RICE UNIVERSITY

Helping America Vote? The Institutional Design of Elections and Recent Reforms

by

Greg William Vonnahme

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

[Signatures]

Robert M. Stein, Chair
Lena Gohlman Fox Professor, Political Science

Mark P. Jones, Professor, Political Science

Keith E. Hamm, Edwards Professor, Political Science

[Signature]

Richard R. Batsell, Jesse H. Jones
Distinguished Associate Professor, Jesse H. Jones Graduate School of Management

HOUSTON, TEXAS

MAY 2008
INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI®

UMI Microform 3309977
Copyright 2008 by ProQuest LLC.
All rights reserved. This microform edition is protected against unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 E. Eisenhower Parkway
PO Box 1346
Ann Arbor, MI 48106-1346
ABSTRACT

Helping America Vote? 

The Institutional Design of Elections and Recent Reforms 

By 

Greg Vonnahme 

This project focuses on the effects of recent electoral reforms on voter participation. I specifically examine three aspects of election administration, which are voter registration deadlines, early voting, and Election Day vote centers. The project builds from recent theoretical advances in the study of turnout to better understand the effects of the reforms and also suggests areas for future research. A new data set on county-level election results and survey responses are also analyzed using matching methods to test the effects of the reforms. The results suggest that registration deadlines and Election Day vote centers increase turnout, particularly for less politically engaged individuals, while early voting seems to negatively affect turnout.
Acknowledgements

Having written the rest of the dissertation, it is with a feeling of humility and great relief that I would like to acknowledge the many contributions that others have made to my graduate education and this project in particular. They have all made the process much more enriching and enjoyable. I was very fortunate to have a committee of excellent mentors. My advisor, Bob Stein, not only chaired my dissertation committee but also first got me started studying election reforms. His energy and creativity have been much appreciated in the course of writing this dissertation and he has also been very generous finding resources to help complete this research. Mark Jones and Keith Hamm have also provided excellent feedback, and this dissertation has been much improved as a result. They have also provided invaluable advice on other research, as well as on my job search and interviews this past year. I am also grateful to Randy Batsell for his comments, as well as Randy Stevenson for his help with this research and for the training he provided in our graduate methods seminars.

Many other faculty members provided help along the way, including (in no particular order) Gina Branton, Melissa Marschall, Rick Wilson, Ashley Leeds, Bill Reed, Ric Stoll, Cliff Morgan, and Lanny Martin. In addition to faculty at Rice, I have been very fortunate to have a wonderful group of friends and fellow graduate students at Rice. Michaela Mattes, Seiji Yamamoto, Rosa Sandoval, Burcu Savun, Valentin Krustev, Johanna Dunaway, Gavin Dillingham, and David Goodman deserve special recognition. Of course, I owe special thanks to Beth Miller for her advice, but more importantly for helping me get away from work every once in a while and helping me keep it all in perspective. Her ability to put up with me throughout this process is a testament to her enormous patience and good humor!
Many other people have helped at earlier points. I would especially like to thank Kelly Kadera, Nick Pudenz, and Beau Brindley. My sister, Julie, and her husband, Joe, have been supportive and wonderful hosts on my wintery trips back to Iowa. And last, I wish to acknowledge my parents who have managed to remain steadfastly encouraging throughout my extended stint in higher education. I’m sure that they of all people share in my excitement (and relief!) to finally be graduating and getting a job.
# Table of Contents

Abstract

Acknowledgements

List of Tables

List of Figures

Chapter 1
   Introduction

Chapter 2
   Evaluating the Effectiveness of Election Reforms

Chapter 3
   Registration Deadlines and Election Day Registration

Chapter 4
   Early Voting

Chapter 5
   Election Day Vote Centers

Chapter 6
   Conclusion

Bibliography

Appendix A

Appendix B

Appendix C
List of Tables

Table 2-1: Probabilities .................................................................................................................. 34
Table 3-1: Number of Observations for Voter Turnout Data by State by Year ......................... 116
Table 3-2: Number of Observations for Voter Registration Data by State by Year ............... 118
Table 3-3: Correlations .................................................................................................................. 120
Table 3-4: Results for Model of Turnout ..................................................................................... 120
Table 3-5: Results for Model of Registration ............................................................................... 120
Table 3-6: Reasons for Not Voting (CPS) .................................................................................... 121
Table 3-7: Reasons for Not Registering (CPS) ............................................................................ 121
Table 3-8: Method of Registration (CPS) .................................................................................... 121
Table 3-9: Matched Sample Analyses for Turnout ..................................................................... 122
Table 3-10: Matched Sample Analyses for Registration .............................................................. 122
Table 3-11: Substantive Effects from Model 1 ........................................................................... 122
Table 3-12: Effect of EDR on Turnout from County Data, 1992-2004 .................................... 123
Table 3-13: Synthetic Control Weights ....................................................................................... 123
Table 4-1: Difference of Proportions ........................................................................................... 148
Table 4-2: Turnout as a Ratio of VAP ......................................................................................... 148
Table 4-3: Balance of Matched Sample ....................................................................................... 149
Table 4-4: Table for Matched Sample ......................................................................................... 149
Table 4-5: Logit Model Estimates .............................................................................................. 150
Table 4-6: Substantive Effects ..................................................................................................... 151
Table 4-7: Model Estimates for Partisanship .............................................................................. 152
Table 4-8: Substantive Effects by Partisanship .......................................................................... 153
Table 4-9: Model Estimates for Interest ....................................................................................... 154
Table 4-10: Substantive Effects by Interest .................................................................................. 155
Table 5-1: Residential Distances to Polling Locations ................................................................. 193
Table 5-2: Average Turnout Over Counties .................................................................................. 193
Table 5-3: Difference-in-Difference Regression ......................................................................... 193
Table 5-4: Proportion of Treatment Cases by County ................................................................. 193
Table 5-5: Matched sample .......................................................................................................... 194
Table 5-6: Results of Matched Sample ......................................................................................... 194
Table 5-7: Logit model estimates .................................................................................................. 195
Table 5-8: Logit estimates with interaction effect ......................................................................... 196
Table 5-9: Controlling for Denver and Douglas counties ............................................................ 197
List of Figures

Figure 3-1: Histogram of Registration Deadlines ........................................... 104
Figure 3-2: Scatterplot of Turnout and Deadline Variables ................................ 104
Figure 3-3: Scatterplot of Registration and Deadline Variables .......................... 105
Figure 3-4: Scatterplot of Estimated County Turnout and Deadlines .................... 105
Figure 3-5: Plot of Turnout over Years .......................................................... 106
Figure 3-6: Scatterplot of State Turnout and Deadlines .................................... 106
Figure 3-7: Correlation Plot with Turnout ....................................................... 107
Figure 3-8: Correlation Plot with Registration ................................................ 107
Figure 3-9: States with Decreases in Deadlines ............................................. 108
Figure 3-10: States with Increases in Deadlines ............................................. 109
Figure 3-11: States with Multiple Changes ................................................... 109
Figure 3-12: Posterior Distribution for Turnout Estimates ................................. 110
Figure 3-13: Simulated Data Sets for Turnout ............................................... 110
Figure 3-14: Posterior Distributions from Registration Model ............................ 111
Figure 3-15: Simulated Data Sets for Registration .......................................... 111
Figure 3-16: Weights for Contagion Simulation ............................................. 112
Figure 3-17: Simulation Results ................................................................. 112
Figure 3-18: Registration Problems and Turnout over Deadlines ....................... 113
Figure 3-19: Proportion Missing the Deadline .............................................. 113
Figure 3-20: Predicted Probabilities of Voting ............................................. 114
Figure 3-21: Predicted Probabilities of Registering ....................................... 114
Figure 3-22: Differences in Turnout between Treated and Synthetic Controls ........ 115
Figure 3-23: Differences for Treated and Non-treated States ............................. 115
Figure 4-1: Priors ..................................................................................... 156
Figure 4-2: Difference between candidates ................................................... 156
Figure 4-3: Variation in candidates .................................................................. 157
Figure 4-4: Campaign information ............................................................... 157
Figure 4-5: Q-Q Plot .................................................................................. 158
Figure 5-1: Distance to Precinct and Vote Center Locations ............................. 198
Figure 5-2: Maps of Precincts, Vote Centers, and Registered Voter Populations ... 199
Figure 5-3: Q-Q Plots of Distance Measure ................................................... 200
Figure 5-4: Propensity Score Densities ......................................................... 200
Figure 5-5: Results of Sensitivity Analysis ..................................................... 201
Figure 5-6: Substantive Effects Controlling for Denver and Douglas ................... 201
Figure 5-7: Polling Place Simulation ............................................................ 202
Figure 6-1: Biological Contagion Results ..................................................... 215
Figure 6-2: Behavioral Contagion Results ..................................................... 215
Chapter 1:
Introduction

This dissertation broadly examines the effects of election reforms on voter turnout. Specifically I focus on how Election Day registration, early voting, and Election Day vote centers affect rates of voter turnout in recent U.S. elections. Research on election reforms is particularly relevant for policy debates as practices related to election administration have received increasing attention from policy-makers, scholars, and advocacy groups. This increased attention has come about in part due to more competitive elections with narrower margins. When the outcome of an election hinges on a small proportion of the votes, administrative practices that might influence the final ballot counts receive greater scrutiny, and recent elections have led to greater interest in election reforms at the federal level that have also had implications for state and local governments as well.

Obviously the most well-known recent example is the 2000 presidential election in Florida where George Bush won by a final certified margin of 537 votes, which was only a 0.009% margin of the total ballots cast in Florida. The 2000 Florida election raised issues of administrative practices related to both the casting and counting of ballots. Legal challenges related to voter intention and dimpled, pregnant, and hanging chads involved issues as to how the ballots were counted. The butterfly ballot on which a disproportionate number of voters were cast for Pat Buchanan in Palm Beach county raised questions about the design of the ballots themselves. Had the election not been so close it is likely that these issues would have remained out of the public spotlight, at least until the next close election when the same problems might have arisen in Florida or elsewhere.
In the wake of the 2000 presidential election, the federal government passed the Help America Vote Act in 2003 in order to modernize U.S. election practices and to avoid similar problems from happening again. The Help America Vote Act (HAVA) included a number of provisions designed to improve administrative practices related to the conduct of elections such as doing away with lever machines and punch cards, modernizing voting systems, requiring a procedure for provisional balloting, creating the Election Assistance Commission (EAC), and the establishment of statewide voter registration databases.

A number of states and counties have also undertaken recent efforts to reform election practices. Reforms such as Election Day registration and early voting have received particular attention from states. By 2007, 31 states had some form of early voting, and in the 2004 federal elections the American National Election Study estimated that 1 out of every 5 voters cast their ballots before Election Day. In the last two years Montana and Iowa have become the seventh and eighth states to adopt EDR, while a number of others such as New York, North Carolina, Pennsylvania, and Texas have considered adopting EDR. Other reform efforts such as mail-in voting, Election Day vote centers, online registration, greater poll-worker training, relaxed absentee voting, and even a state lottery for voters have been considered by a number of states and counties.

In addition to recent efforts to reform election practices there have also been efforts to pass more stringent voter identification laws. While most reform efforts are intended to make voting more convenient and increase participation, voter ID laws are aimed at reducing voter fraud at the risk of disenfranchising some voters. Advocates of more restrictive ID laws argue that stricter identification laws will result in fewer fraudulent votes at the polls while not creating a substantial barrier for legitimate voters. Opponents of voter ID laws
assert that voter fraud is a marginal problem and may disenfranchise voters, particularly the
elderly and lower income voters. While reforms aimed at increasing turnout can be
politically contentious (such as Maryland’s early voting law which failed a legal challenge in
2006), voter ID laws are particularly polarizing as they might differentially affect segments
of the population.

1.1 Motivation

The recent attention to election reform from policy makers creates an opportunity to advance
research on the effects of different reforms with additional resources being more widely
available. While advocates make claims as to the effects of election reforms there is often a
lack empirical evidence to fairly assess these claims. Beyond the contemporary attention
given to the election reform, I argue that there are three other important reasons for scholars
to study the administration of elections.

The first is that, as just argued, the reforms might affect levels of participation.
Participation might have both intrinsic and instrumental value. Some democratic theorists
argue that participation itself is normatively desirable and the intrinsic value of turnout
provides a rationale to pass reforms that increase rates of participation (Putnam 1995, 2000,
2001). Prominent political philosophers have long argued in favor of more participatory
democracy (Mill 1861, de Tocqueville 1904, Dewey 1939), which might be facilitated by
convenience voting reforms. Earlier research in American politics also advanced pluralist
theories of American politics, in which competing interests were thought to be represented in
politics proportionate to their level of popular support (Bentley 1908, Truman 1951, Dahl
1954), although subsequent research raised serious objections to pluralist theories on the
grounds that disincentives to participate would lead to an overrepresentation of smaller interests (Olson 1965), which further suggests that reforms that lower the costs of participation might help overcome the collective action problem. Others have expressed doubts about the intrinsic value of turnout, arguing instead that citizens would rather not have a regular role in politics (Hibbing and Theiss-Morse 2002). There may also be other instrumental values of turnout. In particular, participatory democracy has been argued to be important for preventing wars (Kant 1903), mass-elite linkages (Verba and Nie 1972), individuals' sense of community engagement, and interpersonal trust which in turn facilitates other social interaction (Putnam 2000, 2001). Participation in and of itself has been an active and important area of research and insofar as participation has either instrumental or intrinsic value it is potentially beneficial to know how different electoral procedures affect rates of turnout.

Participation determines not only the size of the electorate but also its composition. With less than full turnout, the population of actual voter might not represent the population of eligible voters and could influence the outcome of at least close elections. In that way, greater voter participation might lead to different election outcomes. It is often argued by political commentators that greater turnout would benefit Democratic candidates, and the NVRA bill of 1994 was vetoed by Republican President George H.W. Bush before being signed into law by President Clinton. Despite the widely held belief that greater turnout would benefit the Democrats, the actual empirical research on the partisan effects of higher turnout are much more qualified. Most of the research finds little partisan benefit to greater turnout (DeNardo 1980, Tucker, Vedlitz, and DeNardo 1987, Nagel and McNulty 1996, Highton and Wolfinger 2001 Citrin, Schickler, and Sides 2003). DeNardo (1980) argues that
the it was the minority parties in Congressional districts that used to benefit from greater turnout, but with the decline in “core” partisan voters in the 1960s the effect had substantially diminished. Tucker and Vedlitz (1987) reconsider DeNardo’s empirical results and argue that Democratic presidential candidates have benefited from greater turnout, but subsequent research has found few effects (Nagel and McNulty 1996, Highton and Wolfinger 2001 Citrin, Schickler, and Sides 2003).

Not only have nonvoters been shown to have a similar partisan distribution as voters, but also have similar issue positions as voters (Highton and Wolfinger 2001). Citrin et. al (2003) devise a novel method of combining exit poll data for voters with survey results from nonvoters to simulate election outcomes under conditions of full turnout. They generally find that while Democrats benefit somewhat from higher turnout there are few elections (they specifically examine Senate elections in 1994, 1996, and 1998) that are sufficiently close to be overturned. They do note, however, that there is substantial temporal and cross state variation in the effects which might warrant future research (for example, Texas has 10% more Democratic nonvoters in 1994 and 1996) and could on a limited basis affect election results. Descriptively, there is a difference between voters and nonvoters, as nonvoters tend to be younger, more residentially mobile, and less wealthy than voters (Rosenstone and Hansen 1993, Highton and Wolfinger 2001). There is also less research on the effects of greater turnout in other types of elections such as primary or local elections in which there might be greater differences between voters and nonvoters as well.

The third motivation for studying election reforms is that it might influence the perceived legitimacy of the government and individuals’ willingness to defer to the governments’ decisions. Recent research argues that effective governments depend on
individuals’ willingness to accept the authoritative decisions of others (Hibbing and Alford 2004). Deference to authority has been shown to depend on the method by which the authorities were chosen. Individuals show a greater willingness to defer to the decisions of meritocratic or democratic leaders than arbitrarily chosen, self-selected, or personally ambitious leaders. Election procedures might thereby affect the perceived legitimacy of government officials, as was expressed by some after the 2000 presidential election who claimed the election was “stolen”. Best practices of election administration might not only affect rates of participation but may also affect the representativeness and effectiveness of governments.

1.2 Previous Research on Election Reforms

The specific contributions of this project vary somewhat across the three different electoral reforms under consideration. Relaxed voter registration laws have been examined in earlier studies, which provide a foundation from which to build. Regarding registration laws, this study mainly seeks to clarify some of the theoretical arguments about the effects of registration policies and to help better identify the specific effects of registration laws. The extant research on early voting has tended to be based on a more narrow set of cases from states that were among the first to adopt early voting such as Texas and Tennessee. This project seeks to extend the analyses to include other observations as well, as 29 states have subsequently adopted early voting. Regarding Election Day vote centers, this project seeks to provide a detailed description of the reform, how it might affect turnout, and to provide empirical evidence about its likely effects.
Cross-nationally the US has comparatively lower turnout than most other advanced democracies with the exception of Switzerland (Powell 1986). Several institutional rules have been thought to affect turnout, including compulsory voting, proportional representation systems, multipartyism, and national registration (Powell 1986, Jackman 1987, Norris 2002, Blais 2006). In each case the US lacks the electoral policies thought to be related to voter turnout. The US does not have compulsory voting where individuals can be fined for not voting. The US also uses majoritarian electoral institutions which can lead to lower turnout, particularly when offices are subject to partisan redistricting which can reduce competitiveness (Jackman 1987). Proportional representation systems allocate seats based on vote proportions, so that one's vote does not have to be decisive regarding which candidate obtained a plurality of the vote in order to affect the outcome of the elections (Jackman 1987). The US also has a two party system which might reduce turnout among individuals with different combinations of political preferences that do not necessarily correspond to parties (Powell 1986). Evidence from cross-national research suggests that the electoral policies in the US might help to explain lower levels of turnout in US elections.

1.2.1 Voter Registration

The effects of voter registration laws on turnout in the US have been studied previously, but the different studies suggest somewhat different effects. Earlier research on registration laws argues that voter registration is a barrier for voter turnout in the US (Wolfinger and Rosenstone 1980; Squire, Wolfinger, and Glass 1987; Teixeira 1992; Patterson 2002). The estimated effect of voter registration laws was around a 10% decline in turnout (Squire, Wolfinger, and Glass 1987; Teixeira 1992). The effects were stronger when focusing on
certain populations such as recent movers and younger individuals (Squire, Wolfinger, and Glass 1987).

Subsequent policy reforms that liberalized voter registration laws have shown few conclusive effects, however. At least two types of registration reforms have been passed, including making registration forms more readily available and lengthening the amount of time that individuals can register in advance of an election (Mitchell and Wlezien 1995, Highton and Wolfinger 1998, Highton 1997, 2004, Berinsky 2005). Some states have notably adopted same-day registration whereby individuals can register at their polling place and be eligible to vote. North Dakota is the only state that does not require voter registration. North Dakota has not required registration since 1951, but allows for municipalities to establish local registration requirements. Closing dates in other states vary, but most are no more than 30 days prior to an election.

While earlier studies suggest that registration laws might be a significant barrier to turnout, studies of registration reforms have found that making voter registration easier has not had a significant impact on voter turnout (Mitchell and Wlezien 1995, Highton and Wolfinger 1998, Highton 1997, 2004, Berinsky 2005). In particular, same-day registration, which virtually eliminates the costs of registration, has not been shown to substantially affect voter turnout. There is some evidence that registration laws can affect certain groups of voters such as the residentially mobile, but also increases the education gap between voters and nonvoters (Highton 1997, Berinsky 2005). While registration reforms have not been shown to have a significant effect of increasing turnout rates, there is some evidence that they might further skew the electorate towards those with higher socioeconomic status.
1.2.2 Convenience Voting

Vote by mail elections are related to relaxed absentee voting insofar as more individuals cast ballots through the mail, but the dynamics are different in vote by mail elections. Vote by mail elections are conducted by mailing ballots to all voters on voter registration lists, and Election Day polling places are not available. Vote by mail elections have been used for smaller bond and special district elections in certain areas, but have also been used more recently for higher level elections.

An early study of vote by mail elections analyzes turnout data from larger cities from 1980 to 1984 (Magleby 1987). Magleby (1987) compares turnout in vote by mail elections to elections that used traditional precinct based locations. Magleby reports evidence of a 19% increase in turnout as a result of mail ballot elections and concludes that mail ballot elections can significantly increase voter turnout. There are several potential limitations to Magleby's study however. One is that he only considers the effects of the reforms on turnout as measured by the ratio of voters to registered voters. Yet the number of registered voters in a county might not be an accurate reflection of individuals that are actually eligible to participate as it can be influenced by individual mobility, mortality, and the timing of purges of the registration lists. A measure of voting age population, while also maybe not ideal (McDonald and Popkin 2001), might provide an alternative measure of voter turnout that could use US Census data which might be more valid than data from registration lists.

More recent studies of vote by mail elections analyze data from Oregon which has subsequently switched to all mail ballot elections. Southwell and Burchett (1997) report on survey data from individuals in Oregon conducted just before and just after the Oregon Senate elections in 1996 that used mail-in ballots. They surveyed 1225 individuals in Oregon
and found that vote by mail was very popular and there were few concerns among individual voters of privacy and security. They also find that vote by mail has few effects regarding partisan advantages and that the composition of the mail ballot electorate is comparable to traditional elections. One possible limitation of the study is that they do not seem to verify self-reported voting, which might overestimate the actual vote though they do not provide a comparison of voting rates among their respondents and statewide-turnout. In a subsequent study, Southwell and Burchett (2000) also report on survey data from Oregon voters and argue that vote by mail increases turnout but does not affect the representativeness of the electorate.

Other studies suggest different results, however, particularly regarding the possibility that vote by mail elections lead to a less representative electorate. Karp and Banducci (2000) examine data from Oregon elections between 1986 and 2000 and find that vote by mail elections lead to higher turnout, particularly among individuals with higher socioeconomic status. Their results do seem to show a significant amount of variation in the effects of vote by mail elections by the type of election, however. In particular, they find that turnout increases by as much as 26.5% for local elections and a 2.9% decrease in midterm elections. This might suggest that the effects vary by the type of election. Berinsky, Burns, and Traugott (2001) also find that vote by mail tends to increase turnout but also makes the electorate less representative of the eligible voting population by making higher income individuals more likely to vote. They also find that there is no evidence that vote by mail leads to an advantage for particular candidates or parties. Their conclusions are primarily based on empirical results from a Weibull model of voting across five elections, which seems to be an unusual model choice for their research question.
The empirical results regarding the effects of vote by mail elections seem to suggest that they have a positive effect on levels of turnout. There seem to be unresolved questions about mail ballot elections, though, regarding the effects on the composition and representativeness of the electorate and possible varying effects across elections.

Relaxed absentee voting policies are related to vote by mail elections in that they provide individuals with a greater opportunity to vote without having to go to a polling place on Election Day. The reform differs from vote by mail in that individuals need to contact local officials in order to receive an absentee ballot and election day polling places are still available for others. Existing research on the effects of relaxed absentee voting show mixed results. Oliver (1996) finds that absentee voting can increase the size of the electorate, conditional on party mobilization. That is, relaxed absentee voting policies can lead to higher levels of turnout, but only when parties and candidates in the election specifically attempt to mobilize voters around the reform (Rosenstone and Hansen 1993, Oliver 1996). In particular, he argues that parties can contact voters to encourage them to vote and inform them of absentee voting availability as well as mail absentee ballot applications to prospective supporters. Oliver (1996) finds that relaxed registration laws can increase turnout, conditioned on party mobilization, such that registration laws have no effect in the absence of mobilization.

In a subsequent analysis of absentee voting policies, Karp and Banducci (2001) find fewer effects on turnout and do not reproduce Oliver's (1996) aggregate level results using individual level data. They find that absentee ballot policies make individuals that were already likely to vote more likely to vote absentee. They find few effects that it increases the
size or composition of the electorate and do not find support for Oliver's conclusion that there
are joint effects of relaxed absentee policies and party mobilization.

Early voting was a reform first adopted in Texas in 1988 that allows individuals to
vote in-person prior to an election at special voting sites. Data from one of the first elections
that offered early voting, the 1992 presidential election, suggests that early voting might have
a positive effect on levels of turnout and the representativeness of the electorate. In particular,
early voting is shown to be related to new registered voters and the size of the Latino
population in the county (Stein and Garcia-Monet 1997). This suggests that early voting
might attract individuals that are otherwise less likely to vote and could increase turnout and
make the electorate more representative of the general population. They also find that these
effects might be conditional on party activities, as the Clinton campaign attempted to
mobilized potential supporters to vote early (Stein and Garcia-Monet 1997). Subsequent
research using data from the 1994 midterm elections in Texas finds that individuals that are
highly likely to vote are also more likely to vote early. That is, individuals that are stronger
partisans, more interested in politics, and older are more likely to vote early than individuals
that are younger, less interested in politics, and less partisan (Stein 1998). This might suggest
that early voting will have less of an independent effect of mobilizing new voters, but could
nevertheless have a joint effect with partisan mobilization.

These results are consistent with the findings from a survey of Knox County,
Tennessee voters in the 1996 presidential elections (Neely and Richardson 2001). Neely and
Richardson (2001) find that individuals that vote early are more likely to be concerned about
politics and attentive to the campaigns, but not necessarily more partisan or better educated.
They conclude that there is little evidence that early voting leads to higher turnout or that it
further biases the electorate (Neely and Richardson 2001). One possible concern about the study is that they sample “likely voters” without describing how they choose the sample. It is also not clear if they verify respondents' self-reported voting behavior, but they report an 89% turnout rate in their sample, when the county-level turnout among registered voters was 69%.

