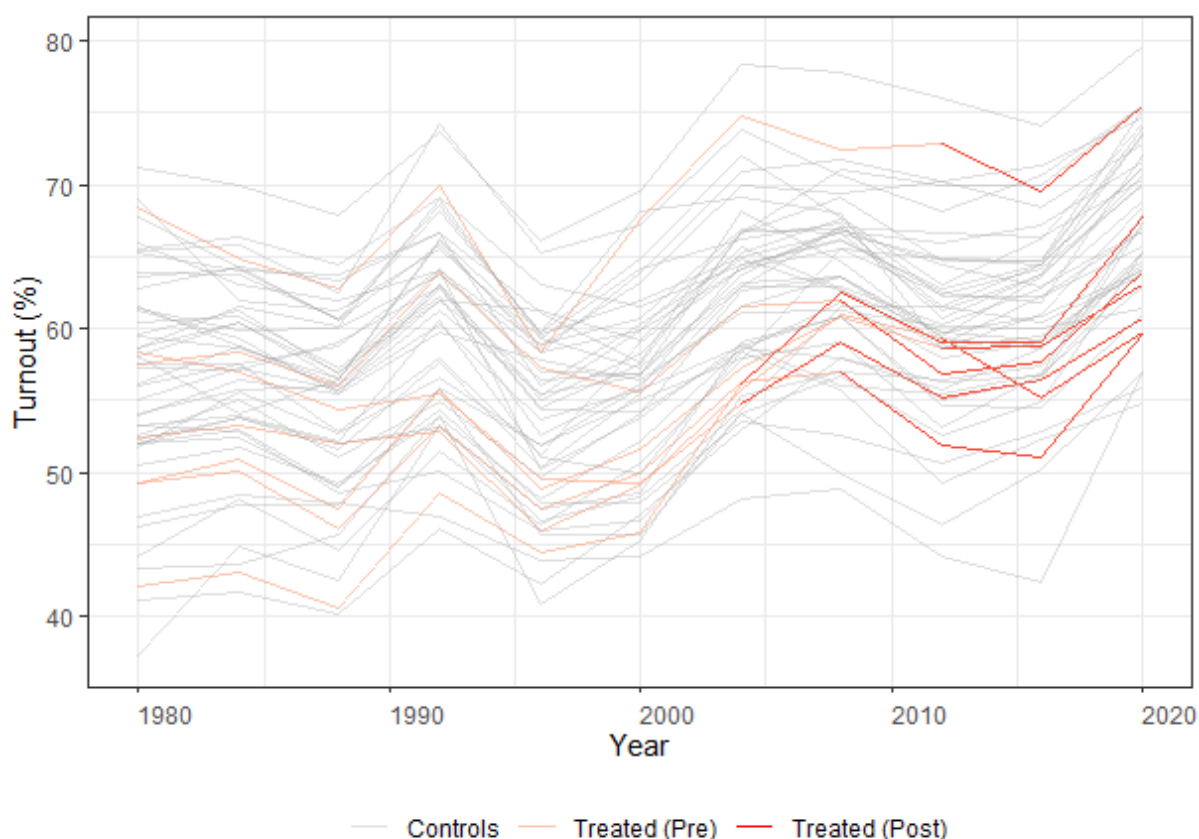


Table 6.1 State-level turnout rates

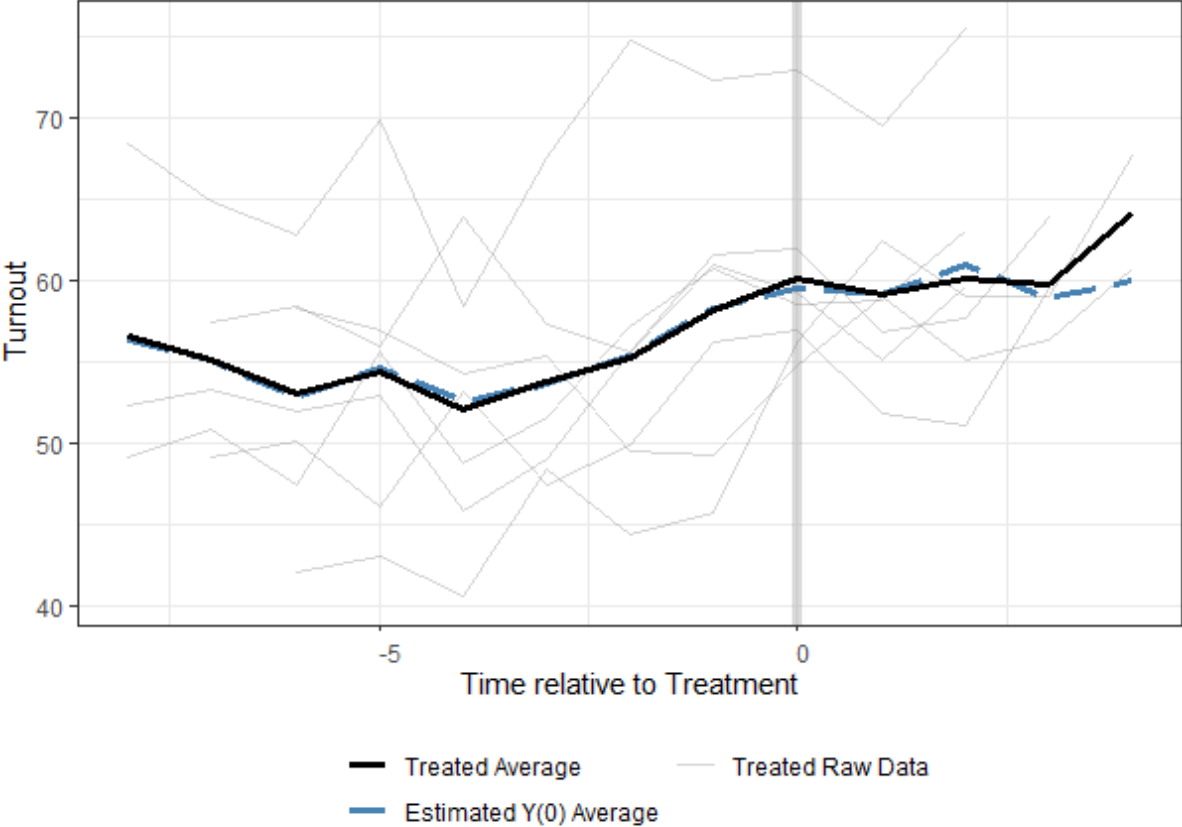
	Mean	Standard deviation	Min	Max
All	59.2	7.6	37.2	79.6
Control group	59.6	7.5	37.2	79.6
Treated group	56.6	7.4	40.6	74.8
Treated group (pre-treatment)	55.4	7.8	40.6	74.8
Treated group (post-treatment)	60.1	4.7	51.1	69.5

6.1.1 The effect of strict photographic voter ID-requirements on turnout

Figure 6.2 shows the results of the matrix completion method. The black line represents the mean turnout over time of the treated states, which are individually plotted in grey. Meanwhile, the dotted line is the estimated counterfactual average turnout, generated using data from the control states. This is the estimated trend of how average turnout among the treated states would have developed had they not implemented voter ID-requirements. The vertical line at zero

represents the time of intervention. As this varies among treated states, their individual outcome graphs are synchronized around this point. Because some states received treatment later than others and thus have fewer post-intervention observations, their time-series end earlier – note how this means the average treated outcome represents fewer states the further it moves post-intervention.

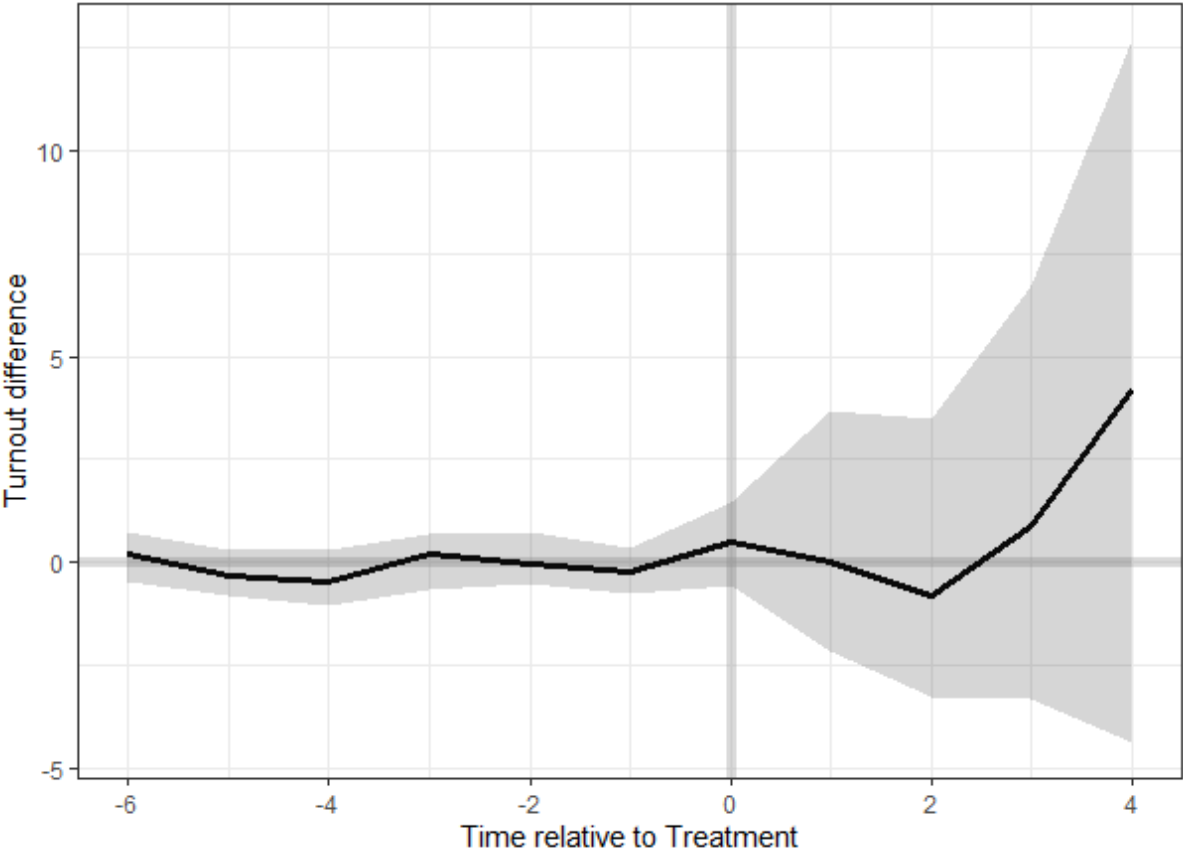
Figure 6.2 Treated and counterfactual average turnout over time



Observe first that the synthetic control closely matches the actual trend in turnout over the pre-treatment period, which increases our confidence in its validity as a post-treatment counterfactual. Post-intervention, there initially appears to be little difference between actual and counterfactual turnout until the fourth election, when average turnout is slightly *higher* than the synthetic control indicates it would have been without ID-requirements in effect. This suggests a small positive effect. Note, however, that this is only based on data from the two states that have been treated for longest, Indiana and Georgia.

Figure 6.3 is a gap plot, showing the year-by-year difference in turnout between treated states and the synthetic control. Essentially, this is the estimated effect of the intervention on turnout – the reader will recall from chapter 4 that in the potential outcomes framework of causality the causal effect of a treatment at time i is calculated as $Y_{i,t}(1) - Y_{i,t}(0)$. Figure 6.3 also shows the 95% confidence interval associated with the estimated effect, represented by the grey band. Again, the good pre-intervention fit is demonstrated by an almost flat trend centered on zero. Post-intervention, however, the effect goes from slightly negative at T+2 to positive at T+3 and T+4. Again, this indicates that turnout was on average higher following the implementation of strict photo ID-requirements, relative to what it would otherwise have been. However, the confidence intervals are large and never exclude zero, meaning we cannot rule out that turnout did not change at all and that the true effect is null.

Figure 6.3 Treated-counterfactual turnout difference, with 95% confidence interval



Turning now to the hard figures, table 6.2 displays the average treatment effect by period relative to the time of treatment. In the seven elections prior to implementation of ID-requirements, we observe no significant effect on mean turnout. This is to be expected, and suggests that the MCM was able to accurately predict treated turnout in the pre-treatment