While the existing empirical results seem to suggest that the effects of early voting might be conditional on party mobilization or on individual characteristics such as attentiveness to politics, important questions seem to remain about the effects of early voting in more recent elections particularly given the increase in the number of states that allow for early voting. Further research might not only help to better generalize the existing results but might also help to answer the apparent empirical puzzle of the spread of early voting despite clear indications that it is effective at increasing turnout.

1.2.3 Election Day Vote Centers

Election Day vote centers are one of the newest electoral reforms to be implemented in the US. They were first adopted in Larimer County, Colorado for the 2003 local elections. Election Day vote centers differ from traditional methods of voting in that there are no precinct polling locations on Election Day. Rather than vote at a designated precinct location, individuals may vote at any county-wide vote center on Election Day. The vote centers are open to all county residents and tend to be larger and more centrally located than designated precinct polling sites.

Larimer county subsequently used Election Day vote centers in the 2004 coordinated primaries and the 2004 general presidential election. Anecdotal evidence from this election
suggests that vote centers were effective at handling a large number of voters, which attracted attention from the news media and other county election officials. Three other counties used Election Day vote centers in the 2005 local elections, including metropolitan Adams county that saw record levels of turnout in a local election. Election day vote centers will be used by eleven counties in Colorado, including Denver, in the 2006 midterm elections, with other counties in other states considering implementing the reform in future elections.

Given the recency of the reform, there are few findings that specifically address the effects of Election Day vote centers. Related studies might suggest possible effects, but with limited generalizability. Existing work on polling locations suggests that more distant locations tend to reduce individual turnout. Haspel and Knotts (2005) argue that distance from one's precinct location increases the costs of voting and thereby reduces turnout, but that this effect varies substantially with whether or not the individual has access to a car. Brady and McNulty (2005) reach a similar conclusion from their analysis of turnout data following precinct consolidations in California for the 2003 gubernatorial recall election. Specifically, they find that an individual's probability of voting declines by .25 percent for every tenth of a mile away between an individual's residency and the precinct location. Similarly, Gimpel and Schuknecht (2003) find that distance from one's polling place also reduces turnout, but that the effect was non-linear. They find that moderate distances reduce turnout while greater distances are related to higher turnout. They argue that the effect is attributable to moderate distances in suburbs that have more traffic congestion, while longer distances might be more common in rural areas with less traffic congestion that might make travel easier.
While these conclusions might suggest that vote centers reduce turnout by increasing individuals' distances from polling place, these works only consider residential distance from distributed precinct sites. Since Election Day vote centers are open to all voters in the county, individuals might be more likely to vote near their workplace, a school, or shopping center. That is, Election Day vote centers allow individuals to vote away from their residence at any site in the county. Individuals can vote near other destinations that might be more convenient, and while increasing one's residential distance, could decrease the distance from individuals' routine travel routes. This suggests that the results from the studies of designated precinct based locations will not necessarily extent to Election Day vote centers.

1.3 Research on Turnout

While there have been many empirical studies of voting and turnout, theoretical models of turnout are not necessarily well-developed (Achen 2006). The section begins with a discussion of Downs' well-known model of turnout and subsequent refinements. This section also reviews related arguments regarding the effects of individual resources and mobilization efforts on turnout.

1.3.1 Downs' Model of Turnout

A number of theoretical models have been provided by scholars that seek to explain individual turnout decisions. One of the earliest and most widely known is Downs' (1957) decision-theoretic model of individual electoral participation. According to Downs' model, an individual will vote when the costs of participation are outweighed by the benefits,
weighted by the probably of casting a decisive ballot. In Downs' model, the probability of casting a decisive ballot is the chance that an individual forces or breaks a tie in the election. The benefits of voting are equal to one's valuation of the material difference between having one's most preferred candidate in office and one's less preferred candidate. The costs of voting are resources that are expended in the act of voting, including the time and effort it takes to travel to the polls and cast a ballot and gathering the information to vote. In the strictest interpretation of Downs' model, an individual's decision to vote is entirely determined by the costs and benefits of voting.

The central problem with Downs' model is that in any large electorate the probability that an individual casts a decisive ballot is small. This weights the benefits of voting downward, suggesting that turnout should be very low. However, even in recent US elections which have historically and cross-nationally low turnout, levels of individual participation are still much higher than would be expected from Downs' theory. These observations gave rise to the paradox of voting where high levels of turnout are observed when theories predict that almost no one should vote.

While Downs' theory focuses on the material benefits of voting, others have considered non-material or expressive benefits of voting. Riker and Ordeshook (1968) provide a refined version of Downs' model that incorporates an additional parameter to account for the expressive benefits of voting. The central problem with Downs' model is that the probably of casting a decisive ballot is so low in a large electorate that it almost eliminates any material benefits from voting. By incorporating expressive benefits from voting, Riker and Ordeshook's (1968) model increases the total benefits of voting and potentially provide a better explanation for observed levels of turnout. The reason that the
incorporation of expressive benefits suggests higher turnout is that it is not conditioned by one's probability of being decisive. Expressive benefits are realized by the act of voting alone and do not depend on breaking or forcing ties between candidates, which allows for higher levels of turnout than in Downs' model.¹

Others have been critical of Riker and Ordeshook's model. While potentially allowing for more consistency with observed levels of turnout, at least two theoretical criticisms of Riker and Ordeshook's model have emerged. One is that the model does not provide an adequate justification for the parameter to capture expressive benefits. Some have argued that the explanation is weak as it suggests that voters that have a preference for voting will vote. Others have criticized the decision-theoretic nature of the model. Some scholars have offered alternative strategic models of turnout wherein the payoff to one voter is conditional on others' actions (Palfrey and Rosenthal 1985).

The game-theoretic models provided by Palfrey and Rosenthal (1985) treat voting as a large team game, where relative group levels of turnout affect the likelihood that one's candidate is elected. The game-theoretic models differ from earlier models of turnout in that individuals' decisions to vote or not are dependent on the others' decisions to vote or not. The mathematical models suggest that higher levels of turnout can be sustained in equilibrium, but only under a restrictive set of circumstances that are not robust to slight deviations from the model (Bendor, Diermeier, and Ting 2003). Furthermore, the assumptions made about individuals and the nature of their interactions, particularly the informational requirements of

¹ While Riker and Ordeshook's model that incorporates the parameter for expressive benefits has not been universally accepted for reasons that will be discussed, the model itself can just be seen as a generalization of Downs' original model. Riker and Ordeshook provide a four parameter model of turnout, whereas Downs' model can simply be seen as a special case of the Riker and Ordeshook model when the fourth parameter is set to zero.
the game-theoretic model are not likely to reflect how individuals actually make voting
decisions. The likely incongruity between the structure of the game-theoretic models and
individual behavior limits the ability of the model to help us understand individual turnout.
More recent theoretical models of turnout have also been provided and are discussed in the
next section.

1.3.2 Individual Resources

Individual resources have also been thought to influence political participation, particularly
educational and financial resources. The argument is that individuals with higher levels of
income and education will have more resources to devote to political participation and are
subsequently more likely to turnout in an election than if they had fewer resources (Verba
Miller and Shanks 1996, Patterson 2002). There are several ways in which higher incomes
might lead to greater participation.² Individuals with higher incomes might be more able to
bear the monetary costs of participation, such as spending time away from work or on
transportation costs. Individuals with higher education levels might also more readily process
political information, and might be more interested in politics as well. In this way, education
can both lower the informational costs of participation and potentially increase the benefits of
voting by leading an individual to be more interested in politics than if they had a lower level
of education (Verba and Nie 1972, Wolfinger and Rosenstone 1980, Popkin 1991,
Rosenstone and Hansen 1993, Miller and Shanks 1996, Patterson 2002). The argument that

² There might be a measurement issue with the use of income as a measure of individual resources, in that one
can have a relatively low income with high access to financial resources, and vice versa. The theoretical
argument seems most consistent with the availability of financial resources, but the empirical research largely
focuses on income, which could introduce some measurement error and affect substantive conclusions.
income and education are related to turnout seems consistent with the Downs' model of turnout, as it builds on the relationship between the costs and benefits of voting.

While the relationship between income and education and voting is well-known, recent research might suggest that the effects are more complex than has previously been thought. Fowler and Kam (2006) argue that individuals might have a personality trait associated with their willingness to bear short term costs for longer term benefits, and that this trait is related to voter turnout. They specifically argue that the costs of voting are realized immediately, while the benefits are accrued over time. This raises the possibility that individual discount rates on future benefits might lead to more or less weight placed on the costs of voting, which could affect turnout. They report empirical evidence that supports their conclusion that individuals that have a higher valuation of future benefits are more likely to turnout that those less willing to delay benefits. The theoretical arguments linking individual resources to turnout are not particularly strong, and these findings might suggest that the relations are confounded, at least in part, by individual personality traits. Existing research seems to support a relationship between willingness to delay benefits and income, education, and voting. That income and education have been found to be related to voting might suggest that the relationship is possibly confounded and the effects of individual financial and educational resources might be lower than previously thought, conditional on an individual's willingness to delay benefits.

1.3.3 Mobilization

3 On possible criticism of Fowler and Kam's model is that they do not specify the benefits of voting in a way that allows certain characteristics such as accrual rates to really be understood. Others have raised the possibility that voting might be associated with both consumption (short-term) and investment (long-term) benefits (Morton 2006).
Mobilization efforts have also been argued to affect individual turnout. Electoral mobilization might refer to efforts by parties, interest groups, and campaigns to increase voter participation in an election by contacting prospective voters and encouraging them to vote. Social mobilization might refer to the effects of informal social networks on political participation. Each might have a positive influence on individuals' likelihood of participation.

Electoral mobilization has recently been shown to have varying effects on individuals' likelihood of participation (Rosenstone and Hansen 1993; Gerber and Green 2000; Green, Gerber, and Nickerson 2003; Niven 2004; Imai 2005). Get out the vote efforts are shown to be most efficacious with door-to-door canvassing, where volunteers go to individual’s houses and encourage them to vote. Mailings have also been shown to have a positive effect, but are weaker than the effects of neighborhood canvassing. The effects of telephone calls has produce mixed empirical results depending on the method of analysis, but the most credible results seem to suggest the telephone canvassing can also increase turnout by a few percentage points (Gerber and Green 2000; Imai 2005).

Aside from get out the vote efforts, a source of recent debate regarding electoral mobilization involves the effects of negative advertising. Some have found that negative advertising tends to demobilize the electorate (Ansolabehere and Iyengar 1995; Ansolabehere, Iyengar, and Simon 1999; Ansolabehere, Iyengar, Simon, and Valentino 1994), while others find that it can be a useful source of information and might increase turnout (Lau and Pomper 2000; Goldstein and Freedman 2002; Freedman, Franz, and Goldstein 2004). Others have found that negative advertising can have differential effects depending on whether the negative messages are about substantive issues, or if advertisements engage in “mudslinging” and focus on personal characteristics (Kahn and Kenney 1999, 2004; Clinton and Lapinski
2004). While the theoretical and empirical results do not uniformly suggest a positive or negative effect of turnout, a growing consensus seems to be that they do not substantially demobilize the electorate (Brooks 2006).

Social mobilization focuses less on electoral activities and more on social engagement and community mobilization. Social context might help to mobilize individuals by facilitating the transmission of political information and could lead individuals to be more active and interested in politics (Verba and Nie 1972, Teixeira 1992, Miller and Shanks 1996, Oliver 2001). One important question that is the subject of ongoing research is the effects of cross-cutting social networks on participation. Cross-cutting social networks refer to an individual's social network in which different affiliates have competing political views. Cross-cutting social networks might provide individuals with greater access to information and thereby make turnout more likely than homogeneous social networks, but might also lead individuals to withdraw from politics because of a desire to avoid conflict (Mutz 2002).

Mobilization might generally serve to lower the costs of voting by facilitating the social transmission of information and increase the benefits of voting to produce higher turnout. While not the main focus of this project, future research on the effects of mobilization on participation might further consider the nature of the benefits of voting. Since the material benefits of voting seem very small, there might be a greater role for symbolic utility in accounting for turnout. Yet the nature of symbolic utility makes it difficult to define, which has limited research on how different mobilization efforts might increase or decrease individuals' symbolic utility of voting. One possibility might be that the act of voting symbolizes other valued outcomes such as having one's preferred party or candidate win the election, or being a part of a democratic political process. The possibility that the act
of voting is symbolic of certain outcomes suggests that the utility of those outcomes might be instantiated back to the act of voting (Nozick 1995). To increase the symbolic benefits of voting, mobilization efforts might be able to increase either the utility of the symbolized outcomes, or the degree to which the utility of the outcomes is instantiated back to the act of voting. Further research might be done on the symbolic nature of voting and might usefully draw on findings and methods in anthropological research regarding symbolic relations (Geertz 1973). It might be worth considering whether different mobilization efforts vary in their efficacy depending on how they affect the utility of symbolized outcomes or the instantiation process.

1.4 Conclusion

The purpose of this dissertation is to help us to better understand the consequences of registration reform, early voting, and Election Day vote centers for voter turnout. These reforms represent the major recent efforts to reform the way that voter registration and elections are conducted. The extant empirical research has examined registration deadlines with mixed findings, early voting to a lesser extent, often relying on more limited evidence, and to a much lesser extent EDVCs.

The remainder of the dissertation is primarily organized around these three election reforms. A separate chapter is devoted to each of the reforms. The first empirical chapter analyzes the effects of the Election Day registration, the second examines early voting, and the third considers EDVCs. Theoretical and empirical considerations specific to each reform appear in the respective chapters. Before discussing the main research findings, the second chapter discusses the rationale for the methods used in this project, describing the challenges
of estimating the causal effects of the reforms from nonexperimental data. I attempt to offer a brief justification for the "quasi-experimental" designs used in the later chapters. The final chapter considers the implications of the results and directions for future research.
Chapter 2:
Evaluating the Effectiveness of Election Reforms

2.1 Introduction

The previous chapter discusses the research questions and extant research. This chapter considers the relevant philosophical and statistical literature to help guide this study of the causal effects of the reforms. Focusing on the causal effect of the policy seems intuitive as we would like to understand what would happen if different policies were chosen. The extant research, however, does not seem to be exclusively aimed at estimating causal effects, and thus the focus of this research might differ from other studies, and addressing causal questions with nonexperimental data will be shown to raise a number of challenges.

Causal relationships have been given greater attention in recent political science research (explicitly and implicitly), and is by no means exclusive to the study of election reforms. Several prominent political science texts argue that causal questions are at the center of qualitative and quantitative political science research (King 1991, Fearon 1991, King, Keohane, and Verba 1994). Psychologists have found that there is a human tendency to seek causal explanations for observed phenomena. The ability to quickly and easily learn about causal relationships has been found in children as young as two years old (Gopnik and Sobel 2000, Kushnir and Gopnik 2005, Schulz, Gopnik, and Glymour 2007), though patterns of causal attribution may differ somewhat across cultures (Nisbett 2005).

While causal relationships seem sensible as reflected in general human tendencies and in applied research settings, it is arguably not the only aim of scientific research (Bunge 1959). Causal studies can be distinguished from normative, descriptive, and predictive
studies. Normative studies involve value judgments, and voter turnout in particular might raise related normative questions such as those discussed in Chapter 1 regarding philosophical arguments about participatory democracy and representative government. Descriptive studies seek to provide detailed information about some phenomenon of interest (i.e. income inequality). There are descriptive aspects to this study, such as the distribution of registration deadlines across the country, but the main purpose of the project is to examine the causal relationships.

Other studies focus primarily on prediction, in that they assess what outcomes are more or less likely to occur in the future. Predictive studies might anticipate election outcomes, the occurrence of international or domestic crises, rates of climate change, weather models, financial markets, and so on. Predictive research is less concerned with causal factors than the accurate prediction of future outcomes and can also be distinguished by their focus primarily on a single (outcome) variable.

In contrast to what some previous political science scholars seem to argue, this project does not claim that the only questions of scientific interest are causal questions, although they seem to be the most common. Some have argued, however, that causal relationships have less scientific value and texts on research methods often provide only limited discussions of causal inference (Pearl 2000). The remainder of the chapter briefly outlines what is meant by a causal relationship for the purpose of this study.

### 2.2 Defining Causality

This section attempts to distinguish between explanation and causality in general, and to consider different possible formulations of causality that might help to guide this study. In
Aristotle's very early discussion of causality, he argues that there are several different types of causes, but does not seem to draw a clear distinction between causation and explanation (Sobel 1995). We might prefer to draw a sharper distinction as research such as historical narratives seem to fit in with a general understanding of explanation, but not necessarily causality. For example, there might be many explanations for the origins of the Cold War, such as the bipolar international system, the development of nuclear weapons, U.S. support for the government of Greece, Truman's presidential campaign in 1948, etc. While historical research might provide an explanation for a particular event it might not address more enduring factors or relationships. Causality might be seen as a subset of explanations but not identical or exhaustive. To better distinguish causation from explanation this section considers two approaches specific to causal relationships.

The first approach to causality that we can consider is regularity theories of causality (Sobel 1995). Regularity theories are consistent with Mill's discussion of "causes of effects". Regulatory theories of causality are concerned with how one can identify factors that cause a certain outcome (Sobel 1995). Hume argues that there are three criteria, that if met, it can be said that something causes something else (Hume excludes the case of self-causation). Hume's three criteria are temporal order, proximity, and constant conjugation. According to the first criteria the cause must precede the effect so that the effect will neither occur before the cause nor will the two occur simultaneously. The second criterion is physical proximity so that the cause will be proximate to the effect. The third criterion of constant conjugation requires that the effect results from the cause in every instance in which the effect is present. The third criterion requires that for a relationship to be causal it must be a general phenomenon rather than one that only happens in a particular instance. Hume provides as an
example of causal relationships one ball in motion colliding with another ball, sending the second ball in motion (Sobel 1995).

Hume further argues that causal relationships do not exist in nature but are rather mental constructs that result from our observation of relationships that meet his three criteria (Sobel 1995). In this way causal relationships are a mental construction with no actual empirical referent. This raises at least one potential problem for the identification of a causal relationship, in that it would seem to exist only insofar as a counter example has not yet been found. As a causal relationship requires constant conjugation and is without empirical referent, causal relationships may be time-specific in that every currently prevailing causal relationship depends on the absence of future counter examples which we may have reason to doubt if there is no actual physical process governing the relationship between the factors.

Hume’s analysis has been critiqued on a number of grounds. The requirement that causes temporally precede effects has been disputed by some that argue that causes and effects can occur simultaneously. Mill uses the example of a ball resting on a cushion and causing an indentation as an example of simultaneous causation. Others have argued that in order for causation to be immediate it must occur simultaneously, otherwise there are intermediate causes (Sobel 1995).

Alternatives have also been offered to constant conjugation, requiring more general types of consistency that might capture differences between necessary and sufficient conditions for an effect to occur. The different regularity theories of causation seem to apply most to the natural sciences, in that they seek to find the causes of effects. These theories seem relevant to some political science studies such as those that have examined why parties exist, how voters’ make decisions, why governments pass certain types of policies, or why
wars occur for just a few examples. While these approaches have been useful in the natural sciences, they might generally be less useful for political science studies. The problem with applying regularity theories of causality to political science is that they require level of knowledge about political phenomena that seems unobtainable. Almost any social outcome of interest is likely to involve a huge number of factors that could not possibly be fully examined.

Regularity theories have also been criticized as not seeming to apply to many scientific and practical instances when we use the word cause (Sobel 1995). Collingwood (1948) argues that the target of a cause is usually something in nature such as a flat tire and its cause is an event or state of things such that if we arranged for the event to occur or the state of things to come about then we would produce that which it is thought to cause, such as running over a nail. This is more consistent with applied research, particularly in an experimental context where we can think about creating certain conditions and then observing the response.

There are a number of ways in which this definition of causality differs from regularity theories. The first is that there is no assumption that the cause is the full cause of the event (Sobel 1995). According to this definition, a nail could still cause a flat tire even though a flat tire could also result from a broken valve, blowout, pinch flat, etc. According to this account of causation those other factors become background conditions, so that we could say in the absence of the other factors, running over nail causes a flight tire.

This also means that there is no direct competition between alternative explanations for an event. While a broken valve can be a background condition if we are interested in whether or not running over a nail can cause a flat tire the two potential causes are not
mutually exclusive. While a broken valve might be a background condition in one study, it might be the cause of interest in another study where running over a nail would be a background condition. We can see this occasionally in "great debates" that emerge in political science over whether some event results from institutional or behavioral conditions, whether roll calls are determined by parties or preferences, whether justices rule based on principles or preferences, and so on. According to this view of causality, these debates are misguided as they seem to construe alternative causes as competing which is not necessarily implied by this view of causality.

This approach to causality also implies a temporal ordering so that the cause precedes the effect at least in terms of the design. The manipulation account involves consideration of a cause that is first set at some value, the background conditions held constant, and the outcome is then observed. While the outcome event might possibly occur simultaneously the observation would occur after the planning and manipulation of the cause of interest.

This perspective on causality also implies counterfactual conditions for causal relations, which helps to resolve the concerns regarding the exclusion of regular, non-causal sequences (Sobel 1995). Events such as night and day exhibit characteristics that seem to fit Hume's criteria for causation but are not usually considered to be causal. The counterfactual condition would seem to exclude cases of non-causal sequences and is also consistent with counterfactual approaches to causality in recent political science research (Fearon 1991, Morgan and Winship 2007).

The manipulation account of causality helps to resolve some of the concerns about regularity theories of causality. These two general approaches to causality are both deterministic, and while there has been some research on probabilistic accounts of causality
many of the issues are similar (Suppes 1970). A recent approach seems particularly promising that analyzes graphical models of causality (Pearl 2000).

While this discussion helps to frame the study it is not entirely clear that it helps to identify the best methods of analysis for the subsequent chapters. Recent statistical research provides a more formal definition for causal relationships that supports counterfactual claims and is consistent with the manipulation account.

The potential outcomes framework has developed in recent decades as a formal expression of causality. According to the potential outcomes framework (or Rubin causal model) the causal effect of X (assumed dichotomous with values 1 and 2) on Y for some unit u in a universe of units is defined as \( Y_{u1} - Y_{u2} \) where the subscript u indicates that the outcome is for unit u and the number indicates the treatment condition under which the outcome is observed (Rubin 1978, Holland 1986, Rubin 1990). The difference represents the causal effect of X on Y for u (the difference is most often taken to be the relevant quantity although in principle there is no reason that a ratio or other comparison could not also be made). It may be worth noting that in this formulation the effect of a treatment is always defined relative to some other treatment (most often taken to be a control). A “full” data matrix describing the causal relationship would have two outcome variables and a treatment indicator for the observed outcome (Rubin 1978).

While the potential outcomes framework provides us with an explicit definition of the causal quantity to be estimated it also raises the fundamental problem of causal inference that the difference is unobservable (Holland 1986). The problem is that we can never simultaneously observe \( Y_{u1} \) and \( Y_{u2} \) as X will either have value 1 or value 2 but never both.
There are two potential solutions to the fundamental problem of causality, one scientific the other statistical (Holland 1986).

The scientific approach resolves the unobservability problem by comparing outcomes under treatment 1 to outcomes under treatment 2 for the same unit at different points in time. The scientific approach is unlikely to be useful for the study of election reforms for a number of reasons. In this approach there is a tradeoff between manipulability and temporal comparability. The more quickly a system can be manipulated the more likely the background conditions are to remain constant and provide valid results. A previously discussed example considers flipping a light switch. If we turn on a light switch, the effect is quickly realized, and once we turn the switch off, it has no further effect on future trials (Holland 1986). However, election laws are not typically manipulable by researchers, and additionally, research findings suggest that voting is habit forming raising concerns as to the transience of the treatment effects. The less transient the treatment effects, the longer the intervals would have to be to isolate the other outcome, yet the longer the interval the less likely it will be that the background conditions will be comparable. The statistical approach is potentially more useful and I consider the conditions in which the statistical approach can be used.

The statistical approach shifts the task away from estimating unit-specific effects to $E[Y_{u1} - Y_{u2}]$ for all $u$ in the universe (Holland 1986). By the rules of probability this is equal to $E[Y_{u1}] - E[Y_{u2}]$. This raises the possibility that observations from one unit can be used to compare the potential outcome under the other treatment for the other units in the universe. This can also be thought of as a missing data problem, where the values are missing completely at random when there is a randomized and known treatment assignment.
mechanism. The statistical solution resolves the fundamental problem of causal inference as units exposed to both treatments can be observed for comparison. One challenge for the statistical solution is the comparison is actually between \( E[Y_{w1}] \) and \( E[Y_{w2}] \) where the \( W \) indicates that the outcome is observed. In order for the comparison to provide a valid estimate of the difference, \( E[Y_{u1}] - E[Y_{u2}] \) should equal \( E[Y_{w1}] - E[Y_{w2}] \). Whether this is true depends on the how the treatments are assigned, which determines which of the outcomes are observed (Rubin 1976, King et al. 2001). This leads to consideration of alternative treatment assignment mechanisms and how different mechanisms might require different assumptions in order to compare the observed differences to the counterfactual differences. The following section briefly reviews the analysis of randomized experiments (where the assignment mechanism is known), followed by a discussion of observational studies in which the treatment assignment mechanism is unknown.

### 2.3 Randomized Experiments

Experimental studies where subjects are randomized to treatment conditions provide the most information for researchers attempting to empirically study the effects of a particular factor (in most cases). Randomized experiments are particularly common in medical and psychological research but are becoming increasingly common in political science as well. Having a clearer view of the role of randomization in experimental research will help us to better understand the consequences of not having random assignment in non-experimental settings.

Modern applications and research on the design of experiments go back to Fisher's research with agricultural experiments (1949). Fisher argues that randomization in
experimental design is the "reasoned basis for inference" and describes at length an example of an experiment involving a lady tasting tea to explain the role of randomization. The example is motivated by a woman who claims that she can taste the difference between a cup of tea infused with milk and one in which the milk was added first followed by the tea. To test the woman's claim Fisher devised a randomized experiment for the woman to taste different cups of tea and classify if the tea or milk was poured first.