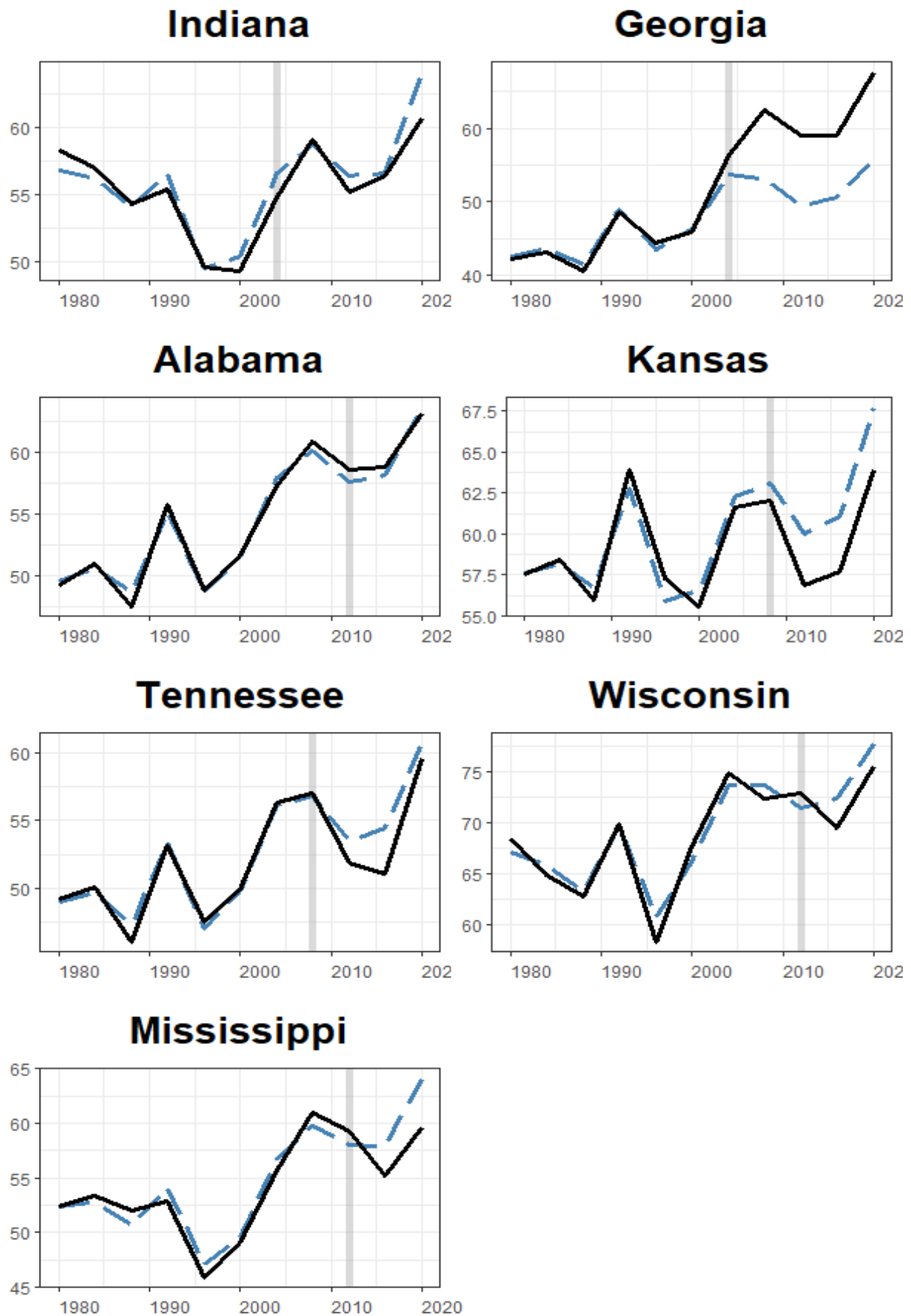
period. The estimated effect hovers around zero for the first three elections post-treatment, but in the fourth election turnout is more than four percentage points higher than the synthetic counterfactual. However, as observed in figure 6.3, standard errors are consistently very large and the effect is never statistically significant. The resulting change in turnout averaged over all periods is 0.32 percentage points, with an associated p-value of 0.872. Based on the 95% confidence intervals, we can be reasonably certain that strict photographic voter ID-laws have not lowered average overall turnout rates by more than 2.9 percentage points, nor increased turnout by more than 5 points. Given this, we cannot reject the null hypothesis of no effect.

Table 6.2 *Estimated treatment effect by period*

Time relative to treatment	ATT	Standard error	CI. Lower	CI. Upper	P-value	N (treated)
-6	0.19	0.32	-0.47	0.76	0.648	0
-5	-0.31	0.29	-0.79	0.31	0.402	0
-4	-0.44	0.36	-1.06	0.30	0.260	0
-3	0.22	0.33	-0.66	0.71	0.666	0
-2	-0.02	0.29	-0.53	0.72	0.954	0
-1	-0.20	0.54	-0.75	0.38	0.530	0
0	0.52	0.51	-0.58	1.46	0.402	0
1	0.01	1.51	-2.15	3.70	0.842	7
2	-0.79	1.77	-3.25	3.50	0.658	7
3	-0.87	2.61	-3.29	6.71	0.790	4
4	4.19	5.86	-4.38	12.71	0.682	2
ATT average		Standard error	CI. Lower	CI. Upper	P-value	
0.32		2.16	-2.86	5.03	0.872	

The largest estimate is found for the fourth post-treatment period, an estimate that – as table 6.6 shows – is based only on data from the two states which have had strict photo ID-laws in effect the longest: Indiana and Georgia. It may therefore be instructive to plot the results for each treated state separately, which is done in figure 6.4. For Indiana, the counterfactual trend displays minor deviations from actual turnout both pre- and post-treatment. Georgia, on the other hand, exhibits a reasonably good pre-treatment fit of the synthetic control, which in turn suggests a large positive effect in the post-treatment period – a more than 10 percentage point increase in turnout relative to the counterfactual scenario. It appears, then, that the positive (though insignificant) effect discovered is largely driven by Georgia.

Figure 6.4 *Treated and counterfactual turnout by state*



Among the remaining treated states, both Alabama and Tennessee have synthetic controls that match well for most of the pre-treatment period. The former shows little meaningful deviation post-treatment, while the latter exhibits a clear *negative* effect in the first and second elections after intervention. Counterfactual Kansas stands out with a similarly large negative effect estimate, but the somewhat worse pre-treatment fit lowers our confidence in its predictive power. In summary, the ATT seems to hide a certain variability in the state-level effects, ranging from strongly positive to moderately negative. This variability possibly also explains why the confidence intervals for the estimated *average* effect are so wide. Note that because standard errors are not generated for the state-level estimates the uncertainty surrounding them is unknown. Strong inferences based on these findings is therefore inappropriate.

6.1.2 Robustness checks

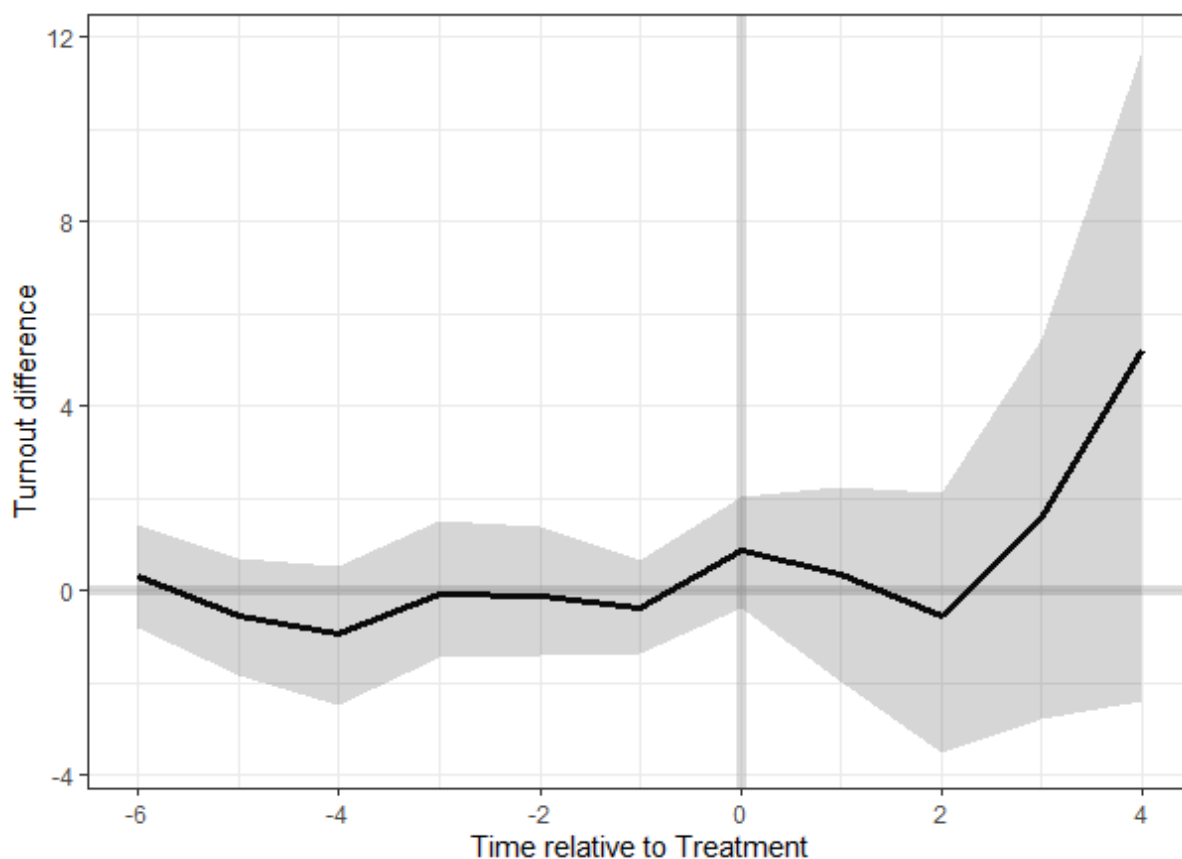
In chapter 4, I explained that the MCM is likely the best estimator given the characteristics of the data. As a robustness check, however, I also estimate the treatment effect using the generalized synthetic control method. Recall that this estimator generates counterfactual outcomes using a latent factor approach with an interactive fixed-effects model. The cross-validation procedure chooses the model with a single latent factor as generating the lowest pre-treatment MSPE. The factor and factor loading are shown in the Appendix.

The results, presented in figure 6.5 and table 6.3, largely mirror those of the main analysis. The estimated effect is initially small before turning positive in the last two periods – again, however, it never reaches statistical significance. The 95% confidence interval for the ATT for all periods is somewhat narrower, between -2.36 and 3.32 percentage point change in turnout. Note, however, that the pre-treatment MSPE is larger here as evidenced by the deviation in pre-treatment fit in figure 6.5, meaning that the GSC did a worse job of approximating the treated states than did the MCM and thus likely constitutes an inferior counterfactual substitute.

Table 6.3 *Estimated treatment effect (generalized synthetic control)*

ATT average	Standard error	CI. Lower	CI. Upper	P-value
0.75	1.43	-2.36	3.32	0.568

Figure 6.5 *Treated-counterfactual turnout difference (generalized synthetic control)*



The independent variable of the analysis is whether a state has implemented strict photographic voter ID-requirements, operationalized as a dichotomous treatment. However, recall from chapter 2 that many states have adopted less restrictive types of voter ID-laws, for instance without requiring photographic ID (non-photo) or with the possibility of making exemptions on Election Day (non-strict). Including these states in the donor pool when generating the synthetic control could bias the estimated effect if these laws also affect turnout. I therefore conduct an analysis where only states without any kind of ID-requirement are included in the control group. Figure 6.6 shows which states remain in the analysis. Figure 6.7 and table 6.4 present the results, which once again closely match those of the main analysis.²³ The effect remains small and insignificant. The estimated ATT for all periods is now slightly negative due to a dip in turnout in period T+2, though we still observe an increase in turnout relative to the synthetic control in the fourth period post-intervention.

²³ The results presented are from the MCM; estimates are substantively unchanged when using the GSC.

Figure 6.6 Treatment status by state and election (restricted controls)

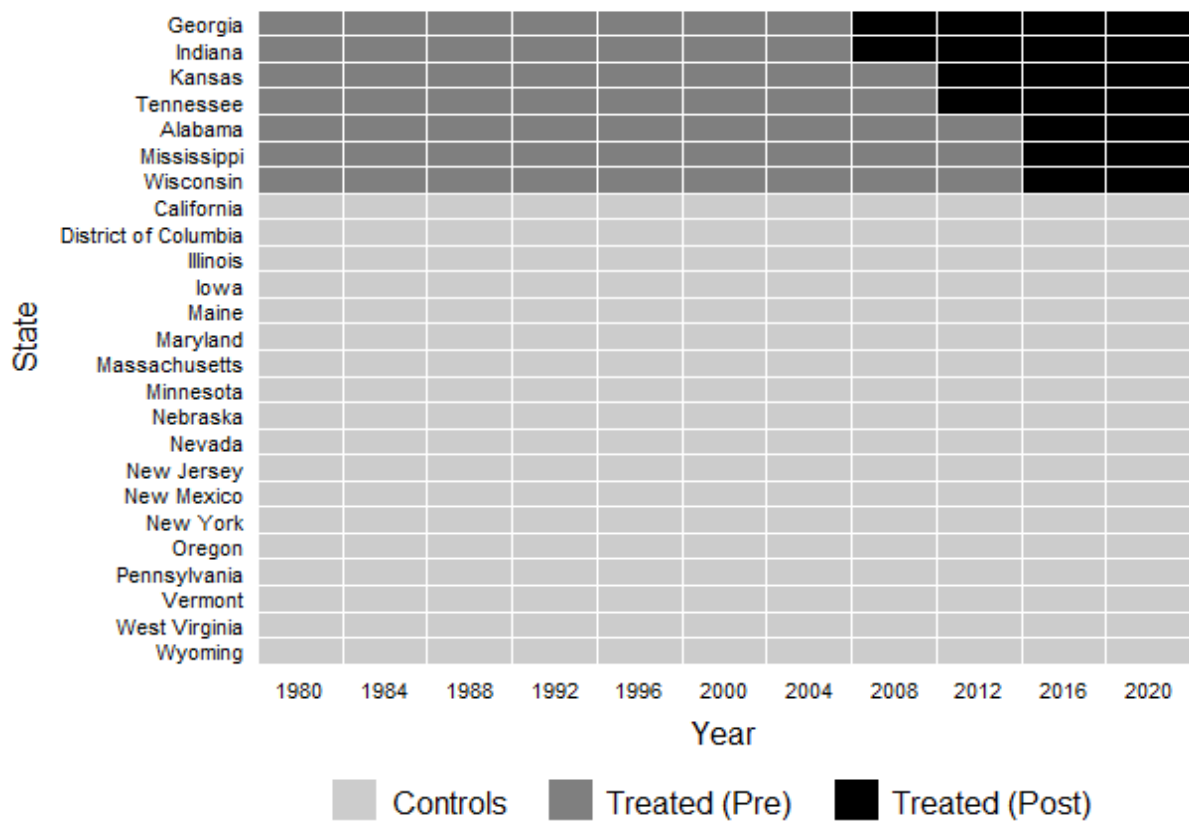


Figure 6.7 Treated-counterfactual turnout difference (restricted controls)