The experiment proceeds by setting up 8 cups of tea, 4 that have milk added first, 4 that have tea added first. The cups are then presented to the woman in a random order. She tastes each of the cups and then classifies each accordingly. Fisher then calculates the number of cups she correctly classifies to assess whether she can taste the difference between the two types of tea. In the experiment, randomization only plays a role in determining the order of the cups.

To use the results to assess the woman's claim, Fisher begins by tentatively assuming that the woman has no ability at all to discriminate between the cups, and makes her classifications solely by guessing their order. This step is Fisher's formulation of the null hypothesis. He then uses the randomization of the sequence of cups to find the probability that the woman would classify X number of cups correctly if she were only guessing and every correct cup was solely a random occurrence. By using the random assignment mechanism under the assumption that the woman has no ability to discriminate between the cups, the probability of the number of correct cups (between 0 and 8) follows a hypergeometric distribution with 8 trials. Table 2-1 shows the probabilities from the hypergeometric density function (there are no odd numbers as missing one cup means a cup of the other type will also be missed).
Table 2-1: Probabilities

<table>
<thead>
<tr>
<th>Correct</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>.014</td>
</tr>
<tr>
<td>2</td>
<td>.228</td>
</tr>
<tr>
<td>4</td>
<td>.514</td>
</tr>
<tr>
<td>6</td>
<td>.228</td>
</tr>
<tr>
<td>8</td>
<td>.014</td>
</tr>
</tbody>
</table>

Fisher uses these probabilities to assess the plausibility of the no effect hypothesis. If the woman has no ability to discriminate the cups there is only a 1/70 chance that she would correctly classify all 8 cups. With a sufficiently small probability, Fisher argues that we can reject the no effect hypothesis and conclude that the woman has some ability to discriminate between the different cups of tea. For other responses, the probabilities are greater. For example, there is a 23% chance that she could guess 6 out of 8 cups even if she could not taste any difference, so we might not rule out the no effect hypothesis if we only observe 6 correct responses. She supposedly classified all 8 cups correctly, allowing us to reject the null hypothesis and providing support for her claim.

This approach described by Fisher forms his "reasoned basis for inference" and the assignment mechanism, which is known by the researcher, provides the "physical basis for the test." The randomization test of the hypothesis of no effect, as described by Fisher, does not require assumptions about independent events, that all of the cups be identical, or that the woman is representative of a larger population (Rosenbaum 2002). Fisher goes on to use this approach to study different experimental data on crop yields. His use of randomization tests for null hypotheses has subsequently been blended with a similar approach advocated by Jersey Neyman, and generalized beyond the analysis of no effect hypotheses using randomization tests.
Randomization tests for null hypotheses follow from our discussion of the potential outcomes framework as the observed units in the treatment conditions may be compared to evaluate the plausibility of the null hypothesis under a given assignment mechanism. Fisher's randomization test for the null hypothesis in terms of potential outcomes begins by assuming that \( E[Y_{u1}] - E[Y_{u2}] = 0 \). Under random assignment and the assumption of no treatment effect, \( E[Y_{w1}] - E[Y_{w2}] = 0 \) as the treatment is without effect and \( p(Y=y | W=w) = p(Y=y) \) for all \( y \) in \( Y \) and all \( w \) in \( W \) (Rubin 1976, King et al. 2001). The plausibility of the observed difference being equal to 0 can thereby be evaluated using a similar procedure for randomization tests as discussed above. While randomization tests are somewhat limited in the evaluation of the no effect hypothesis, they can be justified with virtually no assumptions whereas other analyses can potentially provide more information but with additional assumptions.

The foregoing discussion is intended to show how experimental designs with randomization tests can be used to evaluate the causal relationship between a treatment and outcome variable in a way that is consistent with the potential outcomes model of causality. This method depends on the researcher controlling (or at least knowing) the treatment assignment mechanism. If the treatment assignment mechanism is unknown it is unlikely that the above randomization tests would be a justifiable means of analyzing the data. In the next section we consider alternative means of evaluating non-experimental data with an uncontrolled and unknown assignment mechanism.

2.4 Observational studies
Observational studies differ from experiments in that the researcher does not have the ability to manipulate the treatment condition, and assign the units to particular conditions. Without assigning units to treatments, there will most often be little basis for assuming that the treatments are independent of the outcomes. This raises the possibility that \( P(Y=y \mid W=w) \neq P(Y=y) \) for some \( y \) in \( Y \) and some \( w \) in \( W \), and a randomization test may be inappropriate.

There are two possibilities that we can consider for observational studies, one is a study that is free from hidden bias, so \( P(Y=y \mid W=w \& X=x) = P(Y=y \mid X=x) \) where \( X \) represents a matrix of controls and the outcome is independent of the assignment, conditioned on a set of observed control variables (Rubin 1977, Rosenbaum 1984, 2002). The second case arises when hidden biases are also present, so that \( P(Y=y \mid W=w \& X=x) \neq P(Y=y \mid X=x) \) but the condition \( P(Y=y \mid W=w, X=x, Z=z) = P(Y=y \mid X=x, Z=z) \) where \( Z \) represents a vector of unobserved control variables (Rosenbaum 2002). We will consider methods for studies free of hidden biases first. This type of condition has also been called treatment assignment on the basis of a covariate.

For observational studies in which the treatment assignment depends only on the covariate vector, we could stratify on the \( X \) variables to estimate the treatment effects within each stratum. That is, once the data are collected the units could be stratified so as to be homogenous on the \( X \) variables (Rosenbaum 2002). As treatment assignment depends only on the \( X \) variables any differences in the actual treatment received would be by chance alone and randomization tests could be used to analyze the results.

There are a number of complications that might arise with this approach, particularly for social science research which is likely to have a high dimensional covariate matrix with possibly continuous variable values. With many covariates that have many possible values it
becomes difficult to have a sufficient number of observations to stratify the data for analysis. An alternative to stratifying directly on the covariates when there are many possible vector values is to stratify on estimated propensity scores. The propensity score can be defined as the probability that a unit is exposed to a treatment condition, so that with two values, \( P(w|x) = \prod e(x)^w \{1 - e(x)\}^{1-w} \) where \( e(x) = P(w = 1) \) for every unit (Rosenbaum and Rubin 1983a). The propensity scores can then be used to stratify units with similar probabilities of treatment exposure for randomization inference. Estimated propensity scores for stratification can be obtained from methods such as logit or probit models (Rosenbaum and Rubin 1985a).

A special type of stratification is matching. Matching typically restricts the number of units from each treatment conditions in the strata. Pair matching is simplest method and matches each unit exposed to one treatment condition to a unit exposed to the other treatment condition. There are alternative ways of conducting matching such as many to one matching, one to many matching, and full matching (Rosenbaum 2002). Depending on the configuration of the data, it might be sensible to match one observation from a treatment condition to multiple observations from the other treatment condition. Full matching is a more sophisticated algorithm for selecting the matched units and allows one to many, many to one, and many to many matching.

There are also several ways to perform the matching. While there are a number of different algorithms available, there are several common options. One procedure is nearest-neighbor matching (Rosenbaum and Rubin 1985a). This procedure works by sequentially matching each individual treatment observation to the nearest available control. Once matched, the observations are set aside and the procedure continues for the remaining units in the data. One potential shortcoming of this approach is that it can lead to suboptimal
matched samples. Optimal matching minimizes both the distances between the units taken in sequence as well as the distance in the full sample (Rosenbaum 2002). Treatment cases can also be excluded from the analysis if there are no comparable observations in the data, although some research has found that severe biases can result from incomplete matching (Rosenbaum and Rubin 1985b).

The methods just described depend on observing all relevant control variables so that there are no remaining imbalances between the treatment groups (at least for factors related to the outcomes of interest). In many applied cases, there may be remaining imbalances on relevant control variables. Since there are many variables that might affect an outcome of interest it is unlikely that we will have data on an exhaustive set of variables, without which we cannot directly control for all relevant factors (Rosenbaum and Rubin 1983a, b). We can however assess the sensitivity of the results to different levels of hidden bias.

Sensitivity analyses for hidden biases were first presented by Cornfield et al. (1959). They examined the relationship between smoking and lung cancer and whether the apparent relationship might be confounded by an unobserved hormone secreted by smokers that also raised their risk of lung cancer. Cornfield et al. argue that the contemporaneous results suggest that smokers are nine times more likely to develop lung cancer than non-smokers. They then derive the result that in order for an unidentified hormone to account for the apparent relationship it would have to be at least nine time more prevalent in smokers (even if it were a perfect predictor of lung cancer). Since only an extraordinarily high level of bias could affect the conclusions, they argue that we can be confident that smoking is causally related to lung cancer.
The contribution made by Cornfield et al. is that they replace a qualitative statement that is true for all observational studies (that we cannot rule out hidden biases) with a quantitative statement that is true for a particular study (the amount of bias that would be required to change the conclusion, Rosenbaum 2002). If only high levels of hidden bias would affect the conclusion then we can be more confident in the results. If more plausible levels of hidden bias could affect the results we might pursue additional data collection efforts to address as of yet unobserved variables before reaching definitive conclusions. Methods similar to Cornfield's have since been generalized for other analyses (Rosenbaum and Rubin 1983b), and a dual-simultaneous sensitivity analysis that considers two parameters is reported in the chapter on vote centers.

Sensitivity analyses are a useful means of providing information as to how different levels of biases might affect the results but do not provide any information as to whether hidden biases are likely to be present. At least two other approaches might be used to detect the presence of hidden biases, which are methods of detecting hidden biases and matching with thick description. Methods of detecting hidden biases could include comparing different groups with a common treatment or consideration of a separate outcome variable that would not likely be affected by the treatment. These other analyses might be a useful means of detecting hidden biases, as hidden bias might be present if the groups differ on some other outcome variable that is not thought to be related to the treatments (Rosenbaum 2002).

Matching using thick description might also be a useful way of addressing hidden biases. This approach first matches units similar to the procedures described above. Qualitative information on the units is then compared to assess the quality of the match.
(Rosenbaum and Silber 2001). By incorporating qualitative information the matches could then be redone to obtain more comparable treatment groups for analysis. This approach might not only allow for the inclusion of information regarding currently unmeasured variables but also incorporates information on the units that is unmeasurable (even in future research).

2.5 Conclusion

This chapter reviews the general justification for the methods used in subsequent chapters to determine the causal effect of different election reforms from observational data. More specific details are included in the following chapters. In one way, these methods are an alternative to traditional regression models used in most extant political science research. Recent methodological research, however, argues that matching combined with regression adjustments is superior to using either approach alone. The argument is that regression models are a useful means of adjusting for any remaining observed imbalances in the data, particularly when used with propensity score matching (Rubin 1979, Ho et. al. 2007). By first matching on the covariates we also obtain results that are closer to experimental benchmarks and reduce the dependence of the results on ad hoc modeling assumptions (Rubin 1979, Dehejia and Wahba 1999, Hill, Reiter, and Zanutto 2004, Ho et al. 2007). For several of the analyses I report results both from randomization tests following matching as well as statistical model estimates.

In addition to the reasons just cited, there are several other potential advantages offered by matching over traditional regression-based adjustments. In particular, the theoretical underpinnings of matching seem better developed for causal studies than is true
for regression modeling. The context in which regression models can be treated as estimates of causal effects is not clear from the extant descriptions of regression models (Pearl 2000). For example, a regression equation of X on Y can be algebraically manipulated to produce a regression of Y on X, which does not seem to capture many causal relationships that scholars are interested in studying. There are also other objections such as problems of model-hunting and covariate selection which are additional shortcomings of regression models for causal studies of observational data.

The rationale for causal inference based on regression adjustments seems to rely on population models of a data generating process (King 1991, Morgan and Winship 2007). That is, the specified regression equation is assumed to represent the data generating process for a particular population. In this way, the unknown parameters of regression models might be estimated from the available data to make inferences about the actual data generating process. The problem with this approach is that it seems to require a model of the full data generating process, which in practice will typically be unknown.

A model-based alternative to data generating process arguments has recently been advocated by Pearl (1995, 2000). Pearl develops a graphical approach to causality and provides model estimation procedures that he argues allow researchers to identify causal effects. Pearl's graphical causal models have been argued to be similar to the potential outcomes model of causality, but the methods of inference advocated by Pearl differ from other studies that rely on potential outcomes. In particular, Pearl's approach uses graphical models and estimation procedures to block all paths and back-door paths between confounders and an outcome of interest. By blocking the necessary paths (by conditioning on some set of confounders) we can estimate the causal effect of a treatment variable (as
defined by his *do* operator in the presentation of causal graphs). While I do not directly use Pearl’s methods, I do draw from his consideration of blocking paths and back-door paths between the set of possible confounders to diagnose possible problems and identify ways to possibly improve the existing designs.

A potentially important assumption for the methods used in this project, which is associated with the potential outcomes framework, is the stable unit treatment value assumption (SUTVA). This assumption implies that there is no interference between the units so the treatments received by all of the other units have no effect on a particular unit’s response (Cox 1958, Rubin 1977). An example of a possible violation of SUTVA might be a study of agricultural plots, where a new type of fertilizer might seep into an adjacent control plot and thereby affect the yields of other plots. This assumption might also be problematic for the study of social behaviors such as voter turnout where there might be contextual effects.

As discussed in the chapter on voter registration and the conclusion, SUTVA might not apply to the study of voting, as an individual voter might influence the behavior of those around them. In particular, individuals that are voters might increase the likelihood that those around them vote. It is possible that there could be a suppressive effect on turnout, particularly if individuals have opposing preferences and so engaged in matched abstentions, but regardless of whether the overall effect is positive or negative, social interactions between voters raises the possibility that participation might not be independent of other individuals around the voter.

---

4 This assumption is not required for randomization tests of the no effect hypothesis, but is required for other methods such as the construction of confidence intervals and substantive effects.
While randomization is not feasible we can use the logic of randomized experiments to better design this study. From these considerations the following chapters report results from propensity score matching analyses, a generalized matching procedure for dose-response studies (polychotomous treatment), formal sensitivity analyses, and synthetic control methods for aggregate data, among others. Aside from strategies for analyzing data, this project is also careful to attempt to identify the theoretical mechanisms linking the treatment to the outcome, and testing the full implications of those processes as a way of bolstering our confidence in the results. Apart from methodological tools, this approach has been argued to be one of the most powerful tools for the analysis of observational data (King 1991, Rosenbaum 2002).
Chapter 3:
Registration Deadlines and Election Day Registration

3.1 Introduction

This chapter analyzes the effects of registration deadlines and Election Day registration on voter turnout. Voter registration has been argued to be an important factor for electoral participation, potentially accounting for lower turnout in the U.S. than in other countries, and might also contribute to the electorate being skewed towards individuals with higher incomes and more education (Powell 1986). This chapter makes three contributions to the existing research on voter registration. The first is to offer a resolution to some remaining theoretical critiques of the relationship between registration policies and electoral participation. Specifically, the extant research argues that registration policies lower the costs of voting and thereby increase participation, yet the decision-theoretical models on which the arguments are based are inconsistent with the empirical record. This chapter integrates recent research on turnout into the theoretical discussion to explicitly address these criticisms. Second, new sources of data and statistical methods are combined to provide more reliable estimates of the causal effect of registration policies. Third, the mechanism by which the effect is thought to operate is also tested empirically, which suggests a novel empirical finding that while registration deadlines seem to impact turnout the effect seems to operate in a way that is not fully understood at present.

There are a number of aspects of voter registration that have recently been debated as ways of improving elections. The Help America Vote Act (HAVA) requires states to establish a statewide database of registered voters to improve the maintenance of voter
registration lists and reduce the number of individuals erroneously kept on registration lists and to prevent accidental purges of current voters. The National Voter Registration Act (NVRA) was a precursor to HAVA. NVRA requires that states accept mail-in registration with a common form, and also make registration forms available at departments of motor vehicles and public assistance offices. These provisions of the NVRA are intended to make voter registration more convenient and thereby increase the number of people registered to vote. While these provisions have passed at the national level, other aspects of voter registration policies can also be amended which might have consequences for participation.

The primary focus of this chapter is to assess the effects of registration deadlines and Election Day registration (EDR) on voter registration rates and turnout. While other aspects of registration policies are also important to examine, registration deadlines and same-day registration are two aspects of registration policies that have most commonly been identified in scholarly and more applied work as factors that might be the most likely to produce changes in individuals electoral behavior. The question also seems timely as Montana just implemented EDR for the first time in 2006, and Iowa will be the 8th state to implement EDR in 2008. Other states such as California, New York, and Massachusetts have also considered adopting Election Day registration. Building from previous research, this part of the broader study of electoral institutions hopes to better identify the effects of registration deadlines and EDR on rates of participation and how the possible institutional effects might operate. The remainder of the chapter is organized so that the second section briefly reviews the related literature, and attempts to provide a theoretical explanation for how registration policies affect participation. Three empirical hypotheses are derived from this discussion. These hypotheses are then evaluated empirically in the third and fourth sections of the chapter. The
third section discusses the research design for the first two hypotheses, and the fourth section evaluates the third hypothesis. The fifth section of the chapter reviews the main findings.

3.2 Registration laws and voter turnout

3.2.1 Literature Review

The effects of voter registration laws on turnout in the US have been previously studied, and scholars have examined a number of different questions with some variation in conclusions regarding the effects of registration policies. The two main aspects of registration policies that have been examined are closing dates and Election Day registration (EDR), which are related but not identical. Other questions such as the effects of mail-in registration, the availability of registration at departments of motor vehicles and other offices, and registration purges have also been examined, but to a lesser extent. Earlier research on registration laws argues that voter registration is a barrier for voter turnout in the US (Wolfinger and Rosenstone 1980; Squire, Wolfinger, and Glass 1987; Teixeira 1992; Patterson 2002). The estimated effect of voter registration laws was around a 10% decline in turnout (Squire, Wolfinger, and Glass 1987; Teixeira 1992). The effects were stronger when focusing on certain populations such as recent movers and younger individuals (Squire, Wolfinger, and Glass 1987). While the existing research almost uniformly argues that relaxed voter registration requirements such as EDR and shorter registration deadlines increases participation, there are disagreements in the literature as to the magnitude of the effects, and whether or not the effects vary across subpopulations.

A number of researchers have specifically examined the effects of EDR on voter turnout. The studies have found a positive impact of EDR on turnout, but the magnitude of
the effects varies across studies. Fenster (1994) provides a recent analysis of the first wave of states to adopt EDR, which includes Minnesota, Wisconsin, and Maine, which adopted EDR in the 1974 for Minnesota and Maine, and 1976 for Wisconsin. Analyzing state-level election returns from 1960 to 1992, Fenster finds that overall turnout in non-EDR states declined by about 1%, while turnout in EDR states increased by about 4%, for a net effect of 5%. One potentially unanticipated result of Fenster's analysis (that has not gained much attention) is the conclusion that “day of registration voting is not a transient effect, nor is it cumulative—it leads to a one-shot increase in turnout” (Fenster 1994). Assuming that EDR affects new registrants (rather than existing registrants) we should expect the effect of EDR to manifest itself over time as the population gradually turns over (due to mobility and mortality) and new individuals enter the pool of potential registrants.

These findings are largely confirmed in subsequent study of 2nd wave EDR adoptees, which finds that EDR increases turnout by 3-6% (Knack 2001). This study also finds that the effects tend to be greater in midterm than presidential elections. Knack and White (2000) also examine whether EDR has a uniform effect on individual voters. Examining data from the Census Bureau’s Current Population Study Voter Supplement from 1990 to 1996, they find that younger and more mobile voters are more likely to be affected by EDR, suggesting that EDR might provide a way to reduce the skewness in the American electorate towards individuals with higher SES, and attract younger individuals to the polls. Highton (1997) also considers whether or not there are subpopulation effects of registration policies by comparing the skewness of the electorate in states with no registration requirements (North Dakota) and those with EDR to the rest of the states. Using CPS data from 1980 to 1992,
Highton finds that relaxed voter registration requirements reduces the skewness in the electorate, but does not eliminate it altogether.

In addition to studies of EDR, a number of researchers have also examined the effects of closing deadlines. Wolfinger and Rosenstone argue that eliminating registration deadlines should increase turnout overall by about 9%, and be particularly efficacious for individuals with less education. Subsequent analyses have found evidence of both larger general effects, 14% (Rhine 1996) and somewhat smaller effects, 3% (Rhine 1995). The subpopulation effects of registration deadlines for the less educated have also been disputed. Nagler (1991) argues that Wolfinger and Rosenstone’s result is a function of model specification, and that including a specific interaction term suggests that relaxed registration laws are more efficacious for the more educated. These results are not immediately clear however, as Wolfinger and Rosenstone report substantive effects on individuals’ probability of voting while Nagler reports coefficients without the associated substantive effects.

Expanding on previous studies, Mitchell and Wlezien (1995) consider not only the effects of registration deadlines on turnout, but also how they affect registration. This is an important consideration, as registration deadlines are thought to affect participation through rates of registration. Using CPS data from 1972 to 1982, they find that removing registration deadlines would increase registration rates by 8.2% and turnout by 7.2%. While these findings generally seem consistent with theoretical expectations, the effects might seem closer to one another than we would expect. An increase in turnout of 7.2% derived from an increase of registration of 8.2% seems as though it might be high, particularly when including off-year elections and that new registrants are likely to be younger and more
residentially mobile than older residents. We will return to this issue of the effects of registration deadlines on rates of turnout and registration in the sections below.

While registration deadlines and EDR are related to one another insofar as EDR involves the removal of registration deadlines, they have also been argued to have separate effects. Not only is there more variation in registration deadlines than states with EDR and those without it, but EDR is also thought to have an additional effect on turnout. By allowing individuals to register at polling locations, EDR is thought to further reduce the costs of voting and thus have an additional effect on rates of participation in addition to the removal of registration deadlines (Brians and Grofman 2001).

More recent research has critiqued the research designs used in earlier studies (Ansolabehere and Konisky 2006). Ansolabehere and Konisky argue that previous research has not adequately adjusted for confounding variables. More specifically, the lack of adjustment for alternative explanations of turnout has led to a substantial overestimation of the effects of registration laws. Using data from elections in Ohio and New York when those states first adopted registration requirements for voters in 1965 and 1977, they find that the effect of registration requirements is 3-5% on voter turnout, about half as much as previous estimates of the effects of registration requirements.

This chapter seeks to build from previous research on registration requirements in several ways. While Ansolabehere and Konisky (2006) offer an important critique of existing research on registration policies, it is also difficult to fully address confounding variables as experimental designs are not feasible to examine these questions. Other designs are limited in terms of spatial and temporal domains, and while informative, may not provide definitive answers when considered alone. To further address these issues, and provide
additional evidence to help us better identify the specific effects of registration policies, this chapter aims to report findings from newly collected aggregate data with more cases than has been considered before and for a more recent time interval. The design also examines a more detailed set of hypotheses, which has been argued to be one of the most powerful tools to analyze data from observational studies for causal inference (King 1991, Rosenbaum 2002). The remainder of the chapter proceeds by first making several definitional and conceptual issues more explicit, then describes the research design for closing deadlines and the results. The chapter then closes with a brief review of the findings.

3.2.2 Definitions

This section provides a conceptual definition of the variables discussed in this chapter. The definitions are the basis for both the theoretical discussion and operationalization of the variables, but there are not any particularly difficult issues with the definitions. The actual research design considers a number of other potentially confounding factors and their operational definition is provided below.

**Definition: Registration Deadline** In this study, the registration deadline refers to the number of consecutive days (excluding weekends and holidays when the registration offices are closed) on or before Election Day, in which a state or county's voter registration books are closed to new registrants to be eligible to vote in an upcoming election. The number of consecutive days is relevant as there can be multiple registration offices, none of which are continually open or necessarily for the same days. Satellite registration sites can be open on different days and may have different hours. Furthermore, states with Election Day registration might also have a registration "deadline" prior to an election. For example,
Minnesota allows registration on Election Day, but not within 21 days before the election. By this definition Minnesota would have no registration deadline, as the closing period is not over consecutive days prior to an election.

This provides a tentative definition of registration closing periods. This definition is mainly intended to provide a starting point for the discussion as to how registration laws might affect turnout, and there are certain issues that arise in the measurement of closing periods, which is discussed in more detail below. The closing period is the primary independent variable of interest in this chapter. The main dependent variable of interest is voter turnout, which will simply be taken to be the proportion of eligible voters that participate in an election. There are also particular issues that arise when measuring voter participation, which will also be addressed in the research design section of the paper (McDonald and Popkin 2002).

**Definition: Election Day Registration** Election Day registration refers to a policy that allows individuals to register on Election Day at their polling sites.

This general definition applies to all states that allow for registration at polling locations on Election Day. There might be particular variation in the requirement of each state for EDR, such as identification requirements, residency requirements, and the polling locations themselves might vary such as requiring EDR polling at the county courthouse. While each of these different policies might be an important topic to study as EDR can vary is how it is specifically administered, this definition is most consistent with existing studies of EDR and its implementation. One aspect of this definition is that Oregon would be excluded up to 1985, when they allowed same-day registration, but at a different location than where one could vote, but this does not affect the analyses which begin in 1992.
3.2.3 Registration Deadlines and Participation

Building from these definitions, we can consider how these policies might be related to voter participation. Starting with closing periods, the first way in which they might affect participation is by restricting the opportunities that individuals have to be eligible to vote in an election. By reducing the amount of time in which individuals can register to be eligible to vote, a longer closing period would be associated with less participation. The rationale is straightforward: if a jurisdiction experiences 10 new registrants a day, then increasing the deadline by 5 days would lead to 50 fewer eligible voters in that jurisdiction. While it is likely that individuals would be more likely to register close to Election Day, we can see that even assuming registration to vote at a constant rate, independent of the timing of campaigns and elections, we would expect there to be lower turnout with a longer closing period. The effect might be analogous to a store that closed a number of days before the holidays. While it is likely that they would lose a certain amount of revenue by being closed as they would any time of year, they are also missing a particularly busy season.