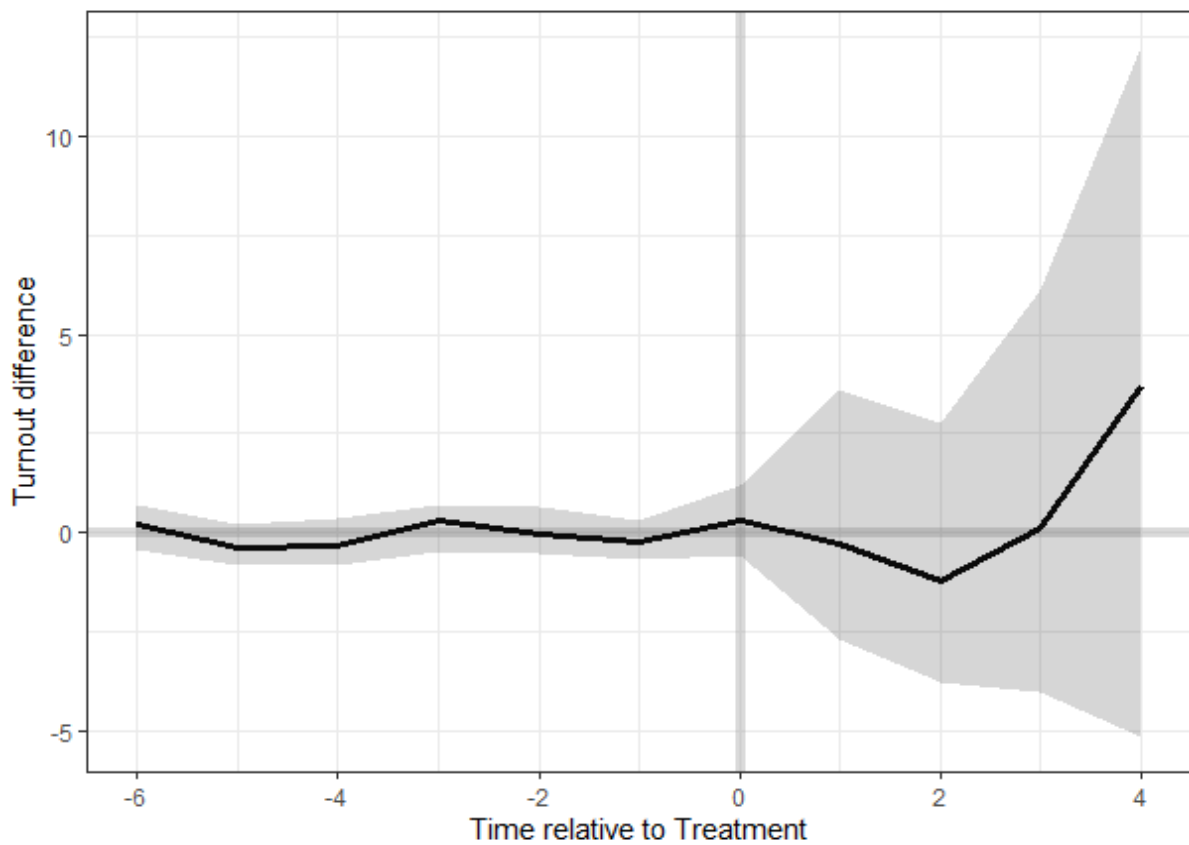


Table 6.4 *Estimated treatment effect (restricted controls)*

ATT average	Standard error	CI. Lower	CI. Upper	P-value
-0.13	2.19	-3.62	4.43	0.866

6.1.3 Alternative analysis – all election years

In chapter 3, I explained how the measure of turnout used is somewhat problematic for midterm election years, as it could, depending on the state and year, refer to a gubernatorial, senatorial, or congressional race. Because turnout is likely higher when a more important seat is to be filled, reported turnout rates in a single midterm election year may be incomparable across states. For this reason, I focus the main analysis on presidential elections only. However, given the uncertainty surrounding the estimated null finding, it appears useful to leverage the full extent of the data available to further validate the inference. I therefore estimate the effect of strict photo ID-requirements using data from both presidential and midterm elections between 1980 and 2020. This has the advantage of drastically increasing the number of observations over time, from $T = 11$ to $T = 20$.²⁴ Crucially, because each midterm election occurs between two presidential election years, additional observations are added both pre-intervention (which yields more data with which to generate and validate the synthetic control) and post-intervention (which yields more data points on which to assess the effect of the treatment). Louisiana is removed from this analysis as it lacks data on turnout for the 1982 midterm election.

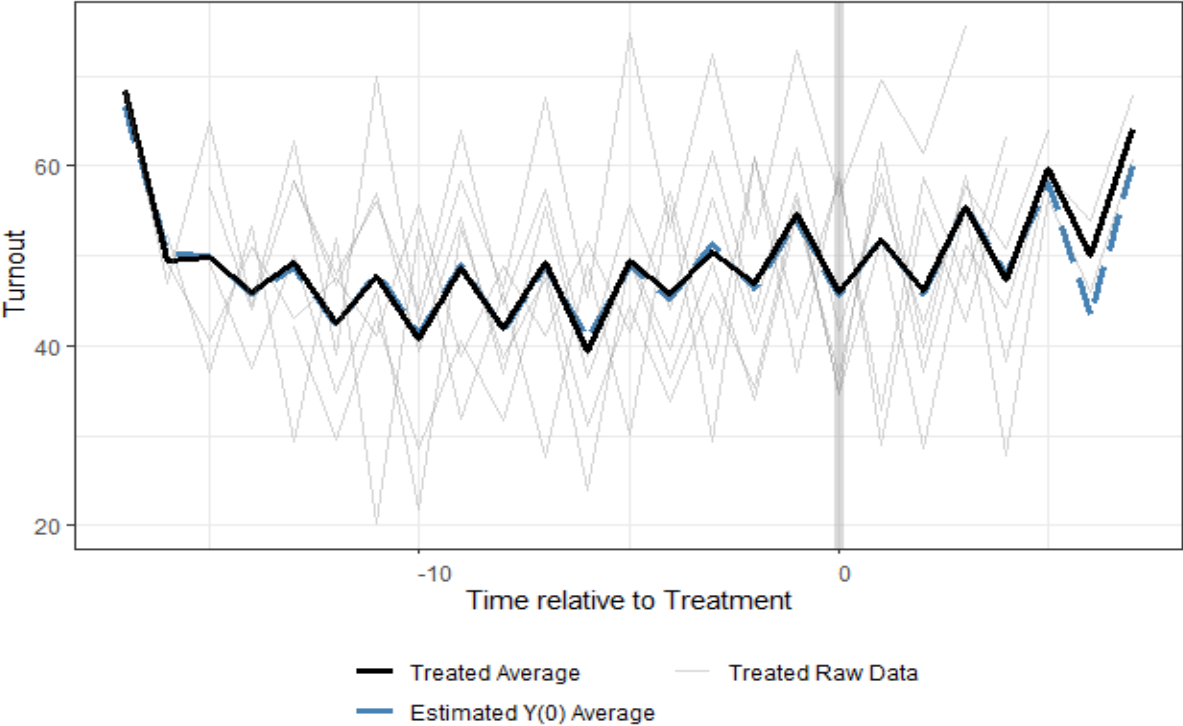
Note, however, that we are introducing additional uncertainty into the model due to the lower degree of comparability of turnout rates across states in midterm election years. This uncertainty is not necessarily captured in the standard errors generated by the MCM and is therefore not immediately apparent from the results.

Figure 6.8 graphs the counterfactual turnout relative to the observed turnout trends, along with the raw data for each treated state. Observe first the oscillation resulting from turnout being consistently higher in presidential elections than in midterms. The MCM captures this variation

²⁴ 1980, 1982, 1984, 1986, 1988, 1990, 1992, 1994, 1996, 1998, 2000, 2002, 2004, 2006, 2008, 2010, 2012, 2014, 2016, 2018, and 2020.

well, and for the most part matches pre-treatment turnout closely. Post-treatment, the synthetic control once again shows no initial effect of the intervention until the end of the time-series, where turnout is once again higher than in the counterfactual scenario, indicating a positive effect. The gap plot in figure 6.9 explicates this deviation. However, it also highlights that the effect does not reach statistical significance at any point. Note also that the pre-treatment fit of the counterfactual falters at some points and is overall noticeably worse than that of the main model (see figure 6.3). Possibly, this is due to the lower comparability of the turnout measure in midterm election-years, which introduces unpredictable variation that makes it difficult for the MCM to generate an accurate synthetic control. Either way, this underscores how our confidence in these inferences are reduced relative to the main analysis. Nevertheless, the conclusion remains the same: state-level turnout rates did not significantly change due to the implementation of voter ID-laws. Averaged over all periods, we can bound the effect between a 1.94 percentage point decrease and a 4.47 point increase, as per table 6.5.²⁵

Figure 6.8 *Treated and counterfactual average turnout over time (all elections)*



²⁵ The null finding remains when applying the robustness checks from the main analysis, with one exception: When estimating with the GSC a significant positive effect is found for the 6th election post-intervention; however, this counterfactual exhibits particularly poor pre-treatment fit. See the Appendix for more.

Figure 6.9 *Treated-counterfactual turnout difference (all elections)*

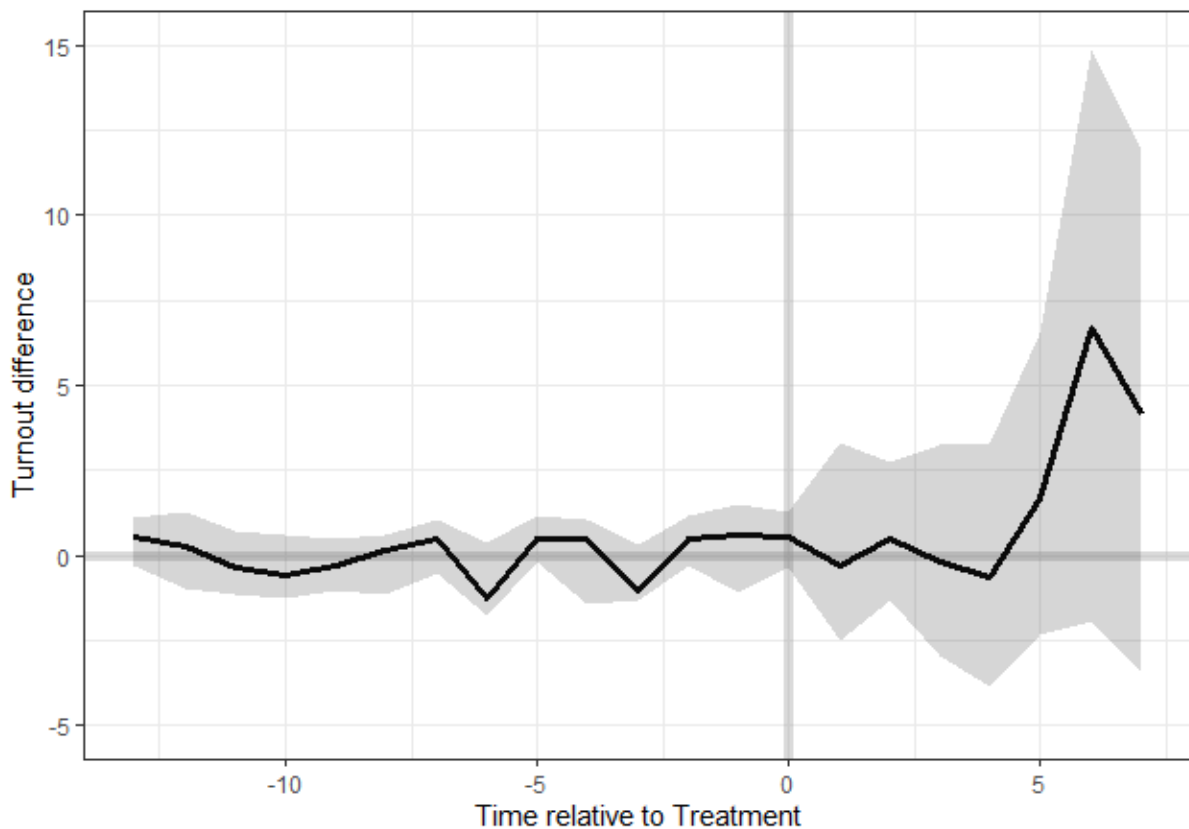


Table 6.5 *Estimated treatment effect (all elections)*

ATT average	Standard error	CI. Lower	CI. Upper	P-value
0.71	1.83	-1.94	4.47	0.68

6.2 Do voter ID-requirements disproportionately affect minorities?

Results from the difference-in-differences analysis

The null finding in terms of an overall effect of voter ID-laws on turnout does not preclude an effect at the group-level, as per hypothesis H2, which could be obscured in the main analysis focused on aggregate turnout rates if the group in question constitutes a relatively small part of the total population. In particular, ethnic and racial minorities are thought to be particularly vulnerable to stringent voter ID-requirements. As discussed in the previous chapter, the lack of official data on turnout among different ethnic groups makes this claim difficult to assess using a synthetic control approach. I therefore employ a difference-in-differences design to

investigate whether strict photographic ID-requirements have negatively affected minority turnout, leveraging survey data from the Current Population Survey. Comparing the change in turnout from 2004 to 2016 among people living in treated states with the corresponding change among those living in the control states yields an unbiased estimate of the treatment-effect, conditional on the parallel trends assumption. Comparing the estimate for minority voters with that for non-minorities allows me to assess whether the effect varies across these groups.

Table 6.6 displays the results of the DID-analysis. I first estimate the overall effect for all respondents, as an equivalent to the main analysis in the previous section. The DID yields an estimate very similar to that of the synthetic control: a small, positive, and insignificant effect. On average, turnout increased by 1.05 percentage points more in the treatment group than in the control group; however, the 95% confidence interval is wide, ranging between -2.21 and 4.31, and the effect is not statistically significant. The overall null finding remains.

Table 6.6 *Difference-in-differences of turnout between treated and control, by group*

	All	Minority	Non-minority
ATT	1.05	0.23	0.56
Standard error	1.66	3.40	1.42
CI. Lower	-2.21	-6.43	-2.22
CI. Upper	4.31	6.89	3.35
P-value	0.53	0.95	0.69
N	159 737	34 385	125 352

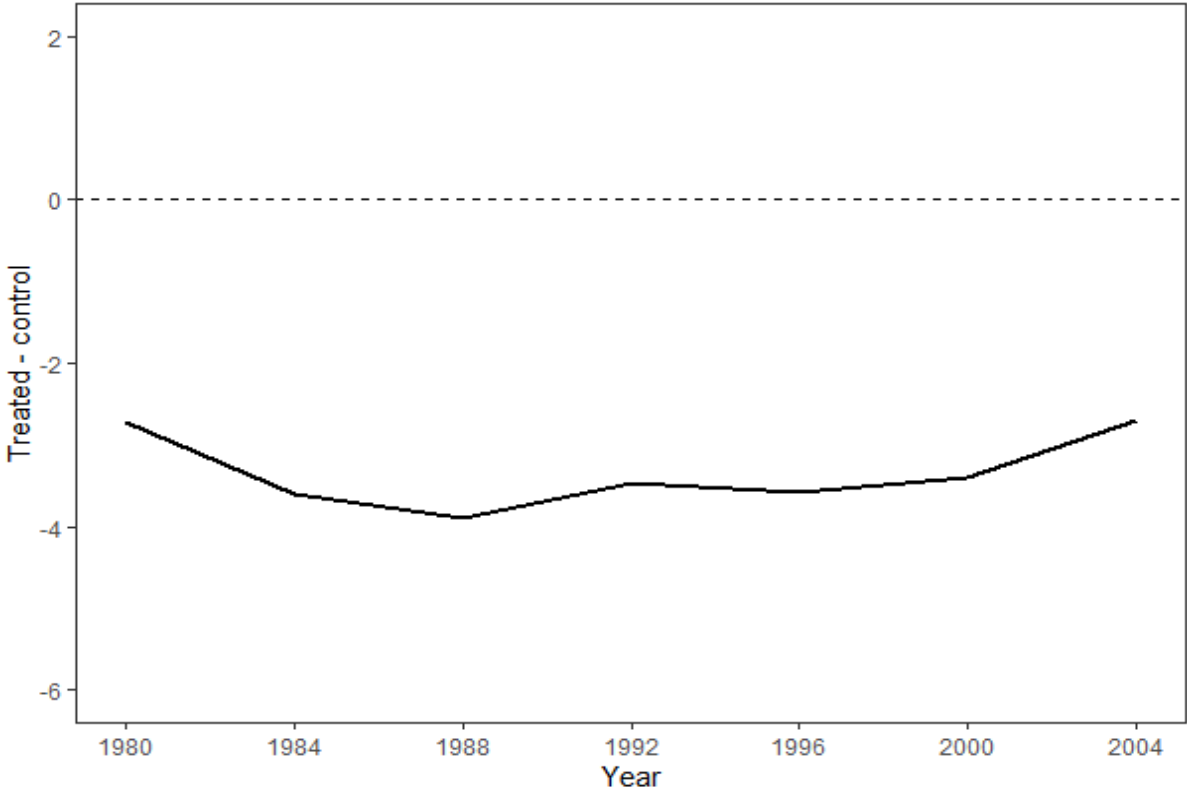
Turning now to the assumed differential impact, H2 predicts that minorities will be disproportionately negatively affected. However, when restricting the analysis to only minority respondents, the estimated effect among the treated remains positive at 0.23. This positive effect on turnout is slightly smaller than the estimate for non-minority respondents (0.56), but the p-values leave both estimates far from statistical significance. This is illustrated by the 95% confidence intervals, which ranges from 6.43 percentage point decrease to a 6.89 point increase in turnout among treated minorities, relative to the change in the control. For non-minorities the estimate ranges from -2.22 to 3.35. The higher degree of uncertainty associated with the minority effect-estimate is likely due in part to the considerably fewer respondents in this group. Nevertheless, the overall conclusion is the same for both minorities and non-minorities in that we do not observe a significant difference in the change in turnout of those living in states with

strict photo ID-laws compared to those living in control states. The difference in differences is statistically indistinguishable from zero, and thus no effect of the treatment is found.

6.2.1 Assumption check: parallel trends

The difference-in-differences approach relies on the parallel trends assumption: that trends in the outcome variable would move in tandem in the treatment and control groups had the intervention not occurred. In other words, we allow for differences in turnout between the two groups so long as they are constant over time. This cannot be tested directly for the post-treatment period, but a second-best alternative is to see whether to assumption holds prior to intervention. To do this, figure 6.10 plots the difference in average turnout rates between respondents from treatment and control states in presidential elections from 1980 to 2004 (Flood et al. 2020). If the trends were exactly parallel, we should observe a flat, horizontal line, though it need not be centered on zero. While not flat, the graph does not indicate any major deviations; pre-treatment turnout was more or less stably a few percentage points lower in the treatment group than in the control. The results do not suggest that the parallel trends assumption would be substantially violated in the post-treatment period in the absence of intervention.

Figure 6.10 *Treated-control difference in turnout, 1980-2004*



6.2.2 Robustness checks - DID

Certain choices made regarding the calculation of the difference-in-differences warrant consideration as robustness checks on the results of the main analysis. As with the synthetic control, I also estimate the DID using only data from states without any ID-requirements in the control group. The first model in Table 6.7 displays the results, which are similar to those of the original analysis in all respects but one: the effect estimate among minorities is now negative at -0.72, the only evidence consistent with hypothesis H2 so far. However, though p-values are somewhat smaller across the board, no estimate comes close to conventional levels of significance – turnout did not change more or less among respondents in strict photo ID-states than it did among those in no-ID states. We thus cannot rule out the null hypothesis of no effect for either group. Note the reduction in N for these models due to the exclusion of respondents from states with “weak” ID-laws.

Table 6.7 *Difference-in-differences (alternative specifications)*

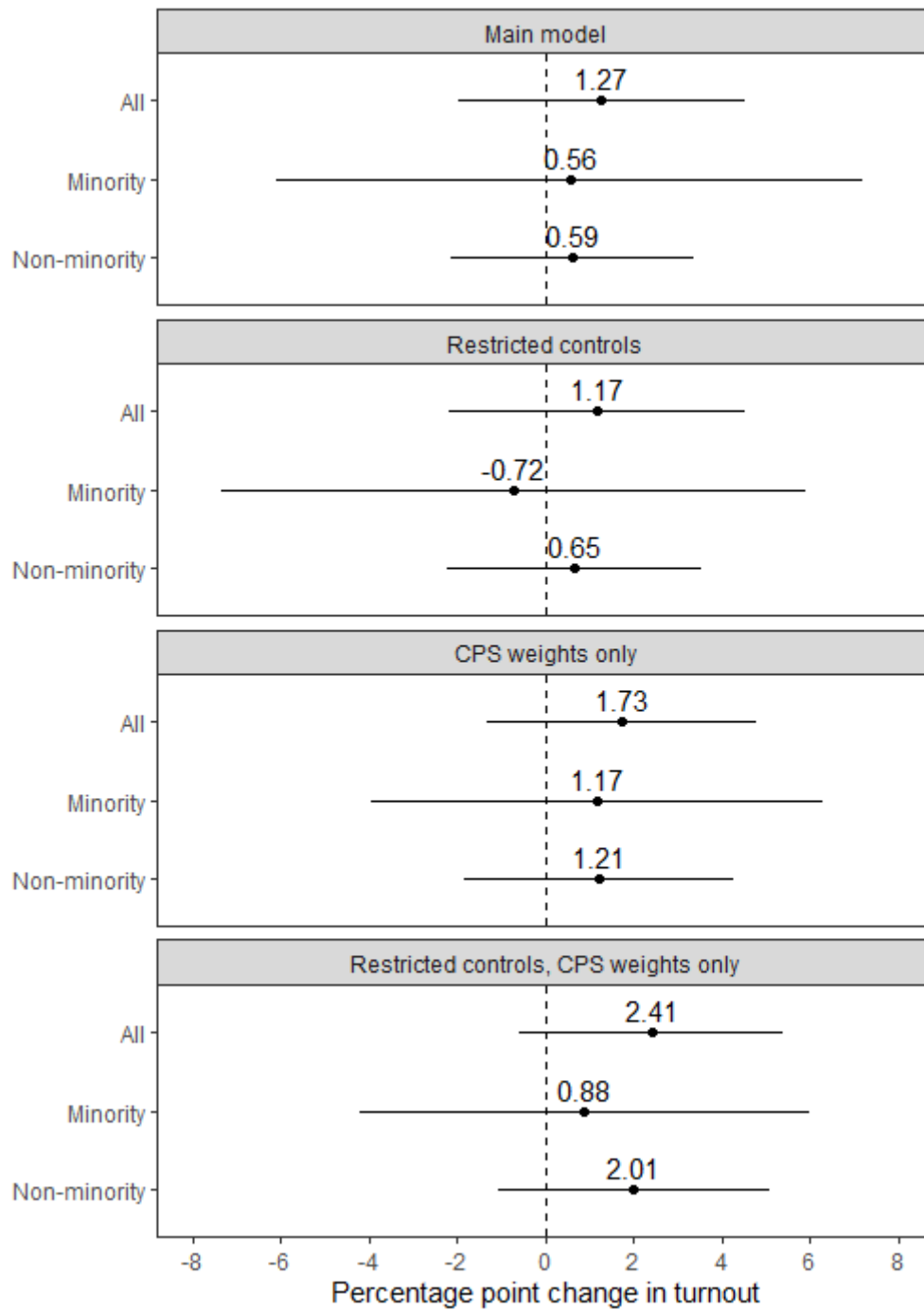
		All	Minority	Non-minority
Restricted controls	ATT	1.17	-0.72	0.65
	Standard error	1.72	3.39	1.48
	CI. Lower	-2.20	-7.35	-2.25
	CI. Upper	4.54	5.92	3.55
	P-value	0.50	0.83	0.66
	N	90 346	20 812	69 534
CPS weights only	ATT	1.65	1.04	1.26
	Standard error	1.58	2.65	1.58
	CI. Lower	-1.44	-4.14	-1.84
	CI. Upper	4.73	6.23	4.36
	P-value	0.30	0.69	0.43
	N	159 737	34 385	125 352
Restricted controls, CPS weights only	ATT	2.41	0.88	2.01
	Standard error	1.53	2.61	1.58
	CI. Lower	-0.59	-4.23	-1.09
	CI. Upper	5.41	5.00	5.10
	P-value	0.12	0.73	0.20
	N	90 346	20 812	69 534

The main DID model combines CPS sampling weights with those suggested by Hur & Achen (2013) to contend with the overreport bias of the CPS. Recognizing that this is not a universal practice, as a second robustness check I also estimate the effect using only the CPS's own sampling weights. The results, presented in the second model of table 6.7, mirror the main model in that the both the overall and group-level effects are slight, positive, and – albeit less so than the main estimates – statistically insignificant.

The final specification combines the two first robustness checks – presented in the last model of table 6.7. The direction of the effect estimates is again unchanged from the main model. However, this time the overall effect and the effect among non-minority respondents becomes both larger and closer to significance – the former is almost significant at the 10 percent-level. This is the closest evidence yet of a significant effect of ID-requirements, and it is positive rather than negative; on average, turnout increased by 2.41 percentage points more in the treated group than among untreated respondents. There is still no evidence of a negative effect among minority respondents, though the positive point estimate for the effect is noticeably smaller than that for non-minorities (0.88 versus 2.01), which arguably constitutes a negative impact among minorities, though in relative rather than absolute terms. However, this estimate remains far from significant. Also, keep in mind that since these estimates only use CPS weights they are arguably more vulnerable to overreport bias, and thus this finding is only tentative.

In sum, then, we observe no significant effect of strict photographic ID-requirements on minority turnout, negative or otherwise. We also cannot conclude that minorities were affected more strongly than non-minority voters. These results hold when excluding respondents subject to other forms of voter ID-laws from the analysis and when using a more conventional weighting scheme. For ease of comparison, figure 6.11 visualizes the effect estimates from the DID-analyses and their respective 95% confidence intervals – note how they all include zero, represented by the vertical line. However, the wide confidence intervals also urge caution in ruling out the possibility of a group-level effect entirely, as the less-than-ideal data available for this secondary analysis attach non-negligible uncertainty to the results.

Figure 6.11 *Group-level effect of strict photo ID-laws, with 95% confidence intervals*



7 Discussion and conclusion

In this chapter, I discuss the findings. I begin by reviewing the results, summarizing their implications for the hypotheses, and answering the research question. I then deliberate on why the analysis finds no effect of strict photographic voter ID-laws on turnout, before considering the implications of this null finding relative to previous studies, theories of voting, and the debate on voter ID, and offer some remarks on policy. In an effort of transparency, I also point out some drawbacks to the analysis. To conclude, I highlight some avenues for future research into the relationship between voter ID-laws and turnout.

7.1 Evaluating the hypotheses and answering the research question

This thesis has investigated the contentious issue of voter ID-laws in the United States, using the following research question:

Have voter ID-laws led to lower turnout?

Using a synthetic control approach, I find no evidence that average turnout in states with strict photographic voter ID-requirements was significantly different than it would have been in a counterfactual scenario in which no such requirement was implemented. The analysis does indicate some variation at the state-level, but the uncertainty associated with these estimates is unknown. Thus:

The analysis found no support for **H1**: *Implementation of strict photographic voter ID-requirements lowers a state's turnout rate, relative to what it would otherwise be.*

Using a difference-in-differences design, I find that the change in turnout from 2004 to 2016 was not significantly different among people living in states that implemented strict photo ID-requirements over this period than among people living in states without such requirements, for both minority and non-minority respondents. Thus:

The analysis found no support for **H2**: *Strict photographic voter ID-requirements lower minority turnout more than non-minority turnout.*

Do voter ID-laws, then, lead to lower turnout? My analysis finds no significant evidence that they do, neither overall nor among the minority voters thought particularly vulnerable. If an effect on aggregate turnout exists, this study suggests that it is likely to be small in magnitude, and possibly even positive, so that requiring voters to show ID in fact leads to a higher turnout rate. The results also indicate that some states may be affected differently by ID-requirements than others. Nevertheless, based on the findings, I cannot answer the research question in the affirmative.