While the effect of the registration closing period can be sustained with a constant voter registration rate, we can also consider whether or not the effect would be greater, if voters are more likely to register as an election approaches. There may be multiple reasons to anticipate that individuals are more likely to register to vote as an election nears.

Influence of Campaigns One reason that more individuals might be interested in registering to vote closer to an election is that campaign activity tends to intensify as an election nears. Political campaigns and advertising is heavier on days closer to an election than those farther away. For a distant election, individuals are unlikely to be attentive to the campaigns, and
those that are attentive are more likely to be partisan and less reliant on specific campaign activities to make their vote choice. Rather, less attentive and less partisan voters might be more likely to be influenced by campaign activities and more attentive to politics as the election nears. This might create greater interest in registering to vote for otherwise eligible individuals close to an election rather than for an election that is farther away.

**Influence of Media Coverage** Not only might political campaigns increase interest in an election and possibly stimulate voter registration nearer to an election, but media coverage of the races also tends to increase as the election draws near. Greater media attention might also help to stimulate greater public interest in the election and voter registration for those that seek to participate (Highton 2004).

**Costs and Benefits of Participation** A third explanation that builds from recent research on turnout suggests that the extent to which individuals’ discount future benefits relative to present costs is related to voter participation. Fowler and Kam (2006) argue that individuals vary in the extent to which they discount future benefits relative to present costs, and that individuals that discount future benefits to a greater extent are less likely to participate in elections. In their paper, Fowler and Kam examine how this personality trait might lead individuals to be more or less likely to vote. Specifically, they use a laboratory experiment to obtain a measure of individuals' willingness to trade off costs and benefits over time.\(^5\) To obtain the measure, Fowler and Cam offered subjects a choice of payoffs they could receive

---

\(^5\) This is not entirely correct. In their experiment, Fowler and Cam do not explicitly include a cost to the subjects, but rather present them will a choice of payments they can receive at different points in the future. In this way, the experiment considers the subjects' responses to monetary gains but does not explicitly examine costs. While this might be taken as a limitation of the study, there is also not clear reason to expect the subjects to respond differently to a short term cost for a future benefit than different levels of gains to be realized at different points in time.
at different times. Subjects were asked to make 20 selections between receiving a payment after 30 days, and versus a greater payment after 60 days. For example, subjects could receive 100 dollars 30 days after the experiment or choose to wait 60 days and receive a higher payment (ranging from 100.17 to 123.07 with 20 categories). After obtaining this measure, they then examined whether or not the subjects voted in the previous statewide primary election (California primary, March 2004). Consistent with their expectations, individuals that were more willing to delay realization of the benefits were also more likely to vote. There is an issue with the measure of individuals' willingness to delay benefits, as the measure is obtained after the election so that it cannot specifically affect individuals previous choice to vote or not. Yet as a measure of a personality trait, the responses are likely to be enduring over longer time spans, so that while it does not a measure of individuals' disposition prior to the election it is likely to be highly correlated.

Fowler and Kam expect this personality trait to affect individuals' likelihood of turning out to vote, as the benefits of voting are realized over time. In this way, the value that individuals attach to future circumstances is an important component of the model as the costs of voting are incurred immediately while the benefits accrue over time. Even though the benefits might outweigh the costs of voting, individuals weight the costs more heavily due to time discounting and might thereby be less likely to vote.

While Fowler and Kam argue that a willingness to delay benefits is specifically related to voter turnout, the argument can also be extended to explain how registration policies affect turnout. The argument is that benefits of participation are realized over time, after the individual has voted. The costs of voting in the short term are taken to be the time and inconvenience of voting. However, voting in the US is a two step process. Individuals
must first be registered in order to vote. Registration, in turn, imposes an additional cost on prospective voters prior to casting a ballot. The restrictiveness of registration policies might therefore effect participation, by increasing the temporal distance between realizing the costs of registration and the benefits of participation. An individual does not need to register for each election, but rather the registration will generally carry over (except in cases where the individual moves or if they have not voted in a sequence of elections they may be removed from the registration book). In this way, individuals might be able to amortize the costs of voting over a number of elections, but each subsequent election will occur farther in the future and is thereby further discounted.

By extending the argument that an individuals' willingness to delay benefits is related to participation, we can see that closing periods might also be related to participation. As individuals discount future benefits of participation in a way that places less weight on benefits that are more distant in the future, restrictions that require individuals to register at earlier dates when the election is farther away might lead to a decrease in participation. While an individual might be willing to bear the short term costs of registration if she can realize the benefits of participation within a week, she might be less willing to bear the costs of registration if the benefits could only be realized in a month. This might also explain why individuals would be more likely to register in the days nearer to an election, as the benefits of participation are discounted less nearer to the election. By considering campaign activities, media coverage, and individuals' personality traits, we might expect the effects of registration restrictions close to an election to decrease voter turnout. From these arguments we can formulate the following hypotheses.

H1: Shorter closing periods should lead to higher levels of turnout.
The means by which registration policies are thought to effect turnout is by influencing rates of registration. We can thus test a second hypothesis that is thought to be directly related to the first.

**H2:** Shorter closing periods should lead to higher levels of registration.

From these arguments, we might expect that more restrictive registration policies in terms of having earlier closing dates, leads to lower voter registration rates and subsequently less turnout. By reducing the opportunities that individuals have to register prior to an election, particularly at a time when interest in registration is likely to be highest, earlier closing dates may lead to lower voter registration, and with fewer individuals eligible to vote, we expect less overall turnout.

It might also be important to note that early voting does not complicate these issues. Early voting is the focus of the next chapter, but it might be thought to have a role in the effects of closing dates as well. This is not the case, however, as registration deadlines apply to early voting as well. That is, in order to be eligible for early voting, an individual must typically be registered the requisite number of days prior to the start of early voting. An individual can also register at a later date and be eligible to vote on Election Day, but not for early voting. As early voting is currently implemented, the arguments in this section would apply to it just as they apply to Election Day voting and thus does not create any immediate complications for this study.

An alternative view of voter registration is that the voters themselves might be responsive to variation in registration deadlines. Instead of registering at a constant rate or more proximate to an election, voters might instead account for registration deadlines and adjust their behavior accordingly. From this perspective, voters are able to account for
deadlines so that there might be an increase in applicants proximate to the deadline, those that are interested in voting will be able to reliably meet the registration deadline be it 30 days or 10 days before Election Day. While this is likely true for some voters, particularly political sophisticates, it is unlikely to apply to all voters, as many seem unaware of their state’s registration policies. For example, according to the 2006 Current Population Survey almost 10% of non-registered voters in Election Day registration states cited missing the deadline as their reason for not being registered. This could introduce some individual heterogeneity into the effects of registration deadlines, but this possibility is unlikely to eliminate the overall effect of registration deadlines and the extent to which certain groups of voters are more or less responsive to registration deadlines would be an interesting area for future research.

3.2.4 Election Day Registration

While registration closing dates might affect levels of participation, Election Day registration might have an even greater effect on participation. The possible effects of EDR might be thought to operate through the absence of a closing date. By allowing individuals to register to vote at the same place and time that they can vote, there is no period in which registration is closed. In some implementations of EDR, the registration books will be closed to new entries immediately prior to an election but individuals are allowed to register on Election Day at polling sites. In this way, EDR does not involve a closing period as the number of consecutive days on or before Election Day in which registration is closed is zero.

As EDR removes registration closing dates, it might be expected to lead to greater electoral participation for the reasons discussed above. However, there might also be added
effect of EDR, as it also lowers the costs of registration. Not only do individuals have a longer time in which to register, particularly at a time when campaign activity and media coverage of elections will be highest, but it is also easier for individuals to register as they can go to their polling site on Election Day and register to vote. Otherwise individuals may have to obtain, complete, and submit a registration form outside of a polling place in advance of an election. The added convenience of EDR, by allowing individuals to register at the same place that they are eligible to vote, should have a greater effect on registration rates than just through the elimination of the closing date alone (Brians and Grofman 2001). By increasing the number of individuals that are eligible to vote, EDR should thereby increase overall rates of voter turnout. This suggests the following hypothesis.

**Hypothesis 3:** Election Day registration leads to higher rates of turnout.

This section provides an explanation for how more or less restrictive voter registration policies can affect levels of participation. The section begins by providing a definition of the relevant aspects of registration policies and then explains how they are connected to individuals' behavior. From these considerations three empirical hypotheses were derived. The main hypotheses are numbers one and three, although hypothesis two will also be important for examining the process by which deadlines affect turnout. The following section describes a research design that is used to evaluate these hypotheses.

### 3.3 Analysis of Registration Deadlines

This section provides a research design in two parts to evaluate the above hypotheses. The first part examines the hypotheses regarding the effects of registration deadlines, and the second looks at the effects of EDR. There are two sources of data used in the analyses;
aggregate elections data from U.S. counties for the main results, and supplementary results from individual level data from the Current Population Study.

The section begins by describing the aggregate data including spatial and temporal domains, how the variables are measured, and the data sources. Using these data, descriptive summaries are reported to show the distribution of the raw data. The descriptive results also complement the inferential analyses by displaying the actual data and showing patterns that help to motivate the choice of statistical models used in the formal analysis.

The second part of this section considers the effects of EDR. The preliminary analyses use the aggregate data to assess the effects of EDR. Along with these results, a recently developed method of conducting comparative case studies is applied to EDR. This approach briefly compares states with EDR to a synthetic control state constructed from the pool of 43 potential control states without EDR. The synthetic control states blend characteristics of the natural controls to create cases that more closely approximate the characteristics of EDR states than any existing natural comparison case alone.

3.3.1 Effects of Deadlines: Evidence from the Counties

The primary source of data is observations at the county level. The data span all regular general elections from 1992 to 2004 when there was at least one statewide-office on the ballot. The spatial domain is all counties (or other local governments that have jurisdiction over the administration of elections, such as some townships in Virginia) in all states, except for Alaska and Washington D.C.⁶ These data were primarily gathered from official sources that were either available from Secretaries' of State websites or by request. When these

⁶ The choice to exclude Alaska was made because elections in Alaska are administered at the borough level, which is not clearly comparable to jurisdictions in other states. Washington D.C. is excluded as it does not have comparable statewide elected offices such as governor and senator.
sources were insufficient, other data were coded from the America Votes series, published biennially by CQ Press. In total, there are 3,146 county or other local political subdivisions in the data, spanning 7 elections from 1992 to 2004, in 49 states. When combined into a time-series cross-sectional data set, there are 22,022 records in the data. The two following tables show the breakdown of the data for the voter turnout and registration data. Of the 22,022 records in the data, voter turnout data was obtained for 98.8% of the possible cases, and voter registration data was obtained for 90.3% of the cases. The specific numbers are broken down by state and year in the tables. Missing data are represented as N/A. The only missing observations for voter turnout are for elections in which there were no statewide races on the ballot. There are more cases with missing voter registration figures due mainly to data availability and the exclusion of North Dakota and Wisconsin. These two states are excluded from the registration data as North Dakota does not have a registration requirement and Wisconsin does not require registration for residents in counties with less than 5,000 people. Voter turnout and registration figures from the counties are used to calculate the dependent variables in this study as discussed in the next subsection.

**Number of Voters** For the number of votes cast, some governments report the total number of ballots cast, while others report total turnout for the top ticket race. Whenever the total number of ballots cast was available, that number is used for the measure of total turnout. When the total number of ballots cast was not available, turnout for the top ticket race was used. For presidential election years, this number was taken to be the votes for president. For mid-term elections, the choice was most often between a senatorial and gubernatorial race. Turnout for these two races was typically comparable, and when both were reported, the race with higher turnout was used. When the total number of ballots cast was available
along with vote totals from these races, roll-off rates for the top-ticket items were generally not substantial.

To calculate the rate of voter turnout, there are several choices that could be made for the denominator. Popular reporting seems to most commonly use the ratio of actual voters over the total number of registered voters. There are several possible short-comings to this measure, however. One is that the registration numbers are used to measure registration rates for the second analyses, and also does not really reflect the number of individuals that were eligible to participate in the election as there might also be deadwood wherein individuals that have moved or are deceased are still in the registration book. In other instances, voter registration is not available, such as counties in North Dakota, which does not require voter registration, as well as for small (<5,000 population) municipalities in Wisconsin (although this law was changed effective January 1, 2006 to require registration of all residents in Wisconsin).

**Voting Age Population** A more common alternative for research studies is to use voting age population (VAP) as the denominator. These data are available from the US Census Bureau, but have also been criticized as providing a misleading measure of turnout, as not all individuals over 18 are eligible to vote (such as ex-felons and non-citizens). Not only might the measure not exactly reflect the number of eligible voters, but has been shown by McDonald and Popkin (2001) to affect conclusions regarding overall trends in turnout rates. They argue that a better option is to use the voting eligible population (VEP), which is the voting age population adjusted for noncitizens, non-eligible felons, and other non-eligible institutionalized persons. While McDonald and Popkin estimate VEP at the state level, corresponding data for the counties is not available, as estimates of the ineligible felon
population is not available at the county-level. However, the largest group of non-eligible voters that accounts for most of the difference between VAP and VEP figures are non-citizens. While a county-level measure of VEP is not available, I collect data on both the VAP and the citizen-VAP (CVAP) to use as denominators in these analyses. Results obtained using the two measures are very similar, and the results reported below are those for the models using VAP.

To construct the measures of VAP and CVAP, data were taken from the US Census Bureau. The measure of VAP was taken from the intercensal population estimates of county VAP. The measure of CVAP was created by obtaining a count of the non-citizen populations in the counties. For this measure, the estimated proportion of non-citizens over 18 was linearly interpolated from 1990 and 2000 US Census data. The CVAP variable was then created by weighting VAP by the estimated proportion of the 18 and over population that were citizens for each county.

Registration counts Similar measures were also created for registration rates. The numerator in this variable is the number of registered voters. Voter registration numbers can be somewhat inaccurate if it contains individuals that have since moved away from the jurisdiction or that are recently deceased. In some instances, voters can also be erroneously removed from registration lists, either through a prolonged period of inactivity or perhaps through administrative errors, though higher registration numbers seem more common. In some counties, voter registration numbers exceed the voting age population. While this suggests the presence of deadweight in the registration book, these counties tend to be

---

7 Calculation of VEP at the county level would further be complicated by changing laws regarding the eligibility of felons and mentally-impaired to vote. While there is no federal requirement that individuals must be citizens in order to vote, every state has a citizenship requirement for voting.
relatively small, and it is possible for registration records to exceed VAP. For example, college students, military personnel, and overseas populations might maintain eligibility to vote in a jurisdiction while not necessarily being counted as residents for Census purposes. The Census counts themselves might also contain some amount of error. While these numbers might reflect inaccurate registration lists, it may also not always be the case.

Some states have sought to address the issue of voters remaining on the registration lists after they have moved or are deceased by creating an inactive voter list (also sometimes called a suspension list). The rules vary by state, but typically a voter will be placed on an inactive voter list after they have not participated in around two consecutive statewide elections. When present, the total number of active registered voters was recorded, and otherwise data on the total number of registered voters was recorded. While active voter registration lists is one way to handle registration deadweight, it is not likely to be perfect, and there is no other clear way to reduce measurement error in the registration data. Using the data on the number of registered voters, a measure of registration rates was created by taking the ratio of registered voters to VAP, and to CVAP. As before, the results were nearly identical between analyses using VAP and CVAP and the reported results are for the models using VAP. The conclusions regarding the effects of registration deadlines on either turnout or registration rates do not depend on which of these measures are used.

The main independent variable of interest is the closing deadline. Registration closing dates are typically set at the state level, except for North Dakota which does not require voter registration, and Wisconsin which does not require registration for municipalities with less than 5,000 residents. The registration deadline variable is recorded as the number of days between Election Day and the last day in which individuals could

---

8 At least not during our analysis period, as mentioned above.
register. These data are recorded from state codes obtained from Lexis-Nexis. Lexis-Nexis also indexes changes in state codes, so that any changes were also recorded between 1992 and 2004.

Some states have different closing dates for different modes of registration. For example, Washington state requires registration by mail 30 days prior to an election, but allows individuals to register in person up to 15 days prior to an election. In these instances (which are a small number of cases), the later date is coded (i.e. Washington’s deadline is recorded as 15 days). Illinois also has different closing dates for counties with more than 500,000 residents than for those with less than 500,000. The date for counties with more than 500,000 residents is recorded, and is always within one day of the closing date for the other counties throughout the analysis period. Some states also set a fixed number of days prior to an election, such as 30, while others describe the closing date, such as the fifth Monday before an election as in Georgia. For those states that did not provide an explicit date, the closing variable was calculated from the date of the election.

A histogram of closing dates across states is shown in Figure 3-1. From this figure we can see that the data ranges from a minimum of no closing date in North Dakota and those with EDR to a maximum of 30 days. The data is approximately bimodally distributed. The largest number of counties have closing dates around 30 days, and the smaller mode is located under 10 days with the others distributed in between. While some states amended their closing dates during this time, the distribution largely remains unchanged when separate histograms were created for each election.

3.3.2 Analyses
The first part of the analysis provides descriptive summaries of the data and the second offers inferential analyses. Descriptive summaries of these data are important for this study for at least two reasons. The first is that this is a new source of data and it is useful to present the raw data to show the relationships in the actual data between the relevant variables. The second is that statistical modeling in the social sciences has been criticized on a number of grounds. There are three particular criticisms that can be identified, which includes ambiguity about the selection of control variables (Pearl 2000), ambiguity about the selection of statistical models (Ho, Imai, King, and Stuart 2007), and model-hunting (Freedman 1991). These criticisms are related insofar as they all involve there not being any “natural” model from which to obtain results that can be used to support causal inferences. Rather, empirical results can be affected by researchers’ choice of statistical models which can leave open questions as to the extent to which there is strong empirical evidence regarding the relationship between particular variables. These descriptive summaries might provide a simple and useful way of addressing these criticisms. Insofar as patterns suggested by the empirical results can also be verified by relationships in the raw data, we can be more confident that the results are not just the product of particular modeling decisions.

Turnout Rates Using the data and measures described above, this section reports several descriptive summaries of the data to be more formally analyzed later in the chapter. The first series of graphs represent estimated turnout for each of the counties, shown in Figures A1-A6 in the Appendix A. Separate graphs are presented for each state, and show the

---

9 These summaries were created by using data on counties with non-missing outcomes, within the allowable range of values, and that had a voting age population of at least 1000. The 1000 VAP criteria was set, as possible measurement error in the interpolated population values are more likely to be influential for smaller counties, when slight deviations can lead to greater changes in turnout ratios. While the criteria of 1000 was somewhat arbitrary, it does reflect a general cut in the data, with more counties above and below, than immediately around the threshold, and only excludes the smallest 52 of 3,146 counties.
distribution of estimated turnout rates for each county in each state. The x-axes in these plots represent the intra-state order of level of voter turnout from low to high, and the y-axes represents estimated county turnout (as a ratio of VAP). The estimates are obtained from a hierarchical linear model, with a random county intercept as the only regressor. The county-intercepts were estimated using WinBUGS, with diffuse normal priors (mean of zero and a variance of 1,000) and a burn-in of 1000 iterations. The results were obtained from three chains with different starting values. Following the initial burn-in, a subsequent 500 posterior draws were recorded (thinned by a factor of 3) to assess convergence of the chains and summarize the results. From these figures, we can see that counties within the states have somewhat similar levels of turnout, though there is variation in levels of turnout both within and between the states.

**Participation and Registration Deadlines** To explore the relationship between registration deadlines and voter turnout, Figure 3-2 plots the data on turnout against the deadlines. The x-axis in this figure represents the number of days between the closing date and Election Day, and the y-axis is turnout as a ratio of VAP. To help spread the data, a small amount of random noise is added to the closing dates in this figure. The green points represent the counties from 1992 to 2004, and the blue line represents the least squares line for these data (estimated without the random noise). From this figure, we can see that the average rate of turnout across the closing dates ranges from a maximum average of about .6 to a minimum average of just below .5. This figure suggests that states with no closing date have just over 10% more of their eligible voters turnout than states with a 30 day registration deadline.

While more restrictive registration deadlines are associated with less voter participation in this analysis period, the relationship is not as strong for rates of registration.
Figure 3-3 shows registration rates across the closing intervals, and while the data show a decline, it is less pronounced than with voter turnout. From this figure, we can see that the least squares line shows an average of about 85% registration rates with no deadline and about a 78% registration rate with a 30 day deadline. While there is still a negative relationship between registration rates and registration deadlines, the difference is not as great as for voter turnout.

In addition to the raw data, Figure 3-4 plots the county-intercepts from the model of turnout against the closing date. This figure shows a similar relationship as the raw data, with not a significant amount of overlap in the estimated county intercepts across the registration deadlines. To further illustrate the uncertainty in the data, a scatterplot of state-level intercepts from a model of turnout is plotted against the deadlines variable in Figure 3-5. This figure includes not only the estimated state intercepts, but lines spanning one standard deviation above and below the estimates. From this figure, we can see that while the lines overlap to some extent, there is still an apparent negative approximately linear relationship between turnout and registration deadlines. An analysis of turnout across the years is also included in Figure 3-6.

Correlates of Turnout We can also consider correlations between variables in the data. Using the county-level data, Figure 3-7 shows pairwise correlations between variables. The variables in the correlation matrix are turnout and deadline as before, along with measures of median age, proportion with a college degree, per capita income, and the proportion of Anglos in the county. Looking at the pairwise correlations, we can see that the registration deadlines have a similar association with turnout as median age and size of the Anglo population, and a stronger association than measures for proportion with a college degree and
per capita income. The figure shows the absolute value of the correlation, and Table 3-3 confirms that the actual correlation is negative, as expected. From the table we can see that the deadlines have the strongest association, although the magnitude is very close to the age and Anglo variables. The other noteworthy result from these displays is the high correlation between the college and income variables, which we will later discuss with model specifications. Similar to the findings from above, the relationship appears to be weaker between closing deadlines and registration rates. Figure 3-8 shows the pairwise correlation matrix with registration rates. The close variable still shows a stronger association with registration rates than the college and income variables, but not as strong as the others. The numeric values are shown in Table 3-3, and while the association is still negative, as we might expect, the association is not as strong as is the case with voter turnout.

The analyses thus far provide a general overview of patterns in the raw data, prior to a more systematic examination. The data reflect a fairly consistent association between voter registration deadlines and turnout. The descriptive analyses show variation in states' registration deadlines, fairly consistent patterns in county-level turnout within states, a stronger correlation between registration deadlines and turnout than other explanatory variables, and a fairly clear negative trend in the graphical representation of the data. While there seems to be a clear association between registration deadlines and turnout, there was a less apparent association between registration deadlines and registration rates.

Having provided a discussion of the aggregate level county data along with descriptive summaries of the variables, the next step in the analysis is to provide more formal analyses of the data that will help with the evaluation of the empirical hypotheses. The first analyses use a subset of the data taken from cases in which the states changed their voter
registration laws. There were 7 states that liberalized their voter registration laws in the analysis time, so that we can assess whether providing voters with more opportunities in which to register increased voter turnout as suggested by the first hypothesis.

The results of the analysis are presented in Figure 3-9. These graphs represent the election year on the x-axis, and turnout rates on the y-axis. The turnout rates are mean differenced so that the vertical axis represents the difference between the state's turnout in that year less the average turnout rate of all the states in that year. The colors are used to separate pre and post-intervention periods, where the red bars represent elections in which there were more restrictive closing deadlines, and the green bars represent less restrictive closing deadlines. The specific change in the registration deadline is labeled across the horizontal axis. From these two figures, we can see that there is not a particularly stark pattern of change in the pre and post-intervention times. While some post-intervention elections show an increase over pre-intervention elections, there is also considerable variation within and between states. That is, there are cases in which post-intervention turnout was higher than for some pre-intervention elections, but was not universally the case, and there are also instances in which turnout was less in post-intervention elections.

We can consider not only whether turnout went up in these states which liberalized their voter registration laws but also whether or not states that moved towards more restrictive voter registration experienced a decline in turnout. Figure 3-10 shows the results from three states that moved towards more restrictive voter registration. As before, the axes represent the election year and differenced turnout rate respectively. In this figure, the graphs begin with the green bars then the red, indicating that change was to a more restrictive registration policy which we expect to be related to a decline in turnout. From this figure, we
can see that there is again no particularly distinct pattern of change in voter turnout following a change in the states' registration laws.

A third group of states changed their voter registration laws twice during this time (no state changed their policies more than three times). Two of these states liberalized their registration deadlines at two different times, while Missouri first adopted a more restrictive deadline in 1994, then made their deadline less restrictive in 1996. This figure (11) again shows mixed evidence of the effects of voter registration laws, as some states show an increase in some elections following relaxation of the deadline, and a decline in others. Overall the findings are somewhat mixed, although the general pattern is in the anticipated direction. Analyzing the data together shows that the correlation coefficient is -.57. The negative correlation is consistent with the hypothesis that closing deadlines are inversely related to turnout.

One of the shortcomings of the initial analyses is that there is no explicit control for potentially confounding variables. Furthermore, the correlation coefficient is negative, which suggests that the deadlines are inversely related to turnout, even if the pattern is not obvious from the graphs. The analysis might further be complicated if the effects of registration laws are thought to take effect over time as the population of voters turns over. That is, with permanent registration the effects of the policies might take effect over time as greater population turnover will increase the number of prospective registrants that might be affected by the policy. To address these issues, we can consider the full data which have policies that have been in place for a longer period of time and also allow for control of potentially confounding variables.
In order to address some of the shortfalls for the pre/post intervention results, we can analyze the full time-series, cross-sectional data. These data were used to construct the scatter plots from the previous section which show an approximately negative linear relationship between registration deadlines and rates of voter turnout. Using the observations of counties over time, we can analyze linear regression models to attempt to adjust for other potentially confounding variables and also to better quantify the effects of registration deadlines on turnout.

**Control Variables** One possible (and important) criticism of the first analysis is that there might be systematic differences across the observations other than changes in registration laws. To attempt to control for other potentially confounding variables, the measures of demographic characteristics of the counties are incorporated into the analysis. Specifically, measures of age, education, income, and racial composition are included, based on data from the US Census. From the census data, measures are created for the median age in the county, the percent of residents with a college degree, per capita income, and the proportion of Anglos in the county. These demographic measures are included in the regression models to adjust for possible imbalances in the data across counties.

**Bayesian Model** These data are then used to estimate a Bayesian normal linear model. Bayesian methods are more appropriate for these data than conventional approaches, as the data do not represent a sample from a population, but rather are closer to the actual population, at least over this time period (Gill 1999, 2002). While some cases such as Alaska and Washington D.C. are not included, a large majority of the cases are observed. With these
data, a Bayesian approach is more appropriate, although both Bayesian and non-Bayesian models return numerically identical results.

The selected Bayesian priors on the parameters are diffuse, which along with the large sample size, ensures that the results are driven by the data, rather than a choice of strong priors. While these results are numerically identical to more conventional non-Bayesian estimates, the technical interpretation of these results are more straightforward. That is, rather than interpreting the estimates relative to a sampling distribution (a probability statement about the data given the estimates) we can make a simpler (and more appealing) statement about the posterior probability distribution of the parameters. This aspect of Bayesian statistics is regularly incorporated in political science research using a popular quasi-Bayesian (though technically incorrect) technique in combination with more conventional likelihood estimates (King, Tomz, and Wittenberg 2000).

The Bayesian models are also easily extended to incorporate hierarchical structures, such as when the data are composed of observations grouped by counties (Western 1999, Gelman and Hill 2006). In these analyses, I report results from a hierarchical linear model as just described that incorporates a county-specific intercept. The model was represents an intermediate point of complexity, although the results are consistent with both simpler and more complex model specifications. The simplest specifications were linear regression models with a constant intercept and measures of closing dates and the control variables. The most complex model was a random intercept, random slope model (for the registration variable). The results reported here were obtained from MCMC estimates of the statistical model. Three separate chains with different starting values were allowed to run for 20,000
iterations, and values were recorded from a subsequent 3000 iterations, thinned by a factor of three.

The results from the analysis are presented in Figure 3-12 and Table 3-4. The main effect coefficients are shown in the table. Summary statistics are also provided for the random intercepts, but are not themselves listed in the table, as there are over 3000 estimates. From this we can see that the estimated linear effect of a change in registration policies can be to increase voter turnout by up to 10%. That is, moving from among the most restrictive deadlines of 30 days to Election Day registration is expected to increase voter turnout by 10%. The figure also shows 50% and 90% confidence intervals, which do not overlap with 0 suggesting a positive effect of shortening closing dates. This effect is also surprising close to other estimates of the effects of registration laws, as Squire et al. (1987) argue that eliminating voter registration might increase turnout by 9%. While Election Day registration removes some of the costs of registration, it is still more restrictive than requiring no voter registration, which might suggest that the effects of doing away with registration or implementing universal registration might be even stronger than anticipated.

Diagnostics To further check the robustness of these results, we can also simulate a "new" dataset using these results (Gelman and Hill 2006). Simulating a new data set will serve at least two important purposes. The first is that it allows us to evaluate the overall quality of the model. If the results of the simulated data closely approximate features of the actual data, we can be more confident that the empirical results are not just an artifact of modeling assumptions. Second, the simulated data can help us better interpret the empirical results, as it might allow us to make a more accurate prediction as to the likely effects of relaxing voter registration laws.
In order to simulate a data set using these results, random draws were made from a multivariate normal distribution centered over estimated mean, and using the estimated variance parameter (Gelman and Hill 2006). Using this distribution, 20,000 new data points were drawn, and this was repeated 5 times, to create 5 new simulated data sets. The results are shown in Figure 3-13. The first plot shows the distribution of the actual data over closing dates, and the second shows the plot of the simulated data over the closing dates. We can see in this figure that the simulated data appear to closely approximate the actual data.

3.3.3 Deadlines and Voter Registration in the Counties

The above results are all consistent with the hypothesis that voter registration deadlines are related to voter turnout. Yet the results cannot be taken as definitive, as the necessarily observational nature of the data would not warrant such strong conclusions without additional evidence. The central problem is that the analyses might be confounded by unobserved variables, so that the results might not actually be produced by an effect of registration deadlines on turnout. While the main concern is that we might erroneously conclude that registration deadlines are related to turnout when they are actually not, it is just as possible that we underestimate the effects of registration deadlines as a result of unobserved bias.

In order to strengthen the empirical evidence in this study, we can further consider the effects of registration deadlines on rates of registration. The argument presented here, and which is consistent with previous policy and scholarly research on the effects of registration deadlines, is that relaxing registration deadlines allows more individuals to register and thereby turnout to vote (e.g. Wolfinger and Rosenstone 1980, Mitchell and Wlezien 1995,
Brians and Grofman 2001, Brown and Wedeking 2006). This suggests a relationship between the two hypotheses about the effects of registration deadlines, so that the evidence regarding one hypothesis will help us to reach more reliable conclusions about the other. Testing the mechanism by which deadlines affect turnout is obviously of central theoretical interest as it directly addresses the research question as to how deadlines affect turnout. The additional test is also of particular importance given the necessity of analyzing observational data. Evidence that confirms both hypotheses is much stronger than evidence that confirms only one. That is, any one analysis using observational data may be subject to the criticism that it could be influenced by unobserved variables, but each related test that shows a similar pattern makes it increasingly difficult to attribute all of the results to unobserved variables. While we might be most directly interested in whether or not deadlines affect turnout, we can be more confident that we understand the nature of its effect when we can validate the mechanism by which deadlines are thought to affect turnout. For example, historical research on the relationship between smoking and lung cancer not only consider if smokers are more likely to get lung cancer than non-smokers, but people who smoke more are more likely to get lung cancer, if people that quit smoking reduce their risk of lung cancer, and other likely implications we would expect to see if there were an actual effect of smoking on lung cancer.

In this study we can assess whether or not registration deadlines affect the number of registered voters. Based on the arguments in this chapter and from previous research, registration deadlines are thought to affect turnout indirectly, as part of the two-step process. The intermediate step is that deadlines have a direct and negative effect on rates of registration, which in turn has a direct and positive relationship with turnout.
For this analysis, registration rates are modeled as a linear function of the treatment variable and the same covariates used in the model of voter turnout. A Bayesian model with diffuse normal priors was again specified and results are obtained using MCMC. The choice of models is more important in this analysis. The results presented below were not fully robust to all model specifications. In particular, a hierarchical model with county-intercepts returned results that had an opposite sign and half of the magnitude of these results. Several other models were estimated including a linear model with only the deadline variable, and a model with a fixed intercept term. The latter model is the source of the results presented here. Both this model, the linear model with only the deadline variable, various subsets of the included controls, and the raw data are all consistent with one another. In addition, the low magnitude of the estimates from the county-intercept model suggests that these reported results are more reliable estimates.

The results from the linear model with diffuse priors are reported in Table 3-5 and Figure 3-14. These results show that the effects of registration deadlines are negative, as anticipated, but the magnitude of the effects are smaller than anticipated. The estimated effect of the deadline variable is that it decreases registration by .006% for every additional day. In other words, the effects of having a 30-day deadline versus EDR is that EDR would increase rates of registration by about 3%. While the effect is in the expected direction, the magnitude contrasts with an estimated effect on turnout of 9% (also comparing a 30-day deadline to EDR), which is puzzling as an increase in the number of registered voters of some percent should increase the number of actual voters by some percentage less than that (or at most the same percentage if every new registrant also voted). There is no clear way to explain a greater effect on turnout than for registration.
We can also use the simulated data to evaluate the effects of registration laws. By rerunning the simulation setting closing dates equal to different values, we can estimate their likely effects. Figure 3-15 shows the results plotted over 5 different values of closing dates, which closely matches the results presented above, as we would expect. The results from the aggregate analysis and simulations provide further evidence of the likely effect of registration laws on voter turnout.

In some sense, both hypotheses 1 and 2 have been confirmed by these results, and drawn into question. Taken separately, the evidence suggests that registration deadlines can be relaxed so as to increase registration and turnout rates. Yet considered together, these two results seem to present an interesting puzzle. That the effect of deadlines on turnout appears to be substantially greater than the effect on registration rates is unanticipated and cannot be accounted for by any existing theoretical explanation. Existing theories imply that the effects of deadlines on turnout should be less (in magnitude) than the effect on registration (e.g. Wolfinger and Rosenstone 1980, Mitchell and Wlezien 1995, Brians and Grofman 2001, Brown and Wedeking 2006).

To better illustrate the puzzle, consider a hypothetical county with a voting age population of 10,000. Assuming that registration deadlines (comparing EDR to a 30-day deadline) lead to an increased rate of registration of 10%, we would expect an increase of 1,000 registered voters in our hypothetical county. If we assume (generously) that 95% of the new registrants turn out to vote, we would expect to see an increase of 950 voters, or 9.5%. In this example, which assumes a 95% turnout rate of new registrants, the increase in turnout is lower than the increase in registration. Yet we observe the opposite in the actual data. From the observed data, the effects on turnout appear to be considerably larger than the
effects on registration rates. That the effects are relatively larger for turnout is a puzzle in light of the prevailing explanations for how the process is thought to operate. There are several possible explanations we might consider.

Possible Concerns with the Analyses The first possible resolution to these findings is that the analyses themselves might not capture the actual effects of registration deadlines. The "gold standard" for statistical inference is randomizing subjects to treatment and control conditions. By controlling the assignment process, subjects are exposed to the treatment condition with a known probability, and can be used to balance all other factors between the treatment and control groups. The rationale being that if the groups are identical to one another on all other factors except the treatment condition, any observed differences on the outcomes can be attributed to the treatment. Without random assignment it is more difficult to rule out alternative explanations. The analyses above included several variables that have found to be related to voter participation, including age, income, and race and ethnicity. Yet there are other factors that are not accounted for that could influence the results. Unobserved variables might lead to the apparent empirical puzzle, which if accounted for, would resolve the discrepancy in the findings. That could be either because the magnitudes of the effects are misestimated or the effects might not be present at all, and these findings might be an artifact of the data.

Yet these possibilities might be unlikely to account for the results. It would seem unusual to confirm both hypotheses alone, but when considered in combination, to conclude that neither are supported by the data. Nevertheless, the possibility is further explored below by supplementing these analyses with a study of individual level data. Before turning to
these analyses, we can also consider alternative mechanisms that might help to account for these findings.

Alternative mechanisms These findings seem to suggest that while registration deadlines have an affect on participation, that the theoretical mechanisms might be misunderstood, or at least incomplete. A better understanding of the complex process by which individuals decide to participate or not in an election might help to better account for our findings. Two possibilities are considered here, which are mobilization and contagion effects, each of which might provide a useful explanation for the results.

Mobilization Existing theories of the effects of registration deadlines have not considered the possibility that relaxed registration laws affect campaign mobilization activities, and thereby voter turnout. Turnout has previously been shown to be responsive to mobilization efforts (Gerber and Green 2000, Imai 2005). If the candidates and parties are responsive to changes in registration policies and thereby increase mobilization efforts, voters might be responding to the additional GOTV activities. This mechanism might account for the findings, as a larger effect on turnout could be sustained. It also builds from previous research that establishes the link between GOTV and turnout, leaving the remaining connection between registration deadlines and mobilization. This specific relationship has, to my knowledge, not been examined in previous research but might help to explain these findings.

Contagion Effects A second mechanism might be that participation is subject to contagion effects whereby individuals' decision to vote is affected by similar decisions of those around them (Kenney 1992, McClurg 2004, Fowler 2005, Nickerson 2008). This process of
interference can create a contagion effect whereby inducing some individuals to vote by directly relaxing registration laws indirectly affects others around them. For example, a college student that is registered to vote will not be directly impacted by relaxing registration laws. Having previously registered, she is already eligible to vote and changing registration laws has no direct effect on her decision to vote. It might, however, affect whether her friends are registered to vote. By increasing the likelihood that her friends are registered to vote, her likelihood of participating might be indirectly affected by deadlines. If individuals are more likely to participate when close associates are known to participate, then this type of contagion might account for the previous findings. Such indirect contagion effects make it at least conceptually possible to explain the stronger relationship with turnout than registration.

While adding a contagion process make it theoretically possible to account for the higher effect of deadlines on turnout, it is not clear that such differences could be sustained by plausible levels of contagion. To initially consider the possibility, I conducted a simple stochastic simulation with a conservative amount of voter contagion. In this simulation, individuals vote with a probability of 50% under control and 55% under treatment. When individuals are taken in isolation, the direct effect of the treatment is about 5%, as expected. We can also simulate turnout when voters make decisions contingent on those around them. For this simulation, individuals are assumed to be affected by both the treatment and the behavior of their closest 20 associates in decreasing order. The weights given to associates' behavior are shown in Figure 3-16. The simulation then proceeds by randomly drawing vote intentions, allowing for one period of contagion where turnout probabilities for the full population are recalculated, then randomly drawing vote intentions based on the new probabilities again. The results are shown in Figure 3-17. This figure was created by
running the simulation 1000 times, and computing the difference in turnout rates between the treatment and control. The baseline turnout rate, when individuals are effectively quarantined, is included as the top graph, while the second graph shows the simulated treatment effect with contagion effects. From this figure we can see that contagion effects might help to account for the observed differences in the effects, as a 5% direct effect increases by an additional point, with only a modest amount of contagion and a single exposure period.

These results suggest that plausible contagion effects can lead the treatment to have a greater effect on turnout. The simulation allows modest effects of associates' behavior and only one round of contagion. Altering levels of interdependence and allowing for more rounds of contagion should increase these effects. The purpose of the simulation is not to definitely establish contagion effects among voters either in general, or a way of accounting for the effects of registration deadlines. Rather, the simulation attempts to establish that contagion effects offer a plausible explanation of the findings in this study.

3.3.4 Analysis of Survey Data
We can also examine individual-level data to further assess the effects of registration deadlines on participation. These data are taken from the Current Population Study, which included supplemental voter participation questionnaire in their November surveys in even numbered years. The CPS data on electoral behavior is a large nationally representative sample of individuals.

The data for this analysis spans 7 elections. Surveys were taken following federal general elections from 1994 through 2006. The total number of responses spanning these
data is 566,548. The surveys include several questions about electoral participation, including whether or not the respondent was registered prior to the November election, and if she voted in the election. In addition to these questions, they also ask respondents to choose from a list the most important reason for why she did not participate. These data can usefully inform this study, as it provides information as to why individuals do not participate. In order for registration laws to increase voter turnout, we should expect current registration laws to be a barrier to voting, and that it should be more significant in states with more restrictive registration deadlines. Among the choices on the CPS survey for why individuals do not vote is registration problems.

Table 3-6 shows the breakdown of the proportion of responses for each of the main reasons that an individual did not vote. We can see that registration problems was not as frequently mentioned by nonvoters in the CPS data as other potential barriers to participation. From this table, 3.87% of nonvoters in the CPS reported that they did not vote due to registration problems, which is significantly lower than the effect that we found with the aggregate analysis.

We can also consider whether this percent varies according to closing dates. Figure 3-18 shows the proportion of nonvoters reporting registration problems across the range of observed closing dates. From this we can again see that the magnitude of the effect is almost uniformly lower than what we would expect from the aggregate level analysis, though the pattern is consistent with the arguments from above. That is, the proportion of nonvoters reporting registration problems is lowest in states with EDR, and increases from there. By about 10 days the proportion appears to remain fairly stable across the years.
We can further examine data on those that were not registered in order to assess the potential impact of registration deadlines. As before, the CPS survey asks individuals to choose from a list of the main reason that they are not registered to vote. The frequencies for the overall data are shown in Table 3-7. These data suggest that the most significant barrier to participation is individuals' interest in politics, which was chosen by over 40% of respondents, rather than registration deadlines, which was chosen by nearly 15% of non-registrants. While these data might suggest that voter registration deadlines matter less for participation, that conclusion might not necessarily follow from the data. The question asks for the most significant reason that individuals are not registered, and while the modal response was "not interested in the election" that does not necessarily mean that registration deadlines would have no effect on their behavior.

Individuals' registration status might likely be affected by a number of factors, and while the modal number of respondents chose a lack of interest in the campaigns from their not being registered, we should not necessarily conclude that closing dates are irrelevant for participation. Not only did a fair number of respondents choose registration problems, but we should also not assume that those individuals are not at all affected by the length of registration deadlines.

Breaking the responses down by closing dates shows a pattern that is consistent with the findings from above. Figure 3-19 shows the frequency of responses by registration deadlines. As we would expect, fewer respondents reported missing the registration deadline in states with EDR, although a striking 10% still claimed to miss the registration deadline when there is effectively no such deadline (except possibly for absentee, military, or overseas voters that are not available to register at polling locations). Once we reach about 10 days,
however, the effect of closing deadlines seems to level off, with a consistent 15% listing the deadline as the reason they are not registered. In addition, individuals in states with EDR report registering at polling locations with more frequency than any other modes of registration. Over 38% of respondents reported registering at a polling location, with another 33% registering at a county government office. The remaining responses are distributed across the six other categories, as shown in Table 3-8.

We can also more formally analyze the available individual data for more insight as to the effects of registration deadlines on turnout and registration. The analysis in this section attempts to adjust for other factors that might reasonably be considered alternative explanations for the differences in participation rates across different levels of registration deadlines. Specifically, we can use the CPS data to adjust for other characteristics of the individuals in the study, including age, sex, race and ethnicity, education, income, and the competitiveness of the election. By adjusting for these factors, we can be more confident that the results we find are actually attributable to the effects of registration deadlines, rather than representing an artifact of the data that might occur if, for example, there were more educated individuals living in areas with shorter registration deadlines.

As discussed in the preceding section, the individual-level data are obtained from the CPS, spanning elections from 1994 to 2006. The control variables taken from the CPS are age, sex, race and ethnicity, education and income. An additional measure for the competitiveness of the election is also obtained from archival data on election returns. Theoretical variables measuring turnout and registration are also taken from the CPS, and the measure of closing deadlines are obtained by matching the state in which each of the
respondents' lives and the year of the interview to the registration deadline in their state for that year. This section briefly describes the measures used in the following analyses.

**Sex:** This variable is a dichotomous measure for the individuals' sex, 0 for male and 1 for female. For clarity, the measure of sex is named female in the presentation of the results to indicate a value of 1 for females.

**Age:** The age variable is measured in years, and spans values of 18 to 90 in the data. The data are limited to this range, as individuals' must be at least 18 years old to be eligible to vote, and the CPS sets a limit of the age of respondents to 90, pooling interviews greater than 90 to an age of 90. This pooling will likely not substantially affect the results, as any additional year of age above 90 is unlikely to substantially effect individuals' participation and reflects a small proportion of the data.

**Race and Ethnicity:** To account for individuals' race and ethnicity, separate dummy variables are created for each group, which includes Anglos, blacks, Latinos, Asians, and other. The other category was created for the least frequently occurring responses in the data, which collapsed responses for Native American, Pacific Islander, other, and multiple races. In the models, variables for blacks, Latinos, Asians, and others are included while the Anglo variable is excluded.

**Education:** The education variable is a 6 point ordinal measure which includes categories for less than high school, high school or equivalent, some college, associates' degree, bachelor's degree, and post-graduate and above.
**Income:** The income variable is a 14 point scale with values starting at “Less than $5,000” then incrementing to “$75,000 and above.” Beginning in 2004, the CPS offered a 16 point scale by adding categories to the top bracket which allowed for responses to $150,000 and above. For comparability, responses in the 2004 and 2006 surveys were collapsed to fit the previous scale.

**Competitiveness:** The final control variable measures the competitiveness of the top-ticket race in the state by year. For each election, the variable is calculated by taking the top vote getters' share of the votes divided by the sum of her share and the share of her nearest competitor. For elections in 1996, 2000, and 2004, the presidential race is used, while in 1994, 1998, 2002, and 2006, the measure is calculated for gubernatorial and senate races. When both offices were up for election at the same time, the more competitive race was used. The reason for using the more competitive race is that it is more likely to be the one to stimulate voter turnout.

**Registration and Turnout:** The two dependent variables are both dummy variables for whether the individual reported voting in that year’s general election, and whether or not the individual reported being registered to vote in the election. Each variable is coded as 0 for not voting or not registered, and 1 for having voted or being registered. One disadvantage of the CPS data is that the responses are not validated, and the registration question is also skipped for voters. To correct the data, responses for the registration question were recoded as 1’s for all respondents that reported voting in that election.

Using these measures from the CPS data, I matched the individuals' to the measures for competitiveness and registration deadlines based on the state in which the subject lived.
The total data set had 566,548 observations. I then randomly selected 5000 cases from the data, and matched them to the nearest 5000 cases in the remainder of the data. The data were subsetted for a number of reasons. The two main reasons are that a smaller data set is more manageable, particular for conducting the matching procedure and the analysis. Furthermore, while the smaller data set leads to somewhat more chance variation in the parameter estimates, an analysis with 10,000 cases is by no means small and if the effects of registration deadlines cannot be detected with a sample of 10,000 they are unlikely to be of substantive interest. This relates to one of the criticisms of null hypothesis significance tests, which is a question of whether one has enough data to differentiate the associations from zero. The second reason is that subsetting the data allows for a better match, as there is a larger pool of potential controls from which to select. As the number of cases increases, the matching procedure will search out cases that are farther apart. In this case, a sample of 10,000 should be more than adequate to detect the effects, and obtaining a good match is of more central concern as imbalances in the data can confound the results. The following section describes the choice of matching techniques, followed by a discussion of the analysis of the matched samples and the results.

Matching methods have recently been proposed as a technique for analyzing observational data in political science studies (as well as potentially addressing some challenges in experimental studies) (Ho, Imai, King, and Stuart 2007). The basic intuition of matching methods is to compare like observations to one another when exposed to treatment and control conditions. In this way, matching methods more closely approximate experimental conditions when randomization is not feasible. The primarily strength of experimental designs is that by randomly assigning subjects to different conditions, with a
large enough sample, the groups will be identical to one another on all other variables. Without random assignment, this might not necessarily be the case. Rather than imposing an assumption of balance on observational data or the appropriateness of a particular statistical model, matching creates more balances samples, akin to what we would expect if subjects had been randomly assigned, on observable pre-treatment covariates. Compared to experiments, matching still has the disadvantage of not adjusting for unobserved pre-treatment variables, which may still affect the results, but in this case experimental control of treatment assignments is not feasible. In contrast to model-based adjustments, however, matched samples have several advantages. Regarding statistical inference, matched samples have been shown to reduce the dependence of empirical results on model specification. In theoretical and applied research, matching has also been shown to provide results that are closer to experimental benchmarks than relying on regression adjustments alone (Ho, Imai, King, and Stuart 2007). Aside from providing a direct improvement in inferences, matching is also an intuitive way of presenting results, as it compares like observations with one another. Matching furthermore discourages model-hunting, which has been argued to be a serious problem in empirical research (Freedman 1991). While researchers must have access to the full data set for regression, matching only requires access to the pre-treatment covariates in order to obtain the observations for analysis. Once the matched samples are created, researchers can then access the outcome variable for analysis. By separating the specific analysis plan from the results, researchers can specify their designs without seeing the results or outcome variables. In this way, matching can be a way of discouraging model-hunting by making it more difficult to conduct many different "runs" than is the case with regression (though no design is entirely immune from these issues).
**Propensity score matching** There are at two general ways of conducting matching, exact matching using the observed control variables, and propensity score matching.\textsuperscript{10} Exact matching is the most straightforward and creates perfectly balanced samples on the observed potentially confounding variables. One practical shortcoming of exact matching is that it can be difficult to implement when adjusting for a large number of variables or when they have a large number of possible responses. The reason is that it can be difficult to find enough exactly matching cases to analyze. When exact matching is not practical propensity score matching can be a useful alternative. In the most common application, propensity score matching uses the observed covariates to estimate the probability that observations are in the treatment or control groups using, for example, logistic regression. Observations can then be matched using some function of the propensity scores. Propensity score matching has been found empirically and theoretically to provide balanced samples on the observed potentially confounding variables (Rosenbaum 2002). Yet neither exact nor propensity score matching can be used for this study, as there are more than two treatment groups (there are 12 different registration deadlines represented in the data).

**Generalized propensity scores for dosage studies** Since there are more than two groups of subjects under investigation, we will need to use a generalized propensity score. There have been two proposed generalizations of the propensity score to polychotomous treatments. One is proposed by Imbens, which is based on estimating propensity scores from a multinomial logistic regression of selection into the different treatment categories. This method allows for multiple treatment conditions, even when they are not arrayed along an ordinal or interval scale. One shortcoming of Imbens' approach is that it does not directly

\textsuperscript{10} Observations can also be variously grouped into strata based on covariate or propensity score values.
represent the causal effect of the treatment without further adjustment. An alternative approach is suggested by Lu, Zanutto, Hornik, and Rosenbaum (2001). Their approach is useful for studies involving dosages that have an ordered ranking. Their method also does not require further adjustment in order to estimate causal effects (in the absence of unobserved bias). Since the treatments (or doses) in this study are arrayed along an interval scale, the generalized propensity score proposed in Lu et al. is applied.

Adopting the generalized propensity score as proposed in Lu et al. involves two changes to propensity score matching with a dichotomous treatment. The first is the method by which the propensity scores are obtained. With a dichotomous treatment the propensity scores are most commonly estimated using a logit or probit model. With a polychotomous treatment, Lu et al. propose using an ordered logit or linear regression model. With the larger number of categories in this case, I chose to use a linear regression model in order to obtain the estimated propensity scores. The second change is the way the matching is conducted. We would like the matched sample to have observations that are as similar to one another as possible on the pretreatment covariates, and as different as possible on the received dose.