7.2 Why do we not find an effect of voter ID-requirements on turnout?

The main takeaway from the empirical examination is the overall null finding. The results from the synthetic control analysis do not indicate that strict photographic voter ID-requirements significantly lowered average turnout in the states that have implemented them, while the difference-in-differences analysis finds no evidence of a differential impact on minority voters. Before discussing the findings further, it is worth noting an important limitation of the methods used. Neither the synthetic control method nor difference-in-differences designs tell us anything about the causal mechanisms involved, and thus we cannot infer from the analysis exactly what processes have or have not occurred to produce the results observed. With this in mind, I attempt some deliberation on how to explain the findings.

The results of the analysis run counter to the prediction of both opponents of voter ID and rational choice theory that voter ID-laws will lower turnout. This is explainable in one of two ways. First, it is possible that there is in fact a net effect on turnout – either negative or positive – but that it is too small to be detected by the tests employed here. Grimmer & Yoder (2021, 2, 14) argue that the true effect of voter ID-laws is likely so small that any study focusing on aggregate turnout will lack the statistical power necessary to capture it, due largely to the small size of the portion of the electorate affected by ID-requirements. Recall that a negative effect on turnout requires the existence of people who would otherwise vote, lack valid ID, and are

unable or unwilling to acquire it (Highton 2017, 156). In chapter 3, I demonstrated that these are all plausible assumptions, but it is possible that the proportion of the electorate exhibiting all three prerequisites is small enough that a decline in turnout among this group is insufficient to noticeably reduce overall turnout rates.

However, the obvious explanation of the null finding is that there is no effect to be found: ID-requirements simply do not affect turnout negatively. Given the expectations derived from the literature on the calculus of voting, this begs the question of why. Several answers present themselves.

One possibility is that the premise behind the hypothesized negative effect itself is faulty. As demonstrated in chapter 3, this prediction implicitly assumes a rational view of voting. While on one hand the supposed marginality of the act of voting is what opens the door to slight cost increases like voter ID-requirements significantly affecting turnout, Aldrich (1993) suggests that voting could be such a low cost/low benefit action that it is simply not worth thinking rationally about. This parallels the prediction of non-rational theories of voting like that of Rolfe (2012), wherein one would not expect turnout to decrease significantly as a result of an increase in the costs of voting, because costs are not an important consideration in the individual's decision to vote or not. Voter ID-laws, then, do not affect voting because voting is not a rational action – at least not in the conventional sense.

A second explanation – more compatible with rational choice theory – is that voters do think rationally about voting, but that voter ID-requirements do not really alter the costs of voting – at least not by much. In other words, the supposed treatment under study is not as strong an intervention as previously assumed. Certainly, in a historical perspective, requiring voters to show ID pales in comparison to some tools of disenfranchisement previously used to restrict the American suffrage (Keyssar 2000). Intuitively, it makes sense that small changes in costs should only be expected to correspond with small changes in turnout. Perhaps having to show ID is not sufficiently costly to impact voters' decision-making.

Alternatively, it is possible that voter ID-laws do, *ceteris paribus*, cause fewer people to vote, but that they also activate processes that operate in the other direction, so that the net effect on turnout is zero. In chapter 3, I outlined two such counteracting forces. First, political elites thought to lose out if turnout declines could engage in counter-mobilization efforts to offset any perceived disadvantage resulting from ID-requirements (Highton 2017). Second, the contentious nature of voter ID-laws may cause people to turn out in protest of the new law. (Valentino & Neuner 2017). Certainly, the partisan divide on the issue of voter ID suggests that Democrats view them as a threat to their electoral interests, while at the individual-level, several studies have found that perceived attempts at disenfranchisement can – somewhat ironically – make people more likely to participate (Valentino & Neuner 2017; Biggers & Smith 2020). Finally, there is the main argument for voter ID-requirements in the first place: voter fraud. Regardless of whether requiring voters to show ID actually reduces instances of fraud, the perception that they do may inspire greater confidence in the integrity of elections among some voters, which in turn could motivate them to vote where they otherwise would have abstained.

Meanwhile, Vercellotti & Andersen (2009) predict that the hypothesized negative effect should be strongest immediately post-treatment – I added that this could apply equally to a mobilizing effect as well. It seems implausible that backlash and elite mobilization efforts should be delayed by multiple elections. Therefore, the finding that the positive-though-insignificant effect estimate grows larger over time could indicate that the demobilizing effect deteriorates at a faster rate than the counteracting forces, or perhaps that the latter are stable over time.

The variation in the effect estimates for the individual treated states presents another puzzle. Among the three states with the best pre-treatment fit of the counterfactual, we observe three very different trends post-treatment: a positive effect in Georgia, a negative effect in Tennessee, and no effect in Alabama. The lack of uncertainty estimates for the state-level effects mean they must be interpreted with caution. However, the results suggest that a closer comparison of these three states could shed some light on the apparent variability in how voter ID-laws affect turnout. The positive effect estimate obtained for Georgia suggests that some mobilizing effect could be in play here. However, the question then turns to why this occurs in Georgia and not,

for instance, in Tennessee – or alternatively, why it is strong enough to reverse a negative impact of voter ID-laws in the former case but not in the latter.

A final possibility concerns the hypothesized differential impact, through which a group-level effect could lie hidden beneath the null finding of the primary analysis. However, the results of the main DID analysis do not indicate any significant decrease in turnout among minority voters either, neither in absolute terms nor relative to non-minorities. This second test thus further compounds the overall null finding. The possible explanations discussed above all apply here as well. In fact, if minority voters perceive themselves to be targeted by the new requirements, there is reason to suspect that an outrage-effect could be particularly strong for this group. Though this secondary analysis is arguably less stringent than the main analysis, the evidence does not support claims that minorities are disproportionately negatively affected.

In summary, there are two main possible explanations for the null finding. First, it may be that there is indeed a net effect on turnout, but that it is so small as to be undetectable by the research design employed. A more intuitive interpretation is that voter ID-laws indeed have no aggregate effect on turnout, either because they do not impact turnout at all, or because some opposing factor negates any negative effect. The research design employed here does not allow us to adjudicate between these possibilities with certainty. For a quantitative approach to investigating such intervening factors within the counterfactual framework of causality, see the work of Imai et al. (2010) on causal mediation analysis. Though subject to the assumption of sequential ignorability, this approach allows both parametric and nonparametric estimation of causal effects in a wide range of applications, and thus could prove a useful addition to the study of voter ID and turnout.

7.3 Implications

In this section, I consider the implications of the findings, first relative to those of previous studies and as they pertain to theoretical explanations of voting, then regarding the issue of voter ID itself and the surrounding controversy. Lastly, I offer some suggestions for policy.

7.3.1 Implications for research

Relative to previous research on the relationship between voter ID-requirements and turnout, the findings of this study are closest to those of Erikson & Minnite (2009), Heller et al. (2019), and Cantoni & Pons (2021), who also find no significant effect. However, because of the degree of uncertainty surrounding the estimates, the results of my analysis are not entirely incompatible with the majority of studies suggesting a small, negative effect (Dropp 2013; GAO 2014; Pryor et al. 2019; Kuk et al. 2020; Grimmer & Yoder 2021). Most dissimilar to the present findings is the study of Hajnal et al. (2017), which finds a large differential impact among minorities – up to a double-digit difference in turnout. In addition to these studies, of particular note is that of Fraga (2018), which, while finding evidence of a negative effect in some cases, indicates that ID-laws may have a *positive* effect on turnout in others. Fraga’s results offer corroborating evidence on the possibility of a mobilizing effect of ID-requirements. Also noteworthy is the study of Cantoni & Pons (2021), which obtains point estimates for the effect that, while never significant, become increasingly positive over time, mirroring the results of the present synthetic control analysis.

Why, then, do the majority of previous studies find evidence of an effect where mine indicates none exists? I argue that the answer is methodological. While some studies make use of a DID design to approximate the counterfactual trend in turnout over time (GAO 2014, Kuk et al. 2020; Fraga 2018), few explicitly employ a rigid counterfactual framework in their research design. This in turn weakens the case for causality in their findings – a criticism which is possibly even more applicable to the studies relying on a regression-based design (Hajnal et al. 2017; Pryor et al. 2019), as they are particularly vulnerable to omitted-variable bias. Additionally, there is variation in the exact classification and types of ID-laws considered, with several studies also examining only a subset of states and elections, often at a time when strict photographic requirements were less than prolific. This could also explain the variability of previous findings. My research design seeks to correct for these potential sources of error; to the degree to which it is successful, it adds meaningfully to existing knowledge.

Recall that the main contribution of this paper is methodological. Operating within the tradition of causal empiricism, it focuses on design-based causal inference using a counterfactual framework. This provides certain advantages relative to other statistical analyses. In a causal

empiricist critique, Samii (2016) argues that conventional regression analyses often are compromised regarding both external and internal validity. He labels the issues respectively as pseudo-generality, by which results based almost entirely on data from a subset of units are taken as valid for the universe of cases, and pseudo-facts, where the results obtained are particular to the exact model specification used (Samii 2016, 943-949). The synthetic control approach as applied here addresses both these points. First, by being explicit about exactly what causal quantity is estimated, it avoids pseudo-generality.²⁶ Second, it constitutes a rigorous approach to causal inference that reduces researcher discretion through automation of key modelling decisions, thus arguably reducing the danger of misspecification relative to a regression-based design.²⁷ This should enhance internal validity and help insure against the generation of mere pseudo-facts. The difference-in-difference design similarly offers a simple, yet powerful way to approximate the experimental ideal when data are observational.

The issue of model dependence – which threatens the internal validity of the causal inferences made in a study – warrants additional attention, as it could help explain the variable findings of the literature. Put simply, model dependence occurs when the causal estimates of a study are sensitive to particular choices regarding model specification, such as control variables, functional form, et cetera, so that changing these factors alters the results. (Ho et al. 2007; King & Zeng 2007). In these cases, significant results are little more than “demonstrations that it is *possible* to find a specification that fits the author’s favorite hypothesis [emphasis in original]” (Ho et al. 2007, 199). This relates to another phenomenon, publication bias, whereby significant findings are favored in the review processes at academic journals and, as a result, are overrepresented in the body of published research (Esarey & Wu 2016). In reviewing the substantive consequences of this skew, Esarey & Wu (2016, 2) conclude that “published estimates of relationships in political science are on average substantially larger than their true value.” In short, model dependency and publication bias could combine to produce exaggerated estimates within a field of study.

²⁶ See the discussion on the average treatment effect (ATE) versus the average treatment effect on the treated (ATT) in chapter 4.

²⁷ Recall from chapter 4 that a key reason for not incorporating additional predictor variables when generating the synthetic counterfactual was eliminating specification searches.

Though it is not certain that research on the effect of voter ID suffers from this ailment, it is, if nothing else, noteworthy that this study, which emphasizes design-based inference and employs a novel methodology precisely to enhance the validity of causal estimates and reduce the possibility of type I errors, fails to reject the null hypothesis where previous studies have done so. Moreover, the most recent and comprehensive study to date, Cantoni & Pons (2021), which has access to a dataset of exceptional size and richness, displays results that are remarkably similar to the ones obtained here. Possibly, the plurality consensus in the literature on the negative impact of voter ID-laws on turnout is – at least in part – an artefact of the research designs employed rather than ID-requirements themselves.

7.3.2 Implications for theory

The object of this thesis is not to test the explanatory power of competing theories of voting. Nevertheless, the results do allow for reflection on the implications for the study of voting more generally. Because rational choice theory seems to predict that voter ID-requirements should lead to lower turnout, finding that this is not the case is arguably a strike against a rational view of voting. If changes in elements of the calculus of voting (i.e. C) only lead to small changes in turnout, this could suggest that rational choice theory is, at best, only a partial explanation of voting, and that other theories may be more powerful (Blais 2000, 11, 137). Charitably, one may counter that small changes in costs should correspond only to small changes in turnout. If voter ID-laws, then, are only a weak intervention, a rational view of voting may still be appropriate for other, stronger treatments.

Furthermore, the possibility that the null finding is in part due to ID-requirements having an indirect, positive effect allows for a scenario in which costs of voting do have an independent effect on turnout, only that in the case of voter ID this effect is masked. However, this possibility also suggests the inadequacy of the calculus of voting as a complete theory of voting, as both mobilization theory and psychological explanations might better account for these counteracting forces. In a sense, this mirrors another fundamental challenge of the calculus of voting: explaining why, in the face of undeniable costs and a negligible probability of affecting the outcome, supposedly rational actors still vote in large numbers. As mentioned in chapter 3, numerous explanations have been offered, but none agreed upon. While Grofman (1993) have suggested that rational choice theory may do better in predicting *changes* in turnout rather than

absolute levels, the findings of this thesis suggest that even with this more limited scope, a full explanation of turnout requires drawing upon alternative theoretical perspectives as well.

7.3.3 Implications for voter ID

What, then, do the findings of this thesis contribute to our understanding of voter ID? First and foremost, the null finding seemingly gives cause for relief among those who fear that voter ID-laws will reduce public participation in elections. Accordingly, the controversy over voter ID may be a case of much ado about nothing (Hood & Bullock 2012).

Recall that this thesis focuses on what is arguably the strongest form of voter ID-law: strict photographic ID-requirements. This influences the inferences we can draw to the study of voter ID and turnout more generally. In one sense, a null finding here means we are unlikely to find a negative effect of other, weaker types of ID-laws. However, if the explanation for the null finding is that outrage over the laws have led to counter-mobilization, this arguably leaves the door open to a negative effect of lesser forms of ID-laws if they, by virtue of being less invasive, are also less controversial and as a result do not spark the same level of counter-mobilization. It thus appears useful to distinguish between direct and indirect effects: lesser types of ID-requirement could have a weaker direct (negative) effect on turnout, but also a weaker indirect (positive) effect. The net impact would depend on the relative strength of the two in a given case.

There is, however, also reason to remain cautious regarding strict photographic requirements. The strong reactions against voter ID center on their supposed threat to democratic fairness and that the differential impact of ID-requirements will alter election outcomes (in favor of Republicans). The results do not necessarily preclude the possibility of this latter scenario. The confidence intervals obtained indicate that the effect is unlikely to be greater than a few percentage points change in overall turnout. However, even such a small change in turnout may be pivotal in close races, especially if this change occurs disproportionately among supporters of one candidate. Investigating the issue of differential impact therefore remains an important and worthwhile endeavor.

Moreover, the normative issue of disenfranchisement remains in spite of the null finding, too. Though the analysis does not support the claim that voter ID-laws will significantly lower turnout, this may be a result of various countervailing forces also resulting from the adoption of ID-requirements. The effect of these moderating forces, however, are mainly observable at the macro-level. If an individual who would otherwise have voted abstains due to lacking ID, while another who would not otherwise have voted, votes – regardless of exactly why – the net change in turnout will be zero. This does not, however, alter the fact that the first individual has effectively had his vote suppressed. This is to say, the null finding in terms of an aggregate effect does not necessarily exonerate voter ID-laws from charges of voter suppression at the individual level.

7.3.4 Implications for policy

Before discussing some policy suggestions, it is pertinent to reiterate and reflect on the partisan dimension of both voter ID-laws themselves and the heated discussion surrounding them, as it influences the interpretation of both the findings themselves and their political implications. While both Republicans and Democrats tout noble reasons for their respective support and opposition of voter ID-requirements – defense of electoral security in the former case and of the right to vote in the latter – it is likely that they are in part motivated by more practical concerns of electoral success. Interestingly, in this interpretation, both parties agree on the assumption that ID-requirements will lead to lower turnout. Though the possibility of a marginal effect remains, the results of this thesis suggest that they are mistaken in their prediction. Whether this is cause for relief or concern depends on partisan point of view. Still, if ID-requirements are merely a tool of electoral competition, the finding that they do not actually affect turnout should weaken support as well as opposition: Republicans have nothing to gain, while Democrats have nothing to fear.

If, however, support for voter ID-laws springs from genuine concerns of democratic fairness, some implications for policy could be agreed upon. First, requiring voters to identify themselves when voting is not an unreasonable policy in itself, despite the low likelihood of impersonation fraud. Opposition on the grounds of potential voter suppression appears unfounded in light of the present study. However, given the normative implications of committing a type II error and erroneously failing to reject the null hypothesis, there is, despite the findings of this thesis, an

argument to be made for erring on the side of caution in terms of ensuring against disenfranchisement. I argue for precisely such a better-safe-than-sorry approach.

While the processes underlying the present null finding remain black-boxed, the hypothesized counteracting, indirect effects of ID-laws highlight the potential usefulness of mobilization efforts. This should inform policy. For example, several studies have suggested that information and get-out-the-vote campaigns could help ensure against demobilization resulting from ID-requirements (Citrin et al. 2014; Bright & Lynch 2017). Through policy initiatives targeted at vulnerable populations, policymakers could effectively have their cake and eat it, too: reducing the potential for fraud while protecting against disenfranchisement.

For some voters, however, the issue lies not in dearth of information or motivation, but rather in their inability to acquire ID. Given that the possibility of ID-requirements affecting these voters negatively is not ruled out, neither by the findings presented here nor those of previous studies, some consideration is warranted on the part of policymakers pursuing ID-requirements. I contend that efforts to combat fraud through the passing of voter ID-laws should be accompanied by equally vigorous efforts to ensure that all those eligible to vote are provided with the documents now necessary to do so. This should, of course, appeal to opponents of voter ID-laws apprehensive about voter suppression; however, supporters of these laws should embrace such initiatives, too, as a signal of sincerity in their concern for *all* aspects of the electoral process.

7.4 Limitations

To end the discussion, I highlight some caveats and potential weaknesses of the analysis. First, it is worth reiterating the scope conditions of the inference. Recall that we are estimating the average treatment effect on the treated (ATT). Since the states that have adopted voter ID-requirements differ systematically from those that have not it is very difficult to infer from these results what the average effect would be if *every* state were to require ID. Thus, to the degree to which the findings are correct, they only speak to the effect voter ID-laws have had on turnout up until this point: just as we cannot predict with certainty how the effect estimate will evolve

among the treated states in the future, we cannot claim to know how turnout will be affected in potential new cases.

Relatedly, a key limiting factor for this thesis concerns data availability. For the synthetic control, more observations – of both T and N_{it} – would mean more precise estimates and increasingly accurate inference. Ideally, this would serve to narrow the wide confidence intervals accompanying the effect-estimate. Regarding the second part of the analysis, the switch to a difference-in-differences design is made necessary by the poor availability of data on turnout among demographic groups, which precludes a synthetic control approach. This is unfortunate given the reliance of DID on the parallel trends assumption, which – despite passing the pre-intervention pseudo assumption check – is always at risk of violation. My analysis is arguably particularly vulnerable due to the relative long interval between the pre- and post-treatment periods: logically, the odds of a time-varying confounder manifesting increases with the passing of time. Additionally, the CPS dataset is not ideal, in that it is mainly representative at the national level and plagued by overreporting of turnout. Though my analysis attempts to correct for these flaws, the solutions are likely imperfect.

Another qualification concerns the cause of interest itself. Though conceptualized here as a uniform intervention, voter ID-laws – even within the category of strict photographic ID-requirements – are not a monolithic treatment. Inter-state, intra-group variation in the exact requirements – like what forms of government-issued ID are acceptable – could therefore muddy the waters when estimating the consequences of strict photo ID-laws taken as a whole.²⁸ Similarly, changes to electoral rules rarely happen alone: ID-laws may be passed as part of a larger bill. Because both methods applied here estimate the causal effect relative to the *time* of treatment, neither can distinguish between effects of the treatment of interest and other interventions occurring at the same time. If some other event occurred simultaneously to the implementation of ID-requirements in a treated state, and this other event influenced turnout, the effect-estimate will be confounded.

²⁸ See the discussion of the stable unit treatment value assumption (SUTVA) in chapter 4 for more.

7.5 Further research

The issue of voter ID is not settled. Though the results of the present analysis indicate that ID-requirements are unlikely to cause major changes in turnout, unanswered questions and remaining obstacles to inference underscore the need for additional research. The longer these laws are in effect, the more data becomes available to researchers. Additional studies in the future are thus intrinsically beneficial in a general sense because they will have a greater empirical foundation from which to draw inferences. Additionally, it is of particular interest to see how the effect (or lack thereof) of voter ID-laws develops over the long term.

More specifically, further research is required at the micro-level. While this thesis conceptualizes causality in terms of actual and counterfactual outcomes, an alternative framework is to emphasize the *causal process* and the mechanisms linking cause and effect (Campaner 2011; Gerring 2010). Studies operating within this tradition could seek to empirically link the implementation of voter ID-requirements with individual voters' decision-making by investigating whether a chain of intervening factors connecting the two can be empirically verified. Concretely, one would seek to ascertain whether voters perceive that their costs of voting are increased as a result of voter ID-laws, and, if so, whether this perceived cost-increase makes them less likely to participate.

The somewhat surprising null finding offers another, more general suggestion for future scholars as well. The possibility that voter ID-laws spur mobilization to a degree that may neutralize or even reverse any negative effects on turnout highlights how election reform – and policy interventions in general – can sometimes have unexpected and unintended consequences (Burden et al. 2014). Likewise, it is a reminder to consider the indirect as well as the direct effects of a particular cause of interest. Further studies could seek to disentangle the two, and further investigate whether ID-requirements really do mobilize voters, and, if so, whether this process occurs through the mechanisms hypothesized here. As mentioned previously, mediation analysis offers a useful strategy (Imai et al. 2010).

Finally, while this study considered the effects of voter ID-requirements in the aggregate, the observed variation in effect estimates between states – including opposite directionality – highlights the need for examinations of individual states to investigate what causes this heterogeneity. Such studies could also help capture the idiosyncrasies of individual states' laws and compensate for the inability of synthetic control estimators to separate the effect of the treatment from other, contemporaneous interventions, by determining whether any such events occurred. The results of the synthetic control in this thesis identifies some potential candidates worthy of a closer look. In this regard, the synthetic control method has fulfilled its promise as articulated by Abadie et al. (2015) of bridging quantitative and qualitative approaches in comparative politics, by using statistical analysis to guide subsequent, more focused studies towards suitable cases for comparison.

References

- Abadie, Alberto & Javier Gardeazabal. 2003. "The Economic Costs of Conflict: A Case Study of the Basque Country." *The American Economic Review* 93(1): 113-32.
- Abadie, Alberto, Alexis Diamond & Jens Hainmueller. 2010. "Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program." *Journal of the American Statistical Association* 105 (490): 493-505.
- . 2015. "Comparative Politics and the Synthetic Control Method." *American Journal of Political Science* 59 (2): 495-510.
- Abadie, Alberto, Susan Athey, Guido W. Imbens & Jeffrey Woolridge. 2017. "When Should You Adjust Standard Errors for Clustering?" *NBER Working Paper Series*: 24003.
- Aldrich, John H. 1993. "Rational Choice and Turnout." *American Journal of Political Science* 37 (1): 246-278.
- Alvarez, R. Michael, Delia Bailey & Jonathan N. Katz. 2011. "An Empirical Bayes Approach to Estimating Ordinal Treatment Effects." *Political Analysis* 19 (1): 20-31.
- American Civil Liberties Union. 2017. "Fact Sheet on Voter ID Laws". Published: May 2017. <https://www.aclu.org/other/oppose-voter-id-legislation-fact-sheet>.
- Angrist, Joshua D. & Jörn-Steffen Pischke. 2014. *Mastering 'metrics: The Path from Cause to Effect*. Princeton, N.J: Princeton University Press.
- Athey, Susan, Mohsen Bayati, Nikolay Doudchenko, Guido Imbens & Khashayar Khosravi. 2018. "Matrix Completion Methods for Causal Panel Data Models." *NBER Working Paper Series*: 25132.
- Barreto, Matt A., Stephen Nuño, Gabriel R. Sanchez & Hannah L. Walker. 2019. "The Racial Implications of Voter Identification Laws in America." *American Politics Research* 47 (2): 238-249.
- Bentele, Keith G. & Erin E. O'Brien. 2013. "Jim Crow 2.0? Why States Consider and Adopt Restrictive Voter Access Policies". *Perspectives on Politics* 11 (4): 1088-1116.

- Bernstein, Robert A., Anita Chadha & Robert Montjoy. 2003. "Cross-State Bias in Voting and Registration Overreporting in the Current Population Surveys". *State Politics & Policy Quarterly* 3 (4): 367-386.
- Bertrand, Marianne, Esther Duflo & Sendhil Mullainathan. 2004. "How Much Should We Trust Differences-in-Differences Estimates?" *The Quarterly Journal of Economics* 119 (1): 249-75.
- Besley, Timothy & Anne Case. 2000. "Unnatural Experiments? Estimating the Incidence of Endogenous Policies." *The Economic Journal* 110 (467): 672-694.
- Biggers, Daniel R. & Daniel A. Smith. 2020. "Does threatening their franchise make registered voters more likely to participate? Evidence from an aborted voter purge." *British Journal of Political Science* 50 (3): 1-22.
- Biggers, Daniel R. & Michael J. Hanmer. 2017. "Understanding the Adoption of Voter Identification Laws in the American States". *American Politics Research* 45 (4): 560-588.
- Blais, André. 2000. *To Vote or Not to Vote? : The Merits and Limits of Rational Choice Theory*, Pittsburgh, Pennsylvania: University of Pittsburgh Press.
- . 2007. "Turnout in Elections." In *The Oxford Handbook of Political Behavior*, eds. Russell J. Dalton & Hans-Dieter Klingemann, 621-635. Oxford: Oxford University Press.
- Blais, André, Jean-François Daoust, Ruth Dassonneville & Gabrielle Péloquin-Skulski. 2019. "What is the Cost of Voting?" *Electoral Studies* 59 (1): 145-157.
- Brady, Henry E, Sidney Verba & Kay Lehman Schlozman. 1995. "Beyond SES: A Resource Model of Political Participation." *American Political Science Review* 89 (2): 271-294.
- Brennan Center for Justice. 2016. "Voter ID History". Published: September 28th 2016. Accessible from: <https://www.brennancenter.org/our-work/research-reports/election-2016-restrictive-voting-laws-numbers>.
- Briffault, Richard. 2002. "The Contested Right to Vote". *Michigan Law Review* 100 (6): 1506-1531.

- Bright, Chelsie L. M. & Michael S. Lynch. 2017. "Kansas Voter ID Laws: Advertising and Its Effects on Turnout." *Political Research Quarterly* 70 (2): 340-347.
- Brodersen, Kay H., Fabian Gallusser, Jim Koehler, Nicolas Remy, and Steven L. Scott. 2015. "Inferring Causal Impact Using Bayesian Structural Time-Series Models." *Annals of Applied Statistics* 9 (1): 247-274.
- Burden, Barry C., David T. Canon, Kenneth R. Mayer & Donald P. Moynihan. 2014. "Election Laws, Mobilization, and Turnout: The Unanticipated Consequences of Election Reform". *American Journal of Political Science* 58: 95-109.
- Campaner, Raffaella, 2011. "Mechanistic causality and counterfactual-manipulative causality: recent insights from philosophy of science". *Journal of Epidemiology and Community Health* 65 (12): 1070-1074.
- Cantoni, Enrico & Vincent Pons. 2021. "Strict ID Laws Don't Stop Voters: Evidence from a U.S. Nationwide Panel, 2008–2018." *The Quarterly Journal of Economics*, forthcoming. doi: 10.1093/qje/qjab019.
- Citrin, Jack, Donald P. Green & Morris Levy. 2014. "The Effects of Voter ID Notification on Voter Turnout: Results from a Large-Scale Field Experiment." *Election Law Journal: Rules, Politics, and Policy* 13 (2): 228-242.
- CNN. 2021. "Exit Polls". Accessed: June 20th 2021. <https://edition.cnn.com/election/2020/exit-polls/president/national-results>.
- Congress.gov. 2002. "H.R.3295 - 107th Congress (2001-2002): Help America Vote Act of 2002." October 29th 2002. <https://www.congress.gov/bill/107th-congress/house-bill/3295>.
- DeBell, Matthew, Jon A. Krosnick, Katie Gera, David S. Yeager & Michael P. McDonald. 2020. "The Turnout Gap in Surveys: Explanations and Solutions". *Sociological Methods & Research* 49 (4): 1133-1162.
- Desilver, Drew. 2020. "In past elections, U.S. trailed most developed countries in voter turnout". *Pew Research Center*. Accessed May 14th 2021. <https://www.pewresearch.org/fact-tank/2020/11/03/in-past-elections-u-s-trailed-most-developed-countries-in-voter-turnout/>.