Thus typical matching procedures have to be amended to account for the difference between the received doses in the measure of distance used to match observations. When comparing a treatment to control group, observations can be matched across the two groups that have the most similar propensity scores. When matching with different doses, we want to match observations that are as close as possible on the propensity score while as far apart on the actual dose received (Lu et al. 2001). Thus the matching involves second factor in
addition to the propensity score difference. The distance used to match observations is (Lu et al. 2001):

$$\Delta(x_k, x_{k'}) = \frac{(\hat{\beta}x_k - \hat{\beta}x_{k'})^2 + \epsilon}{(Z_k - Z_{k'})^2}$$

In this equation, the left-hand side represents the distance between an observation $k$ and a potential match, $k'$. The right-hand side represents the squared difference between the estimated propensity scores from the linear model, divided by the squared difference between the doses received by the two units. The denominator ensures that two observations with the same dose will never be matched, as it returns an infinite distance. The epsilon term is a small positive number, which does not affect the results, other than to doing two things. The first is that it makes sure that no observation is matched to one with the same dose, as the distance will always be infinite, even if the observations have identical propensity scores (rather than 0). The second role of epsilon is to factor in the different doses in the event that potential matches have identical propensity scores. In this case, we would want to match on the observation with the most different dose, and the epsilon insures that the numerator will never be exactly zero (Lu et al. 2001).

Using this distance measure, each of the 5000 randomly selected observations were matched to the available observation that minimized this distance. The matching procedure was conducted using a greedy algorithm that matched each observation sequentially. This procedure returned a sample of 10,000 observations in 5,000 matched pairs. In part to serve as a robustness check for the random sampling, different samples were drawn for the analysis of voter turnout and voter registration. As discussed below, both results are consistent with the theoretical expectations and the aggregate level results.
**Matched sample results** To assess the results of the matched sample we can see how the different dosage groups break down across the potentially confounding variables. For the matched sample of voter turnout, Figures A7 though A13 (in the Appendix A) show dot charts of the control variables across the different registration dosage groups. We can see that there are some groups, such as 1 and 17 that differ from the other observations, but even these are relatively similar. We would also look for systematic departures, such as several of the low groups differing from the general trend among the higher dosage groups, which is not apparent from these figures.

For the matched sample of voter registration, Figures A14 through A20 (in the Appendix A) show the breakdown of the dosage groups for each of the variables. Again, we can see from these charts that there does not seem to be substantial departures on the control variables across the different dosage groups. The data from the matching procedure seems fairly well-balanced across the different dosage groups for both samples, and while does not adjust for unobserved variables, makes it less likely that any observed differences in participation rates can be attributed to differences in the samples on these control variables.

**Analyses and results** The first analysis uses the matched sample data to analyze the impact of registration deadlines on the probability that individuals turn out to vote. For this analysis, models of varying complexity are considered and are reported in Table 3-9. From these results we can see that the results for the registration variable are consistent across models. In all of the models the coefficient for the registration deadline variable is negative, as anticipated. The 90% confidence intervals also do not overlap with 0, suggesting that repeated sampling of the individual level data would also return similar results. In order to provide a clearer interpretation of the magnitude of the effects, rather than just considering
their direction, I simulated the parameter values and calculated predicted probabilities of voting across the range of registration deadlines. For the simulations, I used the estimates from the simplest model, as the estimates for that model had the smallest magnitude. These results should thus be interpreted as consistent with the overall results, but also the most conservative of the model estimates so that the effects might be larger.

The substantive effects for the model of voter turnout are shown in Figure 3-20. The simulations were conducted by drawing 1000 random parameter vectors from the appropriate multivariate normal distribution. Those vectors were then used to calculate predicted probabilities for each value of the registration deadline. The figure shows the median probabilities from the simulation, along with 50% and 90% intervals. The line slopes downward as the registration deadline increases, representing the negative association with turnout. We can see from this figure that the probability of voting ranges from a high of approximately .62 when there is no registration deadline to a minimum of approximately .56 with a 30 day deadline. The rug plot at the bottom of the figure represents the distribution of the data for registration deadlines. While there is some uncertainty in the parameter estimates as represented by the vertical bars, there is not sufficient overlap in the vertical bars to attribute the apparent effects of registration deadlines on turnout to simply chance variation from random sampling of the data.

A similar analysis was conducted for whether or not the individuals' reported being registered. Results from several models of voter registration are shown in Table 3-10. Similar to the analysis of turnout, the results are consistent across model specification, and the effects of registration deadlines are negative as anticipated.
An effect plot was also created for the analysis of voter registration and is shown in Figure 3-21. These simulations were conducted using the simplest model, which also had the coefficient with the lowest magnitude. Thus while all of the models suggest a negative effect, these simulations are likely to be conservative estimates of the actual effect of registration deadlines on voter registration rates. We can see again the downward sloping line, and a change in probability of approximately 8% (comparing the median values) over the range of the data. This is consistent with the results from the model of voter turnout.

The results from the individual analysis if taken alone would still not seem to resolve the puzzle as to the relative size of the effects of deadlines for turnout and registration. A direct comparison of the effects is shown in Table 3-11. From this table we can see that the relative magnitude of the effects is closer to what we would expect, as the effects on individuals' probability of being registered is greater than their probability of voting, but the effects are still closer than expected. The effect on individuals' predicted probability of registering when comparing a 30 day deadline to no deadline is about 8%, while the effect for turnout is about 6%. Thus while the individual analysis might help to account for part of the puzzle from the aggregate results, these results also seem to confirm that the prevailing rationale as to how registration deadlines affect turnout is incomplete. The relative magnitude of the effects of deadline on turnout and registration, in both the analyses of aggregate and individual-level results seems to suggest that there exists an as yet unidentified mechanism by which registration deadlines affect voter turnout. Several possible alternative mechanisms are explored above. The alternative mechanisms might not only be a way to help resolve this empirical puzzle as to the effects of registration deadlines, but might also
help to inform the broader question as to what general factors affect voter participation. Further implications of these findings are discussed in the concluding section of this chapter.

3.4 Effects of Election Day registration

The foregoing analyses all consider the effects of registration deadlines on participation, which has been a different question than much of the existing literature on registration laws has analyzed. Much of the findings to date have analyzed EDR, which is related to registration deadlines, but is not identical. Studies of EDR typically consider only a binary independent variable, comparing jurisdictions with EDR to those without EDR. Yet this simplifies registration laws by not accounting for the variation in registration deadlines in non-EDR states. Studies that compare just compare EDR to non-EDR states can mask the effects of varying registration deadlines on participation.

The second unique feature of EDR is that it allows individuals to register at their polling location. Voters in a jurisdiction with EDR do not (as currently implemented) have to travel to multiple sites in order to register to vote and then cast a ballot. Rather, all EDR states allow registration at polling locations. This might add to the effects of EDR. Not only will the same logic apply as above, but individuals will also save themselves the additional cost of going to a different location to complete the registration applications than where they go to vote. This section of the chapter addresses the question as to whether or not there is an additional effect of EDR over the findings for registration deadlines more generally. This section reports two analyses, that contrary to most of the prevailing findings regarding the effects of EDR, do not find evidence of an additional effect of EDR policies (Brians and Grofman 2001).
3.4.1 Research Design

The first analysis exactly replicates those above, which the inclusion of an EDR dummy variable. This variable has a value of 1 for all elections in which the county offered EDR and includes North Dakota, which has no voter registration requirement. The variable is coded as a 0 for all other elections. The dependent variable is again turnout as a ratio of VAP.\textsuperscript{11} The control variables are also same, and includes the deadline variable from the preceding analyses to capture the potential added effect of same-day registration. The expectation is that the effect of registration deadlines will remain negative while the EDR variable should be associated with an increase in participation.

The first series of models include the EDR variable in several models with different specifications. Linear regression is again used to model turnout and registration rates. The results are shown in Table 3-12. From these results we can see two patterns. The first is that there is a fair amount of instability in the parameter estimates of the effects of EDR depending on model specification. The most basic model that only includes the dichotomous measure estimates the effect of EDR on turnout to be 10%. Including control variables reduces this effect, and in some cases returns a marginally negative estimate. Yet even withholding the parameter instability, the positive estimates of the effects of EDR in models including covariates suggests that EDR has only weak effects apart from registration deadlines. From the full model, the anticipated effect of EDR against a non-EDR state with a 10 day registration deadline is only 3.7%, with 3% coming from the deadline alone and the

\textsuperscript{11} All of the analyses were also run using CVAP as the dependent variable. The results were all similar, and the substantive conclusions do not depend on which of the variables are considered.
additional .7% attributable to EDR. Comparing EDR against a hypothetical 1 day deadline would lead to an increase in turnout of only 1% based on the results of the full model.

3.4.2 Analysis of Statewide Turnout

To supplement these findings, an additional analysis of state-level data is also considered. One of the challenges of examining policies such as same-day registration is that individual level data can be difficult to obtain and is rarely verified to adjust for possible over-reporting. On the other hand, aggregate level data can be easier to obtain and provide more reliable measures. Yet with fewer cases aggregate data makes it more difficult to justify using particular cases for comparison rather that some other case. Particularly with data on U.S. states, it is not always clear that there is a best natural comparison case for any given state in a study. In order to avoid making what is necessarily a somewhat subjective choice of natural controls, we can use a data-driven process to create synthetic controls of the treatment cases to use for comparison (Abadie and Gardeazabel 2003; Abadie, Diamond, and Hainmueller 2007).

The synthetic control method is first used in a recent study of the effects of conflict in the Basque region of Spain on economic growth (Abadie and Gardeazabel 2003). Instead of choosing one of the available regions of Spain that did not have conflict as a basis for comparison, Abadie and Gardeazabel construct a "synthetic" control for the Basque region by using a data-driven process to identify the optimal combination of all of the available controls that is most similar to the region. That is, Abadie and Gardeazabel use all of the other regions to construct an artificial comparison case, which is a weighted combination of the controls, and provides a closer comparison case than any of the individual controls.
The objective of constructing the synthetic comparison case, is to find an artificial region that is as similar to the region under study as possible, on the pretreatment covariates. To do this, Abadie and Gardeazabel provide a function to be minimized, which is:

\[(X_1 - X_0W)'V(X_1 - X_0W)\]

In this equation, \(X_1\) represents a \((K \times 1)\) vector of \(K\) pre-treatment predictors of the outcome variable, \(X_0\) represents a \((K \times J)\) matrix of \(K\) pre-treatment predictors for \(J\) controls. \(W\) represents a \((J \times 1)\) vector of weights, such that all of the weights are non-negative and sum to 1. \(V\) represents a diagonal matrix of non-negative elements that represent the relative importance of the predictors (Abadie and Gardeazabel 2003). The values in \(V\) can either be set by the researcher, or following from previous research, determined by setting the values such that the pretreatment values of the outcome variable for the treated observation most closely match the synthetic control (Abadie and Gardeazabel; Abadie, Diamond, and Hainmueller 2007).

This approach is used in this study to create synthetic controls of the "second-wave" EDR states that first implemented same-day registration in 1994. These states are Idaho, Wyoming, and New Hampshire. The pool of controls includes all states other than these three, North Dakota (which has no registration requirement), Minnesota, Maine, and Wisconsin. These states were excluded as they adopted EDR in the 1970s. Montana also adopted EDR in 2006, and was included in the analysis as it only implemented the treatment in the last observed year, and the effects are thought to manifest themselves over time (as registration is valid across elections, the effects of EDR should be experience over time as the electorate experiences greater turnover).
Using the pool of available controls with the method described above, synthetic controls were created for the EDR states (using software available from Abadie, Diamond, and Hainmueller 2007). Each of the synthetic controls represent a different weighted combination of the non-EDR states, as shown in Table 3-13. The states were then compared against the synthetic states by calculating the difference in the turnout rates for each election from 1980 to 2006. The resulting plot is shown in Figure 3-22. From this figure we can assess the quality of the match as well as the effects of the treatment. All of the states adopted EDR in 1994, which is represented by the gray vertical line. The horizontal line at 0 represents the synthetic control (since the y-axis is treatment-control, the control group is always at 0). Ideally, we would like to see the treatment lines stay as close to 0 as possible in the pre-treatment years. The closer the lines to 0, the better the synthetic control is likely to match the treated observation. Just assessing the differences in the pre-treatment years, we can see that there are some departures, but all of the cases were usually within .1 of the synthetic control for every pre-treatment election.

Assuming that EDR worked as anticipated, we should see the EDR lines to increase after 1994, and likely plateau after the treatment has taken full effect. Instead the lines do not show a consistent pattern of increase or decline, but rather are positive for some elections and negative for others. While it is difficult to discern a consistent pattern of increase or decrease in turnout as a result of EDR from Figure 3-22, there might nevertheless be systematic effects present in the data. While EDR might lead to an increase in some elections and a decrease in other is possible, if not difficult to understand. To further assess this possibility, the method was run for all of the control cases as well.
The rationale for this analysis follows from the previous study by Abadie et al. The approach lets us assess graphically how much natural variation there is in the data using this method, apart from any systematic effects of the treatment. Since the implementation of the treatment in one state is not thought to affect turnout in another state (either an increase or decrease), we can compare the post-treatment outcomes in Idaho, New Hampshire, and Wyoming, to the post-treatment outcomes in the control states. If the treatment is having any effect, even one that is not consistent from election to election, we should see the cases in the treatment states take on more extreme values than those in the control states, where we do not expect any systematic differences in the post-treatment period.

The results of this analysis is shown in Figure 3-23. This figure suggests that EDR has an effect that is not discernable in these states, for recent elections, that is either positive or negative for voter participation. The three curves shown in bold for the treatment states appear to be similar to the other lines for control cases (all of the light gray lines), so that the bold lines blend into those cases in which nothing systematic was known to change about registration laws. The figure shows that it would be difficult to separate the outcomes associated with the treatment cases from the controls, suggesting that EDR does not have an identifiable affect on voter participation.

3.5 Chapter Review

This chapter set out to evaluate three empirical hypotheses:

**H1:** Shorter closing periods should lead to higher levels of turnout.

**H2:** Shorter closing periods should lead to higher levels of registration.

**H3:** Election Day registration leads to higher rates of turnout.
These hypotheses build from considerable existing political science research, but for which somewhat different results are obtained. The prevailing understanding of registration deadlines is that longer deadlines decrease voter turnout, and the process by which deadlines affect turnout is simply by reducing the number of registered voters that are eligible to turn out. The main points of contention regarding the effects of deadlines on turnout have been the specific magnitude of the effects, such as whether elimination of deadlines would produce a 10% or a 5% increase in turnout, and if they affect lower SES individuals to a greater or lesser extent than others. Apart from these issues, the prevailing understanding of the effects of registration deadlines seems fairly well accepted.

The results of this study, however, suggest that scholars might need to reconsider some aspects of our understanding of the effects of registration policies. While previous research has largely been based on state level data and survey results from prior elections (many from more than 15 years ago) from the ANES and CPS, the main results reported in this chapter are taken from county turnout and registration figures from 1992 to 2004. These results both confirm the hypotheses stated above as well as raise new and interesting questions. The effects of registration deadlines are shown to be around 10% for turnout and 6% for registration rates. Based on these results, we can conclude that registration deadlines impact electoral participation.

The problem is that while the results confirm the hypotheses, they also draw into question the theoretical rationale from which the hypotheses are derived and as "simply" stated in several previous studies of registration reform and what seems to be implicitly assumed in other studies. That is, the effects of registration deadlines on turnout operate through registration rates. Yet the results suggest that this theoretical understanding of the
process is at best incomplete, and potentially incorrect. The notable difference in the magnitude of the effects, with turnout showing a larger effect than registration rates cannot be accounted for by any existing theoretical explanation of which I am aware.

The results of the individual analyses using CPS data had several potential drawbacks, but provided a less stark picture of the effects of registration deadlines. The effects from these analyses suggest that deadlines have an 8% effect on registration rates and 6% on turnout. While less striking than the aggregate results, these findings also imply that 75% of new registrants turn out to vote, which seems high when including both presidential and midterm elections and also as new registrants are likely to be disproportionately young and mobile. Taken together, these findings largely confirm the empirical results from previous studies that registration deadlines adversely affect participation, but raise questions about our theoretical understanding of the mechanism that produces this effect. Two possible theoretical extensions are considered above, which are mobilization effects and contagion. Mobilization would seem to be a viable extension of previous research on get-out-the-vote efforts, depending on the response of parties and candidates to changing electoral rules. Contagion effects suggest an alternative process by which the voters interact with one another, which might be of more general theoretical interest, and a simulation was provided that suggests contagion effects might plausibly account for the different effects.

The third hypothesis involved the effects of Election Day registration on turnout as distinct from shorter closing deadlines. The findings from this section suggest that the bulk of the effects of EDR are a result of shorter registration deadlines, and that the availability of registration specifically at polling locations has only a marginal impact on participation. An analysis of state-level data on second wave adoptees of EDR also failed to uncover
compelling evidence of the effects of EDR on turnout. Taken together, these results suggest that while EDR might increase participation against a 30-day deadline, the difference between EDR and shorter deadlines such as 5 or 10 days is substantially lower.
Figure 3-1: Histogram of Registration Deadlines

![Histogram of Registration Deadlines](image)

Figure 3-2: Scatterplot of Turnout and Deadline Variables

![Registration Deadlines and Turnout](image)
Figure 3-3: Scatterplot of Registration and Deadline Variables

Figure 3-4: Scatterplot of Estimated County Turnout and Deadlines
Figure 3-6: Plot of Turnout over Years (+- 1 s.d.)

Registration Closures and Turnout, 1992-2004
**Figure 3-7: Correlation Plot with Turnout**

![Correlation Matrix With Turnout](image)

**Figure 3-8: Correlation Plot with Registration**

![Correlation Matrix With Registration](image)
Figure 3-9: States with Decreases in Deadlines

- California: Change -14 days
- Idaho: Change -10 days
- Illinois: Change -1 days
- Nevada: Change -10 days
- New Hampshire: Change -10 days
- Virginia: Change -2 days
- West Virginia: Change 9 days
Figure 3-10: States with Increases in Deadlines

Arkansas

Colorado

Louisiana

Figure 3-11: States with Multiple Changes

Maryland

Missouri

Vermont
Figure 3-12: Posterior Distribution for Turnout Estimates

![Graphs showing posterior distributions for various factors affecting turnout.]

Figure 3-13: Simulated Data Sets for Turnout

![Graphs comparing actual data with simulated data sets.]

- Actual Data
- Simulated Data 1
- Simulated Data 2
- Simulated Data 3
- Simulated Data 4
- Simulated Data 5
Figure 3-14: Posterior Distributions from Registration Model

![Posterior Distributions from Registration Model](image)

Figure 3-15: Simulated Data Sets for Registration

![Simulated Data Sets for Registration](image)
Figure 3-16: Weights for Contagion Simulation

![Simulation Weights](image)

Figure 3-17: Simulation Results

**No Contagion**

![No Contagion](image)

**With Contagion**

![With Contagion](image)
Figure 3-18: Registration Problems and Turnout over Deadlines

Who Has Registration Problems?

Figure 3-19: Proportion Missing the Deadline

Who Misses the Deadline?
Figure 3-20: Predicted Probabilities of Voting

Figure 3-21: Predicted Probabilities of Registering
Figure 3-22: Differences in Turnout between Treated and Synthetic Controls

![Graph showing differences in turnout between treated and synthetic controls.]

Figure 3-23: Differences for Treated and Non-treated States

![Graph showing differences in turnout for treated and non-treated states.]
Table 3-2: Number of Observations for Voter Turnout Data by State by Year

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>469</td>
</tr>
<tr>
<td>Arizona</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>105</td>
</tr>
<tr>
<td>Arkansas</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>75</td>
<td>525</td>
</tr>
<tr>
<td>California</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>406</td>
</tr>
<tr>
<td>Colorado</td>
<td>63</td>
<td>63</td>
<td>63</td>
<td>63</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>443</td>
</tr>
<tr>
<td>Connecticut</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>56</td>
</tr>
<tr>
<td>Delaware</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>21</td>
</tr>
<tr>
<td>Florida</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>469</td>
</tr>
<tr>
<td>Georgia</td>
<td>159</td>
<td>159</td>
<td>159</td>
<td>159</td>
<td>159</td>
<td>159</td>
<td>159</td>
<td>1,113</td>
</tr>
<tr>
<td>Hawaii</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>Idaho</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>308</td>
</tr>
<tr>
<td>Illinois</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>714</td>
</tr>
<tr>
<td>Indiana</td>
<td>92</td>
<td>92</td>
<td>92</td>
<td>92</td>
<td>92</td>
<td>92</td>
<td>92</td>
<td>644</td>
</tr>
<tr>
<td>Iowa</td>
<td>99</td>
<td>99</td>
<td>99</td>
<td>99</td>
<td>99</td>
<td>99</td>
<td>99</td>
<td>693</td>
</tr>
<tr>
<td>Kansas</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>735</td>
</tr>
<tr>
<td>Kentucky</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>840</td>
</tr>
<tr>
<td>Louisiana</td>
<td>64</td>
<td>N/A</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>384</td>
</tr>
<tr>
<td>Maine</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>112</td>
</tr>
<tr>
<td>Maryland</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>168</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>98</td>
</tr>
<tr>
<td>Michigan</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>581</td>
</tr>
<tr>
<td>Minnesota</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>609</td>
</tr>
<tr>
<td>Mississippi</td>
<td>82</td>
<td>82</td>
<td>82</td>
<td>N/A</td>
<td>82</td>
<td>82</td>
<td>82</td>
<td>492</td>
</tr>
<tr>
<td>Missouri</td>
<td>114</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>810</td>
</tr>
<tr>
<td>Montana</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>392</td>
</tr>
<tr>
<td>Nebraska</td>
<td>93</td>
<td>93</td>
<td>93</td>
<td>93</td>
<td>93</td>
<td>93</td>
<td>93</td>
<td>651</td>
</tr>
<tr>
<td>Nevada</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>119</td>
</tr>
<tr>
<td>New</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>70</td>
</tr>
<tr>
<td>Hampshire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Jersey</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>147</td>
</tr>
<tr>
<td>New Mexico</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>231</td>
</tr>
<tr>
<td>New York</td>
<td>59</td>
<td>59</td>
<td>59</td>
<td>59</td>
<td>59</td>
<td>59</td>
<td>59</td>
<td>413</td>
</tr>
<tr>
<td>North Carolina</td>
<td>100</td>
<td>N/A</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>600</td>
</tr>
<tr>
<td>----------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>North Dakota</td>
<td>53</td>
<td>53</td>
<td>53</td>
<td>53</td>
<td>53</td>
<td>53</td>
<td>53</td>
<td>371</td>
</tr>
<tr>
<td>Ohio</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>616</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>539</td>
</tr>
<tr>
<td>Oregon</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>252</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>469</td>
</tr>
<tr>
<td>Rhode Island</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>273</td>
</tr>
<tr>
<td>South Carolina</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>322</td>
</tr>
<tr>
<td>South Dakota</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>462</td>
</tr>
<tr>
<td>Tennessee</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>665</td>
</tr>
<tr>
<td>Texas</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>1,778</td>
</tr>
<tr>
<td>Utah</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>203</td>
</tr>
<tr>
<td>Vermont</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>98</td>
</tr>
<tr>
<td>Virginia</td>
<td>136</td>
<td>136</td>
<td>135</td>
<td>135</td>
<td>135</td>
<td>134</td>
<td>134</td>
<td>945</td>
</tr>
<tr>
<td>West Virginia</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>385</td>
</tr>
<tr>
<td>Wisconsin</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>504</td>
</tr>
<tr>
<td>Wyoming</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>161</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Total</th>
<th>3,143</th>
<th>2,981</th>
<th>3,144</th>
<th>3,062</th>
<th>3,144</th>
<th>3,144</th>
<th>3,144</th>
<th>21,762</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alabama</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>469</td>
</tr>
<tr>
<td>Arizona</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>15</td>
<td>105</td>
</tr>
<tr>
<td>Arkansas</td>
<td>75</td>
<td>75</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>150</td>
</tr>
<tr>
<td>California</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>406</td>
</tr>
<tr>
<td>Colorado</td>
<td>63</td>
<td>63</td>
<td>63</td>
<td>63</td>
<td>63</td>
<td>64</td>
<td>64</td>
<td>443</td>
</tr>
<tr>
<td>Connecticut</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>8</td>
<td>56</td>
</tr>
<tr>
<td>Delaware</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>N/A</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>Florida</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>469</td>
</tr>
<tr>
<td>Georgia</td>
<td>159</td>
<td>159</td>
<td>158</td>
<td>158</td>
<td>158</td>
<td>158</td>
<td>158</td>
<td>1,108</td>
</tr>
<tr>
<td>Hawaii</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>Idaho</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>44</td>
<td>308</td>
</tr>
<tr>
<td>Illinois</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>102</td>
<td>714</td>
</tr>
<tr>
<td>Indiana</td>
<td>92</td>
<td>92</td>
<td>92</td>
<td>92</td>
<td>92</td>
<td>92</td>
<td>92</td>
<td>644</td>
</tr>
<tr>
<td>Iowa</td>
<td>99</td>
<td>99</td>
<td>99</td>
<td>99</td>
<td>98</td>
<td>99</td>
<td>99</td>
<td>692</td>
</tr>
<tr>
<td>Kansas</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>105</td>
<td>735</td>
</tr>
<tr>
<td>Kentucky</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>120</td>
<td>840</td>
</tr>
<tr>
<td>Louisiana</td>
<td>64</td>
<td>N/A</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>64</td>
<td>384</td>
</tr>
<tr>
<td>Maine</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>16</td>
<td>112</td>
</tr>
<tr>
<td>Maryland</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>24</td>
<td>168</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>14</td>
<td>14</td>
<td>N/A</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>84</td>
</tr>
<tr>
<td>Michigan</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>83</td>
<td>581</td>
</tr>
<tr>
<td>Minnesota</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>87</td>
<td>609</td>
</tr>
<tr>
<td>Mississippi</td>
<td>82</td>
<td>82</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>N/A</td>
<td>164</td>
</tr>
<tr>
<td>Missouri</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>116</td>
<td>812</td>
</tr>
<tr>
<td>Montana</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>56</td>
<td>392</td>
</tr>
<tr>
<td>Nebraska</td>
<td>93</td>
<td>93</td>
<td>N/A</td>
<td>N/A</td>
<td>93</td>
<td>93</td>
<td>93</td>
<td>465</td>
</tr>
<tr>
<td>Nevada</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>17</td>
<td>119</td>
</tr>
<tr>
<td>New</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>70</td>
</tr>
<tr>
<td>Hampshire</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Jersey</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>21</td>
<td>147</td>
</tr>
<tr>
<td>New Mexico</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>33</td>
<td>231</td>
</tr>
<tr>
<td>New York</td>
<td>59</td>
<td>59</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>408</td>
</tr>
<tr>
<td>North Carolina</td>
<td>100</td>
<td>N/A</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>99</td>
<td>100</td>
<td>599</td>
</tr>
<tr>
<td>Ohio</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>88</td>
<td>616</td>
</tr>
<tr>
<td>----------------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>-------</td>
</tr>
<tr>
<td>Oklahoma</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>77</td>
<td>539</td>
</tr>
<tr>
<td>Oregon</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>36</td>
<td>252</td>
</tr>
<tr>
<td>Pennsylvania</td>
<td>67</td>
<td>67</td>
<td>N/A</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>67</td>
<td>402</td>
</tr>
<tr>
<td>Rhode island</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>273</td>
</tr>
<tr>
<td>South Carolina</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>46</td>
<td>322</td>
</tr>
<tr>
<td>South Dakota</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>66</td>
<td>462</td>
</tr>
<tr>
<td>Tennessee</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>95</td>
<td>665</td>
</tr>
<tr>
<td>Texas</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>254</td>
<td>1,778</td>
</tr>
<tr>
<td>Utah</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>29</td>
<td>203</td>
</tr>
<tr>
<td>Vermont</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>14</td>
<td>98</td>
</tr>
<tr>
<td>Virginia</td>
<td>136</td>
<td>136</td>
<td>135</td>
<td>135</td>
<td>135</td>
<td>134</td>
<td>134</td>
<td>945</td>
</tr>
<tr>
<td>West Virginia</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>55</td>
<td>385</td>
</tr>
<tr>
<td>Wyoming</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>N/A</td>
<td>23</td>
<td>23</td>
<td>23</td>
<td>138</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>3,020</td>
<td>2,856</td>
<td>2,686</td>
<td>2,741</td>
<td>2,859</td>
<td>2,859</td>
<td>2,860</td>
<td>19,881</td>
</tr>
</tbody>
</table>
Table 3-4: Correlations