- Doudchenko, Nikolay & Guido Imbens. 2016. "Balancing, Regression, Difference-In-Differences and Synthetic Control Methods: A Synthesis." *NBER Working Paper Series*: 22791.
- Dougherty, Christopher. 2011. *Introduction to econometrics*, 4th ed. Oxford: Oxford University Press.
- Downs, Anthony. 1957. *An Economic Theory of Democracy*. New York: Harper & Brothers.
- Erikson, Robert S. & Lorraine C. Minnite. 2009. "Modeling Problems in the Voter Identification—Voter Turnout Debate." *Election Law Journal* 8 (2): 85-101.
- Esarey, Justin & Ahra Wu. 2016. "Measuring the effects of publication bias in political science". *Research & Politics* 3 (3): 1-9.
- Evans, Jocelyn A. J. 2004. *Voters & Voting : an Introduction*, London: Sage.
- Ferman, Bruno, Cristine Pinto, Vitor Possebom & Burt Barnow. 2020. "Cherry Picking with Synthetic Controls". *Journal of Policy Analysis and Management*, 39(2): 510-532.
- Flood, Sarah, Miriam King, Renae Rodgers, Steven Ruggles & J. Robert Warren. 2020 "Integrated Public Use Microdata Series, Current Population Survey: Version 8.0". Minneapolis, MN: IPUMS. <https://doi.org/10.18128/D030.V8.0>.
- Fraga, Bernard L. 2018. *The Turnout Gap: Race, Ethnicity, and Political Inequality in a Diversifying America*. Cambridge: Cambridge University Press.
- Fund, John H. 2008. *Stealing Elections: How Voter Fraud Threatens Our Democracy*, 2nd ed. New York: Encounter.
- Gaskins, Keesha & Sundeep Iyer. 2012. "The Challenge of Obtaining Voter Identification." Brennan Center for Justice. Accessible from: <https://www.brennancenter.org/our-work/research-reports/challenge-obtaining-voter-identification>
- Gerring, John. 2010. "Causal Mechanism: Yes, But..." *Comparative Political Studies* 43 (1): 1499-1526.
- . 2012. *Social Science Methodology: A Unified Framework*. Cambridge: Cambridge University Press.

- Government Accountability Office. 2014. "Issues Related to State Voter Identification Laws." U.S. Gov. Account. Off., Washington, DC. Accessible from: <https://www.gao.gov/products/gao-14-634>
- Grimmer, Justin, Eitan Hersh, Marc Meredith, Jonathan Mummolo & Clayton Nall. 2018. "Obstacles to Estimating Voter ID Laws' Effect on Turnout." *The Journal of Politics* 80 (3): 1045-1051.
- Grimmer, Justin & Jesse Yoder. 2021. "The durable differential deterrent effects of strict photo identification laws." *Political Science Research and Methods*, published online by Cambridge University Press 28 January 2021: 1-17
- Grofman, Bernard. 1993. "Is Turnout the Paradox that Ate Rational Choice Theory?" In *Information, Participation, and Choice: An Economic Theory of Democracy in Perspective*, ed. Bernard Grofman, 93-103. Ann Arbor: University of Michigan Press.
- Gronke, Paul, William D. Hicks, Seth C. McKee, Charles Stewart & James Dunham. 2019. "Voter ID Laws: A View from the Public". *Social Science Quarterly* 100 (1): 215-232.
- Hajnal, Zoltan, Nazita Lajevardi & Lindsay Nielson. 2017. "Voter Identification Laws and the Suppression of Minority Votes." *The Journal of Politics* 79 (2): 363-379.
- Hansford, Thomas G. & Brad T. Gomez. 2010. "Estimating the Electoral Effects of Voter Turnout." *American Political Science Review* 104 (2): 268-288.
- Hasen, Richard L. 2012. *The Voting Wars: From Florida 2000 to the Next Election Meltdown*. New Haven: Yale University Press.
- Heller, Lauren R., Jocelyne Miller & E. Frank Stephenson. 2019. "Voter ID Laws and Voter Turnout." *Atlantic Economic Journal* 47 (2): 147-157.
- Hicks, William D., Seth C. McKee, Mitchell D. Sellers & Daniel A. Smith. 2015. "A Principle or a Strategy? Voter Identification Laws and Partisan Competition in the American States". *Political Research Quarterly* 68 (1): 18-33.
- Highton, Benjamin. 2017. "Voter Identification Laws and Turnout in the United States." *Annual Review of Political Science* 20: 149-167.

- Ho, Daniel E., Kosuke Imai, Gary King & Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15 (3): 199-236.
- Hood, M. V. & Charles S. Bullock. 2012. "Much Ado About Nothing? An Empirical Assessment of the Georgia Voter Identification Statute." *State politics & policy quarterly* 12(4) 394-414.
- Hur, Aram & Christopher H. Achen. 2013. "Coding Voter Turnout Responses in the Current Population Survey." *Public Opinion Quarterly* 77 (4): 985-993.
- Imai, Kosuke. 2017. *Quantitative Social Science: An Introduction*. Princeton, NJ: Princeton University Press.
- Imai, Kosuke, Luke Keele & Dustin Tingley. 2010. "A General Approach to Causal Mediation Analysis". *Psychological Methods* 15 (4): 309-334.
- Imbens, Guido W, and Donald B Rubin. 2015. *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge: Cambridge University Press.
- IPUMS CPS. 2021a. "About IPUMS CPS". Accessed January 11th 2021. <https://cps.ipums.org/cps/about.shtml>.
- . 2021b. "Voting and registration supplement sample notes". Accessed January 11th 2021. https://cps.ipums.org/cps/voter_sample_notes.shtml.
- Keele, Luke. 2015. "The Statistics of Causal Inference: A View from Political Methodology." *Political Analysis* 23 (3): 313-35.
- Keele, Luke & William Minozzi. 2013. "How Much Is Minnesota Like Wisconsin? Assumptions and Counterfactuals in Causal Inference with Observational Data." *Political Analysis* 21 (2): 193-216.
- Keyssar, Alexander. 2000. *The Right to Vote: the Contested History of Democracy in the United States*. New York: Basic Books.
- King, Gary, and Langche Zeng. 2006. "The Dangers of Extreme Counterfactuals." *Political Analysis* 14 (2): 131-159.
- . 2007. "Detecting Model Dependence in Statistical Inference: A Response." *International Studies Quarterly* 51 (1): 231-241.

- Kochhar, Rakesh & Anthony Cilluffo. 2018. "Incomes of whites, blacks, Hispanics and Asians in the U.S., 1970 and 2016." Pew Research Center. Updated: July 18th 2018. <https://www.pewresearch.org/social-trends/2018/07/12/incomes-of-whites-blacks-hispanics-and-asians-in-the-u-s-1970-and-2016/>
- Kuk, John, Zoltan Hajnal & Nazita Lajevardi. 2020. "A disproportionate burden: strict voter identification laws and minority turnout." *Politics, Groups, and Identities*: 1-9.
- Landman, Todd. 2008. *Issues and Methods in Comparative Politics*. London: Taylor & Francis Group.
- Leighley, Jan E. & Jonathan Nagler. 2013. *Who Votes Now?* Princeton: Princeton University Press.
- Liu, Licheng, Ye Wang & Yiqing Xu. 2020. "A Practical Guide to Counterfactual Estimators for Causal Inference with Time-Series Cross-Sectional Data." Available at SSRN: <https://ssrn.com/abstract=3555463> or <http://dx.doi.org/10.2139/ssrn.3555463>.
- Martinez, Michael D. 2010. "Why Is American Turnout So Low, and Why Should We Care?". In *The Oxford Handbook of American Elections and Political Behavior*, ed. Jan E. Leighley, 107-124. Oxford: Oxford University Press.
- Martinez, Michael D. & Jeff Gill. 2005. "The Effects of Turnout on Partisan Outcomes in U.S. Presidential Elections 1960–2000". *The Journal of Politics* 67 (4): 1248-1274.
- McDonald, Michael P. 2010. "American Voter Turnout in Historical Perspective". In *The Oxford Handbook of American Elections and Political Behavior*, ed. Jan E. Leighley, 125-143. Oxford: Oxford University Press.
- . 2021a. "Voter Turnout". *United States Elections Project*. Accessed February 1st 2021. <http://www.electproject.org/home/voter-turnout/voter-turnout-data>.
- . 2021b. "CPS Vote Over-Report and Non-Response Bias Correction". *United States Elections Project*. Accessed January 13th 2021. <http://www.electproject.org/home/voter-turnout/cps-methodology>.
- . 2021c. "What is left out of the voting-eligible population?" *United States Elections Project*. Accessed January 8th 2021. <http://www.electproject.org/home/voter-turnout/faq/leftout>.

- . 2021d. “What is the ‘Vote for Highest Office’ and ‘Total Ballots Counted?’” *United States Elections Project*. Accessed January 8th 2021. <http://www.electproject.org/home/voter-turnout/faq/numerator>.
- . 2021e. “Why should I care if turnout rates are calculated as percentage of VAP or VEP?” *United States Elections Project*. Accessed January 8th 2021. <http://www.electproject.org/home/voter-turnout/faq/vap-v-vap>.
- Minnite, Lorraine C. 2010. *The Myth of Voter Fraud*. Ithaca: Cornell University Press.
- Morgan, Stephen L. & Christopher Winship. 2007. *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. New York: Cambridge University Press.
- National Conference of State Legislatures [NCSL]. 2017. “Voter ID History”. Updated: May 31st 2017. <https://www.ncsl.org/research/elections-and-campaigns/voter-id-history.aspx>.
- . 2020. “Voter Identification Requirements | Voter ID Laws”. Updated: August 25th 2020. <https://www.ncsl.org/research/elections-and-campaigns/voter-id.aspx>.
- . 2021. “Voter Verification Without ID Documents”. Accessed: May 17th 2021. <https://www.ncsl.org/research/elections-and-campaigns/voter-verification-without-id-documents.aspx>.
- Norwegian Directorate of Elections. 2021. “Identification”. Accessed: May 16th 2021. <https://www.valg.no/en/election-in-norway/you-and-the-election/election-material/identification/>.
- package version 1.0.9*. <https://CRAN.R-project.org/package=gsynth>.
- Pryor, Ben, Rebekah Herrick & James A. Davis. 2019. “Voter ID Laws: The Disenfranchisement of Minority Voters?” *Political Science Quarterly* 134 (1): 63-83.
- Riker, William H. & Peter C. Ordeshook. 1968. “A Theory of the Calculus of Voting.” *American Political Science Review* 62 (1): 25-42.
- Robitzsch, Alexander & Simon Grund. 2020. “miceadds: Some Additional Multiple Imputation Functions, Especially for ‘mice’”. *R package version 3.10-28*. <https://CRAN.R-project.org/package=miceadds>.

- Rosenstone, Steven J. & John Mark Hansen. 1993. *Mobilization, Participation, and Democracy in America*, New York: Macmillan.
- Rubin, Donald B. 1974. "Estimating causal effects of treatments in randomized and nonrandomized studies." *Journal of Educational Psychology* 66 (5): 688-701.
- Samartsidis, Pantelis, Shaun Seaman, Anne Presanis, Matthew Hickman & Daniela Angelis. 2019. "Assessing the Causal Effect of Binary Interventions from Observational Panel Data with Few Treated Units." *Statistical Science* 34 (3): 486-503.
- Samii, Cyrus. 2016. "Causal Empiricism in Quantitative Research." *The Journal of Politics* 78 (3): 941-55.
- Schaffer, Frederic Charles, & Tova Andrea Wang. 2009. "Is Everyone Else Doing It? Indiana's Voter Identification Law in International Perspective". *Harvard Law & Policy Review* 3 (2): 397-412.
- Schulz-Herzenberg, Collette. 2019. "To Vote or Not? Testing Micro-level Theories of Voter Turnout in South Africa's 2014 General Elections." *Politikon* 46 (2): 139-156.
- Shapiro, Stuart & Deanna Moran. 2019. "The Costs of Voter ID Requirements." *The Regulatory Review*. Updated: January 8th 2019. <https://www.theregreview.org/2019/01/08/shapiro-moran-burden-voter-identification/>
- Sobel, Richard & Robert Ellis Smith. 2009. "Voter-ID Laws Discourage Participation, Particularly among Minorities, and Trigger a Constitutional Remedy in Lost Representation". *PS: Political Science and Politics* 41 (1): 107-110.
- Stewart, Charles III. 2013. "Voter ID: Who has them? Who shows them?" *Oklahoma Law Review* 66 (1): 21-52.
- . 2021. "How We Voted in 2020: A Topical Look at the Survey of the Performance of American Elections." MIT Election Data & Science Lab. Accessible from: <https://electionlab.mit.edu/research/projects/survey-performance-american-elections>.
- Stewart, Charles III, Stephen Ansolabehere & Nathaniel Persily. 2016. "Revisiting Public Opinion on Voter Identification and Voter Fraud in an Era of Increasing Partisan Polarization". *Stanford Law Review* 68 (6): 1455-1490.

- United States Census Bureau. 2019. "Frequently Asked Questions". Accessed January 19th 2021. <https://www.census.gov/programs-surveys/cps/about/faqs.html#Q3>.
- United States Department of Justice. 2020. "Jurisdictions Previously Covered by Section 5". Updated: September 11th 2020. <https://www.justice.gov/crt/jurisdictions-previously-covered-section-5>.
- Valentino, Nicholas A. & Fabian G. Neuner. 2017. "Why the Sky Didn't Fall: Mobilizing Anger in Reaction to Voter ID Laws." *Political psychology* 38 (2): 331-350.
- Vercellotti, Timothy & David Andersen. 2009. "Voter-Identification Requirements and the Learning Curve." *PS: Political Science & Politics* 42 (1): 117-120.
- Von Spakovsky, Hans. 2011. "Voter Photo Identification: Protecting the Security of Elections". *The Heritage Foundation Legal Memorandum* 70: 1-9.
- Xu, Yiqing. 2017. "Generalized Synthetic Control Method: Causal Inference with Interactive Fixed Effects Models." *Political Analysis* 25 (1): 57-76.
- Xu, Yiqing & Licheng Liu. 2018. "gsynth: Generalized Synthetic Control Method". *R package version 1.0.9*. <https://CRAN.R-project.org/package=gsynth>.