<table>
<thead>
<tr>
<th>Variable</th>
<th>Turnout</th>
<th>Registration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deadline</td>
<td>-.340</td>
<td>-.090</td>
</tr>
<tr>
<td>Age</td>
<td>.314</td>
<td>.325</td>
</tr>
<tr>
<td>College</td>
<td>.179</td>
<td>.033</td>
</tr>
<tr>
<td>Income</td>
<td>.124</td>
<td>.067</td>
</tr>
<tr>
<td>Anglo</td>
<td>.328</td>
<td>.238</td>
</tr>
</tbody>
</table>

Table 3-5: Results for Model of Turnout

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
<th>5%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>-.0032</td>
<td>.00013</td>
<td>-.0034</td>
<td>-.0033</td>
<td>-.0032</td>
<td>-.0031</td>
<td>-.003</td>
</tr>
<tr>
<td>Age</td>
<td>.0045</td>
<td>.00029</td>
<td>.004</td>
<td>.0043</td>
<td>.0045</td>
<td>.0047</td>
<td>.005</td>
</tr>
<tr>
<td>Income</td>
<td>-8.2 e^-7</td>
<td>2.2 e^-7</td>
<td>1.2 e^-6</td>
<td>9.6 e^-7</td>
<td>8.1 e^-7</td>
<td>6.7 e^-7</td>
<td>4.7 e^-7</td>
</tr>
<tr>
<td>Anglo</td>
<td>.15</td>
<td>.0076</td>
<td>.13</td>
<td>.14</td>
<td>.15</td>
<td>.15</td>
<td>.16</td>
</tr>
<tr>
<td>Intercept</td>
<td>.3</td>
<td>.009</td>
<td>.28</td>
<td>.29</td>
<td>.3</td>
<td>.3</td>
<td>.31</td>
</tr>
</tbody>
</table>

Table 3-6: Results for Model of Registration

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.D.</th>
<th>5%</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Close</td>
<td>-.0006</td>
<td>.0001</td>
<td>-.0007</td>
<td>-.0006</td>
<td>-.0006</td>
<td>-.0005</td>
<td>-.0004</td>
</tr>
<tr>
<td>Age</td>
<td>.0088</td>
<td>.00024</td>
<td>.0084</td>
<td>.0087</td>
<td>.0088</td>
<td>.0090</td>
<td>.0092</td>
</tr>
<tr>
<td>Income</td>
<td>-3.8 e^-7</td>
<td>1.9 e^-7</td>
<td>6.9 e^-7</td>
<td>5.0 e^-7</td>
<td>3.7 e^-7</td>
<td>2.5 e^-7</td>
<td>6.9 e^-8</td>
</tr>
<tr>
<td>Anglo</td>
<td>.092</td>
<td>.0054</td>
<td>.083</td>
<td>.089</td>
<td>.092</td>
<td>.096</td>
<td>.101</td>
</tr>
<tr>
<td>Intercept</td>
<td>.41</td>
<td>.0087</td>
<td>.41</td>
<td>.42</td>
<td>.42</td>
<td>.43</td>
<td>.44</td>
</tr>
</tbody>
</table>
Table 3-7: Reasons for Not Voting (CPS)

<table>
<thead>
<tr>
<th>Main reason for not voting</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Illness or disability (own or family member’s)</td>
<td>20.25</td>
</tr>
<tr>
<td>Out of town or away from home</td>
<td>12.08</td>
</tr>
<tr>
<td>Forgot to vote (or send in absentee ballot)</td>
<td>7.05</td>
</tr>
<tr>
<td>Not interested, felt vote wouldn’t count</td>
<td>11.27</td>
</tr>
<tr>
<td>Too busy, conflicting work or school schedule</td>
<td>22.42</td>
</tr>
<tr>
<td>Transportation problems</td>
<td>3.05</td>
</tr>
<tr>
<td>Didn’t like candidates or campaign issues</td>
<td>6.83</td>
</tr>
<tr>
<td>Registration problems</td>
<td>3.87</td>
</tr>
<tr>
<td>Bad weather conditions</td>
<td>1.44</td>
</tr>
<tr>
<td>Inconvenient hours, polling place</td>
<td>1.75</td>
</tr>
<tr>
<td>Other</td>
<td>9.98</td>
</tr>
</tbody>
</table>

Table 3-8: Reasons for Not Registering (CPS)

<table>
<thead>
<tr>
<th>Main reason for not being registered</th>
<th>Percent (Total)</th>
<th>Percent (EDR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Did not meet registration deadlines</td>
<td>14.74</td>
<td>8.96</td>
</tr>
<tr>
<td>Did not know where or how to register</td>
<td>5.09</td>
<td>5.94</td>
</tr>
<tr>
<td>Did not meet residency requirements</td>
<td>4.28</td>
<td>3.39</td>
</tr>
<tr>
<td>Permanent illness or disability</td>
<td>5.00</td>
<td>4.96</td>
</tr>
<tr>
<td>Difficulty with English</td>
<td>1.03</td>
<td>.58</td>
</tr>
<tr>
<td>Not interested in the election</td>
<td>43.58</td>
<td>46.10</td>
</tr>
<tr>
<td>My vote would not make a difference</td>
<td>3.56</td>
<td>4.06</td>
</tr>
<tr>
<td>Not eligible to vote</td>
<td>5.38</td>
<td>3.31</td>
</tr>
<tr>
<td>Other reason</td>
<td>17.34</td>
<td>22.69</td>
</tr>
</tbody>
</table>

Table 3-9: Method of Registration (CPS)

<table>
<thead>
<tr>
<th>How did you register to vote?</th>
<th>Percent (Total)</th>
<th>Percent (EDR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>At a DMV</td>
<td>24.19</td>
<td>6.08</td>
</tr>
<tr>
<td>At a public assistance agency</td>
<td>1.68</td>
<td>.37</td>
</tr>
<tr>
<td>Registered by mail</td>
<td>12.45</td>
<td>2.37</td>
</tr>
<tr>
<td>At a school, hospital, or on campus</td>
<td>7.53</td>
<td>5.15</td>
</tr>
<tr>
<td>Went to county government office</td>
<td>29.48</td>
<td>33.90</td>
</tr>
<tr>
<td>Registration drive</td>
<td>8.67</td>
<td>5.23</td>
</tr>
<tr>
<td>Registered at a polling place</td>
<td>9.67</td>
<td>38.40</td>
</tr>
<tr>
<td>Other</td>
<td>6.34</td>
<td>8.51</td>
</tr>
</tbody>
</table>
Table 3-10: Matched Sample Analyses for Turnout

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deadline</td>
<td>-.009 (.001)</td>
<td>-.012 (.002)</td>
<td>-.012 (.002)</td>
<td>-.012 (.002)</td>
<td>-.012 (.002)</td>
</tr>
<tr>
<td>Age</td>
<td>.034 (.001)</td>
<td>.035 (.001)</td>
<td>.038 (.001)</td>
<td>.039 (.001)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>.117 (.044)</td>
<td>.012 (.044)</td>
<td>.077 (.045)</td>
<td>.077 (.045)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>.473 (.018)</td>
<td>.380 (.018)</td>
<td>.382 (.019)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>.149 (.006)</td>
<td>.102 (.007)</td>
<td>.101 (.007)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>.094 (.082)</td>
<td>.022 (.081)</td>
<td>.181 (.084)</td>
<td>.164 (.085)</td>
<td></td>
</tr>
<tr>
<td>Latino</td>
<td>-.505 (.086)</td>
<td>-.354 (.088)</td>
<td>-.304 (.089)</td>
<td>-.349 (.089)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>-.609 (.132)</td>
<td>-.610 (.135)</td>
<td>-.683 (.136)</td>
<td>-.721 (.137)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>-.007 (.173)</td>
<td>-.052 (.173)</td>
<td>.092 (.176)</td>
<td>.046 (.178)</td>
<td></td>
</tr>
<tr>
<td>Presidential</td>
<td></td>
<td></td>
<td></td>
<td>.555 (.05)</td>
<td></td>
</tr>
<tr>
<td>Competitiveness</td>
<td></td>
<td>.878 (.216)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>.513 (.031)</td>
<td>-2.43 (.103)</td>
<td>-2.36 (.095)</td>
<td>-3.21 (.112)</td>
<td>-3.85 (.145)</td>
</tr>
</tbody>
</table>

N= 10,000

Table 3-11: Matched Sample Analyses for Registration

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deadline</td>
<td>-.016 (.001)</td>
<td>-.010 (.001)</td>
<td>-.010 (.002)</td>
<td>-.010 (.002)</td>
<td>-.010 (.002)</td>
</tr>
<tr>
<td>Age</td>
<td>.013 (.001)</td>
<td>.014 (.001)</td>
<td>.015 (.001)</td>
<td>.015 (.001)</td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>.158 (.042)</td>
<td>.113 (.042)</td>
<td>.138 (.042)</td>
<td>.133 (.042)</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>.257 (.017)</td>
<td>.224 (.018)</td>
<td>.214 (.018)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>.059 (.006)</td>
<td>.037 (.006)</td>
<td>.036 (.006)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>.065 (.073)</td>
<td>.056 (.073)</td>
<td>.112 (.074)</td>
<td>.122 (.074)</td>
<td></td>
</tr>
<tr>
<td>Latino</td>
<td>-.421 (.077)</td>
<td>-.352 (.078)</td>
<td>-.328 (.078)</td>
<td>-.29 (.078)</td>
<td></td>
</tr>
<tr>
<td>Asian</td>
<td>-.824 (.132)</td>
<td>-.896 (.134)</td>
<td>-.931 (.134)</td>
<td>-.87 (.135)</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>.142 (.154)</td>
<td>.176 (.154)</td>
<td>.214 (.155)</td>
<td>.233 (.156)</td>
<td></td>
</tr>
<tr>
<td>Presidential</td>
<td></td>
<td></td>
<td></td>
<td>-.257 (.049)</td>
<td></td>
</tr>
<tr>
<td>Competitiveness</td>
<td></td>
<td>-1.40 (.183)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>1.44 (.03)</td>
<td>-1.24 (.087)</td>
<td>-1.42 (.083)</td>
<td>-1.71 (.095)</td>
<td>-1.02 (.118)</td>
</tr>
</tbody>
</table>

N= 10,000

Table 3-12: Substantive Effects from Model 1

<table>
<thead>
<tr>
<th>Predicted Probabilities</th>
<th>0 Days</th>
<th>30 Days</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turnout</td>
<td>.626</td>
<td>.562</td>
<td>.064</td>
</tr>
<tr>
<td>Registration</td>
<td>.808</td>
<td>.724</td>
<td>.084</td>
</tr>
</tbody>
</table>
Table 3-13: Effect of EDR on Turnout from County Data, 1992-2004

<table>
<thead>
<tr>
<th></th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
<th>Model 4</th>
<th>Model 5</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDR</td>
<td>0.102</td>
<td>-0.021</td>
<td>-0.01</td>
<td>-0.001</td>
<td>0.007</td>
<td>-0.0005</td>
<td>0.014</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Deadline</td>
<td>-0.005</td>
<td>-0.004</td>
<td>-0.003</td>
<td>-0.003</td>
<td>-0.003</td>
<td>-0.003</td>
<td>-0.003</td>
</tr>
<tr>
<td></td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td>(0.0001)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>0.01</td>
<td>0.007</td>
<td>0.009</td>
<td>0.009</td>
<td>0.01</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0002)</td>
<td>(0.0002)</td>
<td>(0.0002)</td>
<td>(0.0002)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% College Degree</td>
<td>0.308</td>
<td>0.287</td>
<td>0.634</td>
<td>0.6</td>
<td>0.668</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td>(0.01)</td>
<td>(0.017)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Anglo</td>
<td>0.114</td>
<td>0.12</td>
<td>0.111</td>
<td>0.129</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td>(0.005)</td>
<td>(0.005)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td>-6.8e-6</td>
<td>-7.3e-6</td>
<td>-6.3e-6</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(2.67e-7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>0.494</td>
<td>0.616</td>
<td>0.188</td>
<td>0.174</td>
<td>0.141</td>
<td>0.125</td>
<td>0.157</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.003)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N=</td>
<td>21759</td>
<td>21759</td>
<td>21743</td>
<td>21071</td>
<td>21071</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3-14: Synthetic Control Weights

<table>
<thead>
<tr>
<th>Treated</th>
<th>Control</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wyoming</td>
<td>Connecticut</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>Massachusetts</td>
<td>.663</td>
</tr>
<tr>
<td></td>
<td>Vermont</td>
<td>.332</td>
</tr>
<tr>
<td></td>
<td>Iowa</td>
<td>.001</td>
</tr>
<tr>
<td>New Hampshire</td>
<td>Connecticut</td>
<td>.75</td>
</tr>
<tr>
<td></td>
<td>Florida</td>
<td>.25</td>
</tr>
<tr>
<td>Idaho</td>
<td>Montana</td>
<td>.556</td>
</tr>
<tr>
<td></td>
<td>Massachusetts</td>
<td>.444</td>
</tr>
</tbody>
</table>
Chapter 4:
Early Voting

4.1 Introduction

Among the reforms to voting procedures, early voting is a particularly recent and significant change. There has been a proliferation of early voting laws in the past 15 years, with some sources putting the number of states with some form of early voting at 31 (electionline.org). Not only has early voting taken off among state election officials, but voters are also increasingly casting their ballots early. According to the ANES, in the 2006 Congressional elections more than one in five voters cast their ballot prior to Election Day.

Early voting is one of the most significant election reforms, in that in a dynamic political setting it fundamentally alters the way that elections are held. In the absence of early voting, the procedures are clear. Qualified electors are allowed to cast their ballots on Election Day. That is, at the conclusion of the campaigns, voters are thought to weigh the characteristics of the candidates and choose their most preferred candidate. The timing of the election is not inconsequential. With Election Day voting there is an end to the campaigns that is designated a priori. The outcome of the election depends on voters’ evaluations at that time. Early voting complicates elections. On every day of early voting, individuals are permitted to cast ballots that represent their preferences at that time. But we know that campaigns are dynamic, so that candidates’ positions might shift, their electoral fortunes might change, they might drop out of the race, and so on. This complicates elections because it no longer reflects the attitudes of voters at a particular time (the designated end of the
campaigns). Rather it reflects the aggregation of voters’ expressed preferences over a longer period of time.

There are several related questions we can consider about early voting. The first is whether or not early voting increases turnout. The added convenience of voting might make it more attractive to voters than Election Day voting. If there is an increase in early voters, we might consider whether or not this affects their campaign choices. Campaigns are dynamic and the information about the candidates can change over time, so we might also consider the stability of individuals’ preferences over time. If voters’ preferences change over time, then early voting might also be consequential in that how an individual would vote prior to Election Day may not be the same as how they would vote on Election Day. This raises the possibility that early voting might not only increase the likelihood that an individual votes, but could also affect how they vote. It also complicates the interpretation of the election results as it is no longer solely indicative of popular support at the conclusion of the campaigns but rather represents popular support as measured at various points in the campaign. Whether or not these factors are of importance for turnout, election outcomes, or individuals voting patterns is not clear from the existing research on early voting.

This chapter analyzes how early voting affects rates of voter turnout and if there are conditional effects of early voting. The argument put forward in this chapter is that individuals will self-select into early voting based on the strength of their preferences for the candidates. Individuals with stronger preferences will be more likely to vote early as their preferences are unlikely to change as a result of campaign dynamics, while electors with preferences that are not as strong will be more likely to wait until Election Day, when the campaigns have concluded. The next section of the paper reviews the extant literature on
early voting. I then discuss theoretical mechanisms linking early voting to turnout, and test the hypotheses in the research design section. In the last part of the chapter I review the main results and discuss possible directions for future research.

4.2 Literature

Early voting was a reform first adopted in Texas in 1988 that allows individuals to vote in-person prior to an election at special voting sites. Data from one of the first elections that offered early voting, the 1992 presidential election, suggests that early voting might have a positive effect on levels of turnout and the representativeness of the electorate. In particular, early voting is shown to be related to new registered voters and the size of the Latino population in the county (Stein and Garcia-Monet 1997). This suggests that early voting might attract individuals that are otherwise less likely to vote and could increase turnout and make the electorate more representative of the general population. They also find that these effects might be conditional on party activities, as the Clinton campaign attempted to mobilized potential supporters to vote early (Stein and Garcia-Monet 1997). Subsequent research using data from the 1994 midterm elections in Texas finds that individuals that are highly likely to vote are also more likely to vote early. That is, individuals that are stronger partisans, more interested in politics, and older are more likely to vote early than individuals that are younger, less interested in politics, and less partisan (Stein 1998). This might suggest that early voting will have less of an independent effect of mobilizing new voters, but could nevertheless have a joint effect with partisan mobilization.

These results are consistent with the findings from a survey of voters in Knox County, Tennessee in the 1996 presidential elections (Neely and Richardson 2001). Neely and
Richardson (2001) find that individuals that vote early are more likely to be concerned about politics and attentive to the campaigns, but not necessarily more partisan or better educated. They conclude that there is little evidence that early voting leads to higher turnout or that it further biases the electorate (Neely and Richardson 2001). One possible concern about the study is that they sample likely voters without describing how they choose the sample. It is also not clear if they verify respondents' self-reported voting behavior, but they report an 89% turnout rate in their sample, when the county-level turnout among registered voters was 69%.

While the existing empirical results seem to suggest that the effects of early voting might be conditional on party mobilization or on individual characteristics such as attentiveness to politics, important questions seem to remain about the effects of early voting in more recent elections particularly given the increase in the number of states that allow for early voting. Further research might not only help to better generalize the existing results but might also help to answer the apparent empirical puzzle of the spread of early voting despite clear indications that it is effective at increasing turnout.

4.3 Theory

The theoretical discussion identifies two processes that are central to the effects of early voting. The first is convenience, which suggests that early voting should increase turnout. The second is the informational tradeoff associated with early voting (early voters forgo opportunities to acquire additional information as the campaigns are still ongoing), which suggests possible conditional effects of early voting.
4.3.1 Definition: Early voting

For the purpose of this study, early voting means one-stop, no-excuse, in-person voting prior to Election Day. From this definition, a state has early voting if it allows any qualified elector to vote in-person before Election Day. This definition does not include various forms of absentee voting, such as those that require an excuse to vote absentee, absentee mail-in ballots, or a process that is not one-stop. This definition does not constrain policies that allow for the establishment of satellite early voting sites, extended hours, or the length of the early voting period. As with previous chapters, voter turnout at the aggregate level is taken to be the proportion of the eligible electorate that casts a ballot in an election and for individuals reflects whether they voted in a given election.

4.3.2 Early Voting and Voter Turnout

There are two factors to consider in the relationship between turnout and early voting. The first is convenience, and the second is information about the candidates. This section considers each of the factors in turn. While the added convenience of early voting suggests that it should increase voter turnout, early voting also takes away opportunities for voters' to acquire information about the candidates and thereby creates a cross-pressure that might offset some of the effects of convenience. Greater convenience suggests that early voting should increase voter turnout, while information effects might create other conditional effects.

The first way in which early voting might be related to voter participation is by making voting more convenient. As discussed in the previous chapter there are a number of reasons to anticipate that convenience voting will increase turnout. By lowering the barriers and the costs to participation, early voting might thereby increase turnout rates.
There are several ways in which early voting might be a more convenient mode of voting. The primary way is that early voting allows individuals to cast their ballots in person for several days to several weeks prior to Election Day. By offering voters the opportunity to cast ballots for a number of days, rather than only between working hours on Election Day, early voting might be more convenient as voters’ can choose to cast their ballot when the opportunity cost of expending the time and effort to go to the polls is lower.

In addition to providing individuals with greater opportunities to vote on different days, early voting also allows individuals to vote at multiple polling locations. Just as individuals may be busy on Election Day which creates a barrier to participation, they may also be away from their residential polling location. While precinct-based Election Day voting requires individuals to go their designated polling location, early voting sites are open to all individuals in a county. Having the choice of multiple polling locations might also lower the costs of voting, particularly for mobile voters that commute longer distances and if early voting sites are more centrally located. The other benefits of centralization such as having more equipment, better trained poll workers, and larger facilities to accommodate a greater number of voters that were discussed in the previous chapter may also apply to early voting. By linking early voting to the costs and convenience of participation we should expect that the policy is positively related to rates of voter turnout. From these arguments we would also expect there to be a positive relationship between the length of early voting and the number of early ballots cast. This hypothesis follows from early voting increasing the convenience and accessibility of voting. The more opportunities individuals have, the more early voters we should observe.
We may also anticipate that the effects of early voting vary based on the characteristics of individual voters. In particular the timing of individuals’ vote choice seems to be particularly important for whether or not they are more likely to turnout as a result of early voting. While early voting may be more convenient and attractive for early deciders, voters that make their vote choices later in the election cycle might be less responsive to early voting. In order to effectively increase turnout, the early voting period should correspond to when voters have decided on which candidate to support. Late deciders may be less affected by early voting as the early voting period may precede the timing of their choices. By analogy, prompt service at a restaurant is not necessarily better if the customers are not given time to decide on their orders. Thereby early voting might have greater consequences for how early deciders vote (although we could also argue that by allowing early deciders to vote away from Election Day, polling places might be less congested and thereby indirectly make voting easier for late deciders).

To further explore this possibility, I analyze a model of voter decision-making. Ideally we would like to analyze a full stopping model of voter decision making to derive when voters will stop seeking new information about the candidates and choose whom to support. While a fully specified stopping model is left for future research I do consider the initial parts of a full theoretical stopping model to tentatively assess the timing of vote choices.

An assumption of the model considered here is that voters’ decisions are uncertain. It is assumed that they do not have complete information as to which of the candidates would perform better if elected, but rather form (uncertain) beliefs. To represent the voters’
uncertainty regarding the candidates, I assume that their beliefs follow a normal probability distribution for each of the candidates (with different mean and variance parameters).

If voters knew the parameters of the distributions, they could just choose the candidate with the highest expected value. Yet it is unlikely that voters will know the parameters governing the quality of the candidates. While voters may not be fully informed as to the parameters of the distribution they are also not totally uninformed as to how the candidates might perform if elected. Although candidates are often less than candid and sometimes make ambiguous statements, voters can assess information from the campaigns to update their beliefs. After obtaining information about the candidates they decide whom to support.

To capture the information acquisition and updating processes, I use Bayes’ rule to combine both individuals’ prior beliefs about the candidates as well as information they acquire from the campaigns. Other studies in political science have used Bayesian models to examine the role of information and experiences in the development of other types of opinions and attitudes (Bartels 1993, Alvarez and Brehm 2002, Green, Palmquist, Schickler 2002, Achen 2006). The likelihood component of the Bayesian model follows a normal distribution as just discussed. The priors are assumed to be distributed normal inverse gamma. The normal inverse gamma distribution is conjugate with the normal likelihood, which means that the posterior distribution will also be distributed normal inverse gamma. We can then consider the marginal posterior distribution for the expected values (of the distributions for the candidates). This probability distribution represents the voters’ beliefs as to which candidate has a higher expected value. When there is more separation, the choice
is assumed to be clearer to the voter and she is thereby less likely to change her preference as a result of subsequent campaigning.