Appendix

A.1 Factor & loading from generalized synthetic control

Figure A.1 *Latent factor*

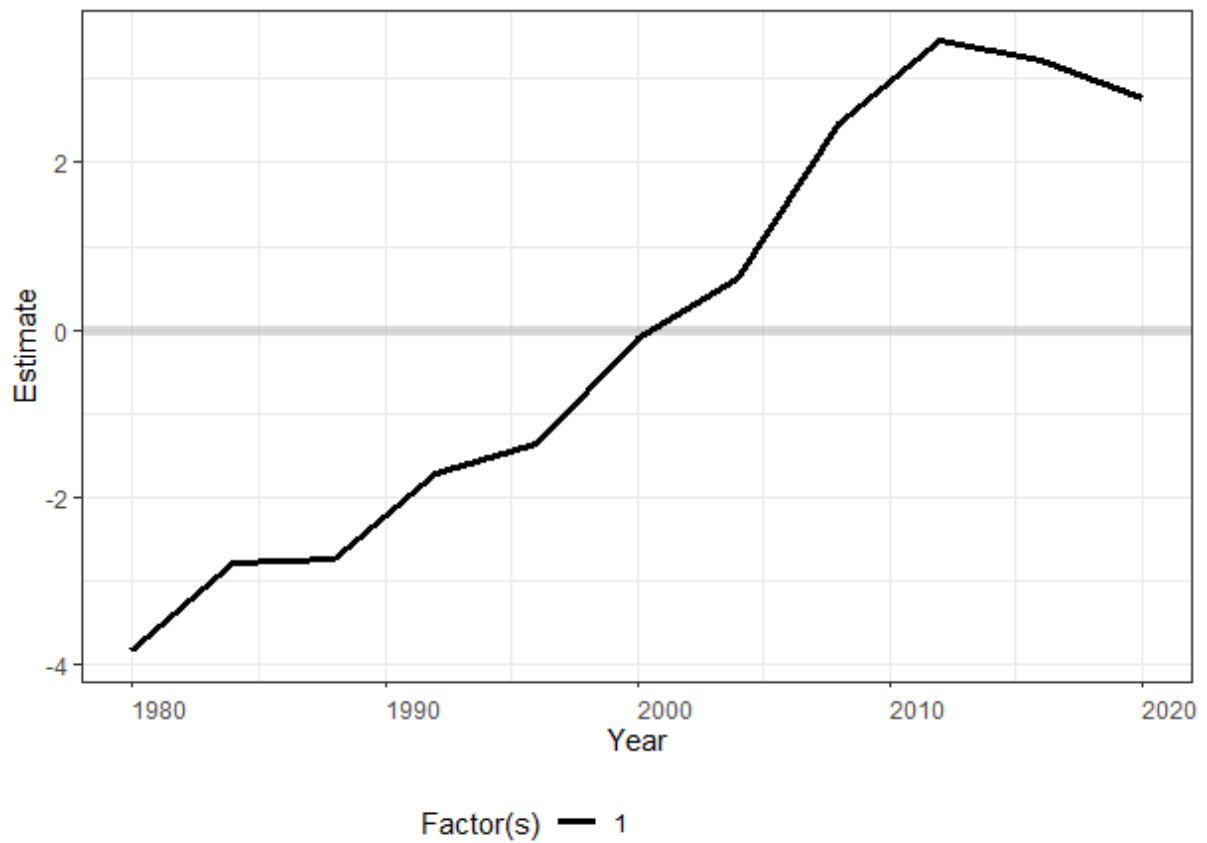
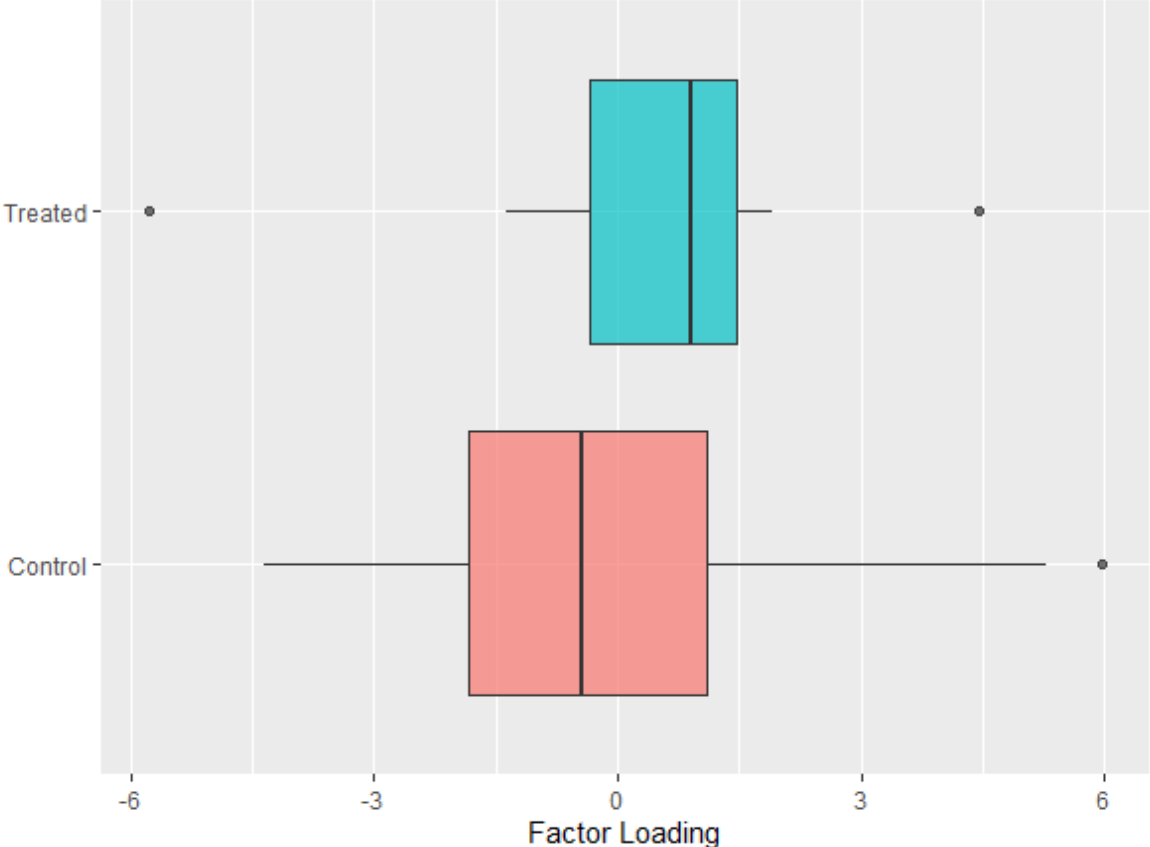


Figure A.2 Factor loading



A.2 Robustness checks for analysis with all elections

Figure A.3 Treated-counterfactual turnout difference (all elections – GSC)

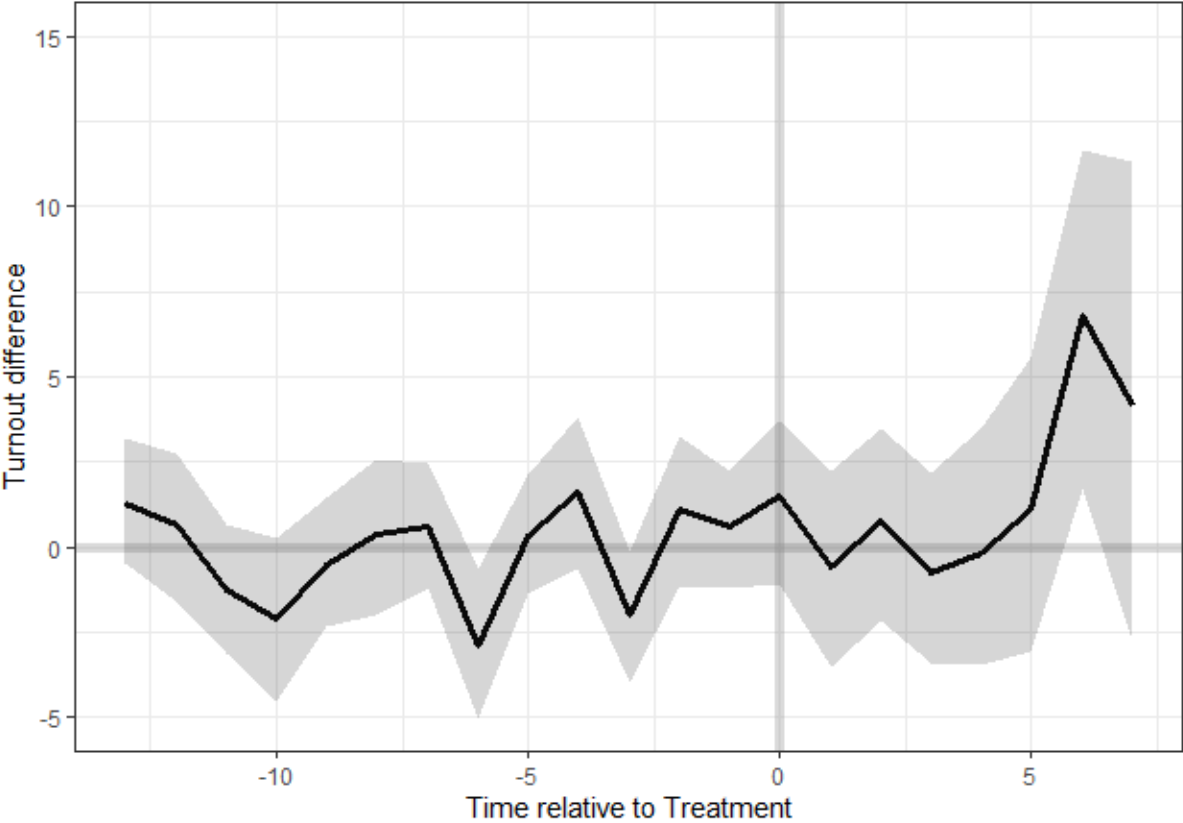


Table A.1 Estimated treatment effect (all elections – GSC)

ATT average	Standard error	CI. Lower	CI. Upper	P-value
0.60	1.23	-1.81	2.88	0.66

Figure A.4 *Treated-counterfactual turnout difference (all elections – restricted controls)*

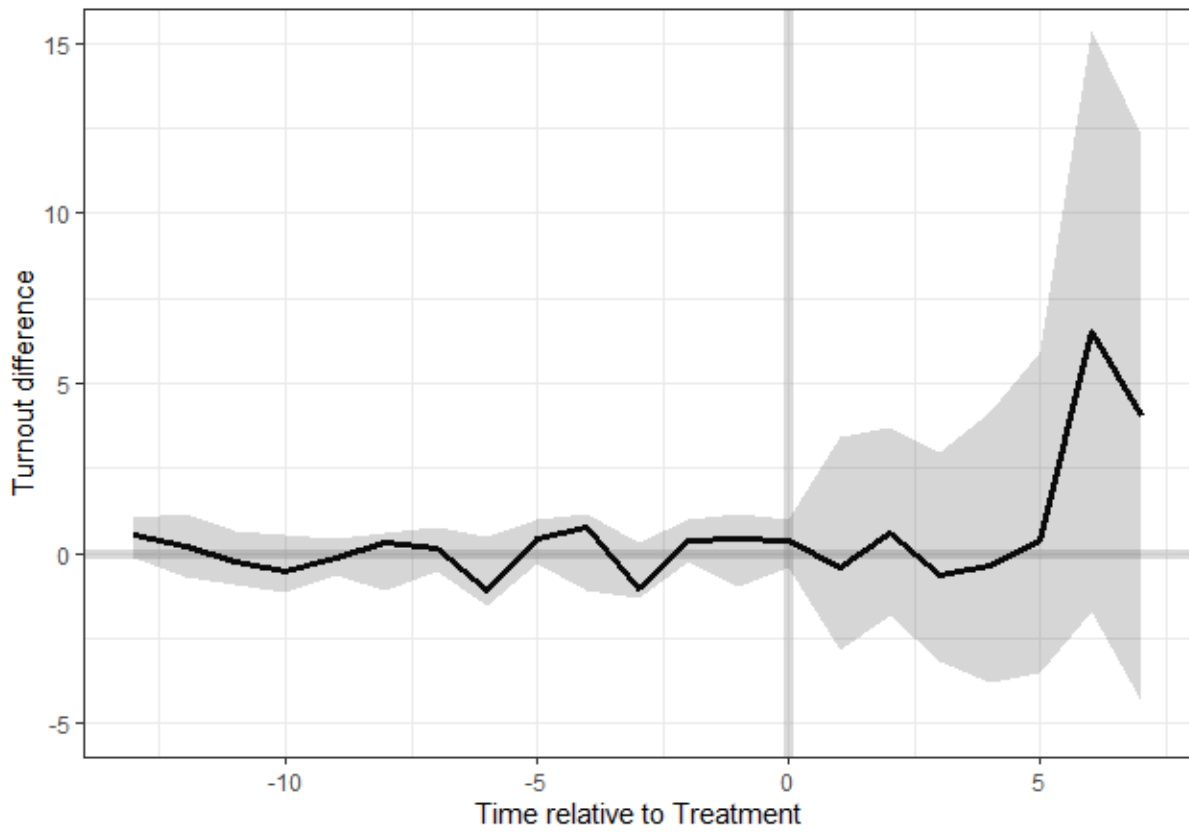


Table A.2 *Estimated treatment effect (all elections – restricted controls)*

ATT average	Standard error	CI. Lower	CI. Upper	P-value
0.49	1.98	-2.38	4.80	0.83