When the posteriors show a substantial amount of overlap, however, the voter is assumed to be more tentative in her vote choice. With greater overlap of the posteriors the difference between the candidates is less clear and a voter might be more likely to seek additional campaign information than when there are large differences. Voters’ uncertainty over the candidates might in turn be affected by a number of different factors. In the sections below I present results from the normal inverse gamma posterior regarding the effects of voters’ priors, the difference between the candidates, the variability of campaign information, and the amount of campaign information. The simulations were all conducted using MCMC. These results suggest tentative empirical hypotheses as to the timing of individuals vote choices and possible conditional effects of early voting.

The first parameter that I vary is the strength of the priors. Here, the strength of the priors refers to the variance of the prior distribution of $\mu_1$ and $\mu_2$. I set this parameter at 250 different values and record the posterior probability that $\mu_1$ is greater than $\mu_2$. The results are shown in Figure 4-1. We can see from this figure that there is a positive and diminishing relationship between the variance of the priors and the amount of overlap for the full range of the graph. That is, more informative priors (a lower variance) are associated with more distinct posteriors. The probability increases with the variance of the prior but seems to plateau at higher variance values. These results suggest that individuals will have stronger preferences when they have stronger prior beliefs about the candidates, than when they have few prior beliefs about the candidates.
The second parameter that I consider in the model is the distance between the candidates, again ranging over 250 different values. We would expect that the strength of individuals’ beliefs would be greater if the candidates were very different from one another than if they were more similar. Consistent with these expectations, we can see in Figure 4-2 that there is an inverse relationship between the distance between the candidates and the strength of voters’ beliefs. As shown in Figure 4-2, the probability falls off rapidly as there are greater differences between the candidates.

In addition to the variation between the candidates, we can also consider variation for the candidates themselves. Campaigns often make ambiguous or uncommitted statements when campaigning, so that there might be more variation in the information that voters’ receive about the candidates. To consider how the variation in the information voters’ receive about the candidates affects the probability of switching, I set the information variance parameter at a number of different levels and tracked the probability of switching. The results shown in Figure 4-3 shows that the probability of switching is low and consistent at low levels of variation. As the variation increases, however, the probability of switching both increases and becomes much more diffuse. Whereas switching is consistently low when there is little variance, as the variance increases the probability of switching increases and also shows larger fluctuations.

Voters might also have differential access to information for a variety of reasons. The amount of information that voters have about the candidates was also varied in the model, and the results are plotted in Figure 4-4. These results show that as the amount of information increases the probability of switching decreases. At particularly low levels of information individuals are susceptible to switching whereas with greater access to
information that probability declines. This suggests that in a more information rich environment, individuals will be less likely to switch their candidate preferences.

4.3.3 Hypotheses

The theoretical results suggest a number of empirical hypotheses. The first is that early voting should increase voter participation overall by lowering the costs of voting.

**H1a:** Early voting increases voter turnout.

**H1b:** A longer early voting period increases the proportion of early voters.

The first hypothesis follows from the initial theoretical discussion as to how early voting may increase the convenience of voter participation and thereby increase turnout. These arguments also lead to the second hypothesis that a longer early voting period should lead to more early voters.

**H2:** Early voting should have a greater effect for strong partisans.

The second hypothesis follows from the effects of individuals’ priors on the strength of their candidate preferences. Stronger partisans are thought to have a stronger prior for the candidates, whereas less partisan voters might start with weaker beliefs about the candidates and instead rely more heavily on specific campaign information. This might make less partisan voters more sensitive to the informational costs of early voting.

**H3:** The effect of early voting should be stronger for individuals that are more attentive to politics.
The third hypothesis follows from the analysis of individuals’ exposure to campaign information. As we saw in the theoretical section, individuals with greater exposure to political information will have stronger preferences regarding the candidates and might thereby be more likely to vote early.

4.4 Research Design

This section describes the research design used to evaluate the empirical hypotheses from above. I first discuss the general effects of early voting on voter participation to assess whether there is evidence that early voting increases turnout in general. The second part considers whether there are subpopulation effects of early voting.

The first section presents two sets of results, one analyzing aggregate data the other analyzing individual data. The results from these analyses are consistent and suggest that early voting seems to have a negative effect on turnout, contrary to expectations. While inconsistent with the arguments from above, the empirical analyses seem suggest that early voting decreases participation. Possible explanations for this result are considered in the conclusion to this chapter.

The second section analyzes individual-level data to assess possible subpopulation effects. The results presented in this section suggest that early voting has a greater effect on partisans and the politically engaged, but the overall effect of the reform remains negative.

4.4.1 General Effects

There are three sources of data used to evaluate the hypothesis that early voting should increase voter turnout. The first were collected at the state-level on whether or not the state
has early voting, the second provides information on turnout rates that were collected for a large number of counties, and the third is survey data from the ANES. Each data source is briefly described below.

**Early Voting** The main independent variable is whether or not the state offers early voting. While the concept seems straightforward, a consequence of having decentralized elections is that states may adopt a range of different policies regarding voting opportunities. Thus we need to define what is meant by early voting to be able to correctly classify state election laws. As before, early voting allows for one-stop, no-excuse, in-person voting prior to Election Day. That is, early voting is taken to be the opportunity for voters to appear in-person at a designated polling site, prior to Election Day, and cast a ballot on-site. This is distinct from states that offer mail-in early voting, restricted in-person absentee voting, or drop-off sites for ballots that must first be requested or distributed by mail.

Following this definition, each state was coded for whether or not it offered early voting. The variable was coded a 1 if state laws allow early voting and as 0 otherwise. Data for this variable were coded from electionline.org (2007) and state codes from Lexis-Nexis. Data were also recorded for the year in which the state first implemented early voting, the total number of days of early voting that is required, and the number of consecutive days up to and including Election Day in which early voting is unavailable.¹²

**Aggregate Turnout** The first source of information on voter turnout is taken from U.S. counties from 1992-2004. The data are the same as those described in the chapter on voter

¹² Some states allow for early voting opportunities to continue through Election Day. For those policies, the days prior variable is recorded as 0.
registration. These data contain observations for each county (or municipalities that administer their own elections) from 1992 to 2004, excluding Washington D.C. and Alaska. The measure of voter turnout in these counties is taken to be the ratio of the number of ballots cast over voting age population. Data from the number of ballots cast was recorded using official election returns supplemented with data provided by the America Votes Series (published biennially by CQ Press). In instances when information on the total number of ballots cast was not available vote totals for the top of the ticket race were recorded. Data for this variable were collected for over 98% of the possible observations (21,762 of 22,022).

**Individual Turnout** The other source of data on voter turnout is the American National Election Survey. Data from the ANES are analyzed from 1990 to 2006. The primary variable of interest is whether or not the individual reported voting. Additional information on whether the individual reported voting early is also considered, along with a number of control variables that are discussed below.

The first analysis considers whether the average rates of turnout differ between early voting and non-early voting counties. A difference of proportions test provides an initial assessment of the differences between the groups. While informative, the difference of proportions test does not adjust for potentially confounding variables and might not represent the causal effect of early voting on rates of participation.

The results from the test of proportions are shown in Error! Reference source not found. From this table we can see that the initial evidence seems to run counter to the first hypothesis. The average rate of turnout in counties with early voting actually seems to be lower than in counties that do not have early voting. These results suggest that early voting
might actually have a suppressive effect on turnout, contrary to the theoretical expectations from above.

To adjust for potentially confounding variables a number of statistical models were estimated. I chose to model voter turnout in the counties using a normal-linear model. The reason that I chose to use a linear model is that rates of voter turnout appear to be roughly normally distributed, linear regression models are particularly robust to violations of modeling assumptions, and the linear model seemed to fit the data reasonably well. Robust standard errors were also estimated to adjust for heteroskedasticity and potential non-conditionally-independent observations.

The control variables included in the model are median age, median years of education, percent Anglo (of voting age population), a dichotomous indicator for presidential election years, and a measure of the competitiveness of the top of the ticket race. These models improve on the difference of means test as they help to adjust for intervening variables and should provide a more reliable estimate of the causal effect of early voting on turnout.

The results from the model including the control variables are shown in Table 4-2. From this table, we can see that turnout remains lower in counties that have early voting. The results from the model can be interpreted as marginal effects, and we can see that the magnitude of the effect from the statistical model is similar to the difference of means test. That is, early voting appears to decrease participation by about 2.2%, with a 90% interval between 2.0 and 2.4 (both negative values suggesting a decrease in participation). The
difference of means test as well as the model estimates seems to show that early voting decreases turnout by around 2.5%.\textsuperscript{13}

In addition to the analysis of county data, I also analyze individual data. These data are taken from the ANES. After combining these data with the measure of early voting, we can compare the rates of turnout to individuals with and those without early voting.

To control for potentially confounding variables in the individual analysis I also consider a matched sample. Matching provides an alternative to model-based adjustments to account for confounding variables. Generally, the procedure works by identifying like observations in the treatment and control samples for comparison. The combined sample of similar cases across the treatment and control groups are then compared to assess any differences in the responses. The advantages of matching are that it has been shown to reduce dependence of the results on arbitrary modeling assumptions, provide estimates that are closer to experimental benchmarks, reduces the problem of model-hunting, and provides a more intuitive means of conveying the design and results.

While matching has been shown to offer a number of advantages over traditional model-based alone, the analyses in this section seeks to control for a number of confounding variables. As the number of variables increases, along with the number of possible values of each variable, the number of possible combinations of the control variables grows rapidly. With many different possible combinations of the control variables it is generally not possible to find exact matches. To address this issue, I use propensity score matching which

\textsuperscript{13} The results are robust to a number of alternative specifications. Specifically, I also estimated a reduced model with only early voting, reduced and full models using citizen voting age population as the outcome, full and reduced models using turnout as a ratio of registered voters, and each of those models comparing early voting to non-early voting states as well as analyzing early voting states only (over time). In all of the models, the estimated effects are negative, significant at conventional levels, and have similar magnitudes (from 2-4%). The results from these models are included in the appendix.
reduces the number of potentially confounding variables into a single continuous propensity score on which observations can then be matched.

In this case, the propensity scores are obtained by first estimating a logit model of exposure to early voting using the control variables. The propensity scores are then created by returning the predicted probabilities of exposure to treatment. The estimated propensity scores are then assigned to every observation in the data, and the treatment subjects are matched to the nearest control subject for analysis. The unmatched controls are discarded from the analysis as they may not have a clear analogue in the treated data set.

The analysis of the matched cases proceeds in several steps. The first is to evaluate the quality of the match. This is done by comparing the balance improvement between the treatment and control groups in the matched and unmatched data sets. As we can see in Table 4-3, the matched data are generally better balanced on the control variables than the unmatched data. I also report the distribution of estimated propensity scores for the matched and unmatched data, as shown in the Q-Q plots in Figure 4-5. We would like the points to track as closely to the diagonal line as possible, indicating a similar distribution of propensity scores for the treatment and control groups. While each graph departs from the line somewhat, particularly at higher values of the propensity scores, we can see that in general the matched sample tracks the line closer than does the unmatched data.

Using these data, I then compare the proportion of turnout for the early voting and non-early voting groups. The table is reported along with a test of significance. The results are shown in Table 4-4. From these results we can see that turnout was actually 3.6% greater in non-early voting states. The null hypothesis of no effect can also be rejected at conventional levels of significance as reported by the p-value in the table.
We can see from the balance statistics, however, that propensity score matching has not removed all of the observed differences from the sample. To further account for the remaining differences between the groups (other than exposure to early voting) I also report the results of several statistical models.

The estimates are reported in Table 4-5. Using the matched sample, results are reported for model estimates that include only the treatment variable, one that also includes the propensity score distances, and a third that includes the full set of control variables. From these results we can see that the estimated effect of the treatment is consistently negative, and the 90% confidence interval for the third model ranges from -.263 to -.012. These results also suggest that early voting has a negative effect on turnout.

While the model estimates suggest that early voting is negatively related to turnout, the coefficients reported in Table 4-5 cannot be interpreted as marginal effects of the treatment variable. The reason is that the expected value of the dependent variable is modeled as a non-linear function of the explanatory variables. The marginal effect of the treatment variable is reported in Table 4-6. This figure reports the expected probability that a typical individual would vote under treatment and control conditions, along with a 90% confidence band. We can see from this figure that the expected probability of voting is around 1.6% greater for individuals not exposed to early voting.

The results from the individual analyses and the aggregate analyses are generally consistent. Across methods of analysis and sources of data, early voting seems to be lead to a 2-3% decrease in voter turnout. This finding contradicts the theoretical argument presented in this chapter and in previous research on early voting. There is not an obvious explanation
for these results, but the consistency of the results suggests that we should reconsider the theoretical argument. Possible areas for future research are discussed in the conclusion.

4.4.2 Subpopulation Effects

In addition to an overall effect on turnout the theoretical argument also suggests subpopulation effects. In particular, the arguments suggest that early voting should be more efficacious for more partisan and politically engaged voters. The results in this section suggest that while there do seem to be differential subpopulation effects as expected, the overall effects are still negative.

The data used to test for subpopulation effects are the same matched sample data as those used above. The primary treatment and outcome variables are also the same. In order to assess the possible conditional effects of early voting on turnout I incorporate several interaction terms in separate models of turnout.

In order to test the third hypothesis, I interact early voting with partisanship. Partisanship is measured on a four-point scale with higher values representing stronger partisans. The lowest value of partisanship represents independents with no political leanings and ranges up to individuals that identify as strong partisans (either Republicans or Democrats). The expectation is that the interaction between early voting and partisanship should be positive indicating that the effects of early voting are greater for stronger partisans. The results are shown in Table 4-7.

As we can see in Table 4-7, the coefficient estimate for the interaction between early voting and partisanship is positive as expected. The 90% confidence interval also indicates that the estimate does not overlap with zero suggesting that early voting does have a positive
conditional relationship with partisanship. However, we can also see that the overall treatment effect is still negative, suggesting that the negative direct effect might outweigh the positive conditional effects even for strong partisans. To assess the substantive effects, simulation results are presented in Table 4-8. Consistent with the estimate, we can see that the effect of early voting increases for stronger partisans, but only at the highest value for partisanship does the point estimate become marginally positive. While these results are consistent with the third hypothesis, there still seems to be an overall negative effect of early voting on turnout from these results.

To assess the fourth hypothesis, we can consider the variable that measures individuals’ interest in the campaigns. The ANES includes about how interested the respondents were in the campaigns (not much interested, somewhat interested, and very interested). I interact the interest variable with early voting to assess whether political interest conditions the effect of early voting. The numerical values of the variable increase as the respondent expresses greater interest, and we should thereby expect a positive conditional relationship as the reform should be more efficacious for individuals with greater interest.

The results from the analysis are presented in Table 4-9. From these estimates we can see that the interaction between early voting and voter interest is positive, as expected. In the first two models we can see that the coefficient estimate is positive and the standard error is relatively small. In the third model the estimate is still positive but the standard error is greater and the confidence interval overlaps substantially with negative values. While the conditional relationship is most likely positive, based on these results we cannot rule out the possibility of a null or negative relationship. Substantive effects for the full model are also reported in Table 4-10. From this table we can see that the effects of early voting increase
for more interested voters, but also that most of the confidence intervals include negative values as well.

The results from these analyses suggest that early voting is more efficacious for partisan voters and perhaps for more interested voters as well. The results are also consistent with the overall negative effect of early voting on turnout. These empirical findings suggest that not only might early voting decrease turnout but may also exacerbate disparities between voters and non-voters. Insofar as early voting is more effective for likely voters, the reform might result in an electorate that is more partisan and politically interested than the population.

4.4.3 Discussion

The empirical findings presented in this chapter confirm parts of the theoretical argument and disconfirm others. The main hypothesis, regarding the effects of early voting on turnout, did not receive support in any of the analysis. Using data from both U.S. counties and the American National Election Study, early voting was consistently related to lower rates of turnout. The magnitude of the effect was also comparable across analyses and suggests that early voting is associated with around a 3% decrease in turnout. The effect is sustained across several data sources, units of observation, and methods of analysis.

This finding is even more puzzling in the context of the anecdotal evidence about the popularity of early voting and findings for the second hypothesis. Anecdotally early voting has increased in a number of states in recent years and has been popular in those states. The second hypothesis also considers whether longer early voting periods are associated with a greater number of early voters. The results show an increase in early voting with longer
early voting periods. These results seem to suggest that voters are in some ways responsive to early voting, yet the effects on turnout still appear to be negative.

The third hypothesis argues that more partisan voters might be more likely to vote early as they are expected to make vote choices sooner than less partisan voters. The results tentatively confirm the hypothesis. While the overall effect of early voting remains negative, we do see an increase in the treatment effect for more partisan voters. The fourth hypothesis suggests that the effects of early voting are conditioned by voters' interest in the campaigns. The findings again seem to tentatively confirm the hypothesis, as we observe an increase in the treatment effect for highly interested voters. The results again suggest a negative overall effect of early voting, but that might be greater for highly interested voters. These results should be taken as more suggestive, as the standard errors were particularly large relative to the point estimates. The effects were in anticipated direction, but a null or positive effect would also be plausible.

4.5 Conclusion

This chapter examines the effects of early voting on voter participation. The theoretical arguments presented in this section suggest that early voting should increase voter turnout, particularly for more partisan voters. While the results suggest that the effects are stronger for partisans the overall effects on turnout are negative. These findings are consistent across data sources and methods of analysis, and suggest that early voting decreases turnout by 2-3%.

There is no clear theoretical rationale that might account for a negative effect of early voting on turnout. The direct effect of the policy itself is to afford voters greater access to
polling facilities. By increasing the opportunities for individuals to vote, the theory suggest that early voting should increase participation yet the opposite appears to be occurring. There are a number of possible explanations for this result which we might consider and that may warrant future study.

One is that early voting might take away from administrative and campaign resources for Election Day. While partisans are more likely to take advantage of early voting the administration of early voting requires election officials to expend resources to provide early voting opportunities that might take away from resources available on Election Day. Given that early voting seems to mostly serve a targeted population of strong partisans that are already likely to vote, early voting might decrease turnout by diverting resources away from Election Day voting and thereby make participation less accessible and convenient at a time when it is most necessary, for less engaged voters.

Similarly, campaign resources might also be more diffused away from Election Day as a result of early voting. Given opportunities to vote early, campaigns may engage in GOTV activities prior to elections, particular if they target stronger partisans who might be more predictable supporters of the candidate. This might in turn take campaign resources away from GOTV activities targeting Election Day voters, which might also decrease turnout.

Another possible explanation might involve the symbolic utility of voting as discussed in the first chapter. One of the challenges for explanations of turnout that involve symbolic utility is how to define symbolic utility in a way that is conceptually distinct from the behavior, so the symbolic utility does not reduce to whether or not an individual voted (which is the outcome of interest). As discussed above, symbolic utility might be present when an individuals’ act comes to symbolize some other outcome such that the utility to the
individual of realizing that outcome gets instantiated by to the individuals’ act (Nozick 1995). In the context of an election, symbolic utility might be present if an individuals’ act of voting symbolizes a valued outcome such as full participation (or full contribution to the public good).

There are two parts to the process of symbolic utility to consider. The valuation of the outcome, such as a participatory democracy with full turnout, would not likely be affected by early voting. That is, it is not clear how early voting could affect people’s core political values. It might affect the instantiation process, however, if by spreading the election out over several days participation becomes more temporally diffuse and the act of voting become more individualistic and correspondingly less symbolic of wider social participation.

The results from this chapter confirm theoretical expectations regarding subpopulation effects of early voting showing that strong partisans and individuals with a high level of political interest are more likely to vote early than nonpartisans and individuals with lower political interest. Contrary to theoretical expectations, the results also suggest that there is a negative overall effect of early voting on turnout. This result is robust to numerous methods of analysis. Currently there is no clear explanation for this finding, and future research might consider possible effects of early voting on symbolic utility and resource allocations for the administration of early and Election Day voting by local officials as well as political campaigns.
### Table 4-15: Difference of Proportions

<table>
<thead>
<tr>
<th></th>
<th>Turnout / VAP</th>
<th>Turnout / CVAP</th>
<th>Turnout / Registered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-Early Voting</td>
<td>.514</td>
<td>.522</td>
<td>.637</td>
</tr>
<tr>
<td>Early Voting</td>
<td>.485</td>
<td>.500</td>
<td>.602</td>
</tr>
<tr>
<td>Difference</td>
<td>-.029</td>
<td>-.022</td>
<td>-.035</td>
</tr>
</tbody>
</table>

| 90% Confidence Interval | -.017, -.041 | -.011, -.034 | -.024, -.047 |

### Table 4-16: Turnout as a Ratio of VAP

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>S.E.</th>
<th>90% Confidence Interval</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Voting</td>
<td>-.022</td>
<td>.001</td>
<td>-.024, -.020</td>
<td>-.024</td>
<td>-.020</td>
</tr>
<tr>
<td>Age</td>
<td>.007</td>
<td>.000</td>
<td>.006, .007</td>
<td>.006</td>
<td>.007</td>
</tr>
<tr>
<td>Income</td>
<td>.000</td>
<td>.000</td>
<td>.000, .000</td>
<td>.000</td>
<td>.000</td>
</tr>
<tr>
<td>Anglo</td>
<td>.156</td>
<td>.004</td>
<td>.150, .162</td>
<td>.150</td>
<td>.162</td>
</tr>
<tr>
<td>Competitiveness</td>
<td>.172</td>
<td>.007</td>
<td>.160, .184</td>
<td>.160</td>
<td>.184</td>
</tr>
<tr>
<td>Presidential</td>
<td>.126</td>
<td>.001</td>
<td>.124, .128</td>
<td>.124</td>
<td>.128</td>
</tr>
<tr>
<td>Constant</td>
<td>-.045</td>
<td>.006</td>
<td>-.056, -.035</td>
<td>-.056</td>
<td>-.035</td>
</tr>
</tbody>
</table>

N= 21071
Table 4-17: Balance of Matched Sample

| Variable      | Unmatched data | | | Matched data | | |
|---------|----------------|-----------------|-----------------|-----------------|-----------------|
|         | Treated | Control | Difference | Treated | Control | Difference |
| Distance | .368 | .341 | .027 | .368 | .367 | .002 |
| Age     | 45.334 | 46.215 | -.881 | 45.334 | 45.503 | -.169 |
| Female  | .532 | .534 | -.002 | .532 | .540 | -.008 |
| Income  | 2.882 | 2.922 | -.040 | 2.882 | 2.891 | -.009 |
| Education | 4.397 | 4.260 | .136 | 4.397 | 4.399 | -.002 |
| Partisan | 2.896 | 2.819 | .077 | 2.896 | 2.916 | -.020 |
| Interest | 2.108 | 2.097 | .012 | 2.108 | 2.130 | -.022 |
| Discuss | .822 | .801 | .021 | .822 | .828 | -.006 |
| Newspaper | 3.404 | 3.614 | -.210 | 3.404 | 3.432 | -.028 |
| TV News  | 3.819 | 4.149 | -.329 | 3.819 | 3.865 | -.045 |
| Efficacy | 2.880 | 2.863 | .016 | 2.880 | 2.877 | .002 |
| Married | .551 | .565 | -.014 | .551 | .552 | -.001 |
| Competition | .446 | .454 | -.009 | .446 | .451 | -.005 |
| Presidential | .635 | .635 | .000 | .635 | .643 | -.009 |
| Minority | .317 | .215 | .102 | .317 | .320 | -.003 |
| Informed | 2.722 | 2.770 | -.048 | 2.722 | 2.700 | .022 |

Table 4-18: Table for Matched Sample

<table>
<thead>
<tr>
<th></th>
<th>Non-Voters</th>
<th>Voters</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>No Early Voting</td>
<td>711</td>
<td>1774</td>
<td>2485</td>
</tr>
<tr>
<td></td>
<td>28.6</td>
<td>71.4</td>
<td>100</td>
</tr>
<tr>
<td>Early Voting</td>
<td>801</td>
<td>1684</td>
<td>2485</td>
</tr>
<tr>
<td></td>
<td>32.2</td>
<td>67.8</td>
<td>100</td>
</tr>
<tr>
<td>Total</td>
<td>1512</td>
<td>3458</td>
<td>4970</td>
</tr>
<tr>
<td></td>
<td>30.4</td>
<td>69.6</td>
<td>100</td>
</tr>
</tbody>
</table>

Prob($\chi^2$) = .006
<table>
<thead>
<tr>
<th></th>
<th>Coef. (S.E.)</th>
<th>Coef. (S.E.)</th>
<th>Coef. (S.E.)</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Voting</td>
<td>-.171 (.062)</td>
<td>-.171 (.062)</td>
<td>-.137 (.076)</td>
<td>-.263</td>
<td>-.012</td>
</tr>
<tr>
<td>Age</td>
<td>.026 (.003)</td>
<td>.021 (.030)</td>
<td>.224 (.486)</td>
<td>.127</td>
<td>.261</td>
</tr>
<tr>
<td>Female</td>
<td>.355 (.079)</td>
<td>.224 (.486)</td>
<td>-.012</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>.194 (.041)</td>
<td>.127 (.261)</td>
<td>.418</td>
<td>.350</td>
<td>(.041)</td>
</tr>
<tr>
<td>Education</td>
<td>.218 (.029)</td>
<td>.171 (.266)</td>
<td>.697 (.063)</td>
<td>.594</td>
<td>.801</td>
</tr>
<tr>
<td>Partisanship</td>
<td>.350 (.041)</td>
<td>.281 (.418)</td>
<td>.359 (.101)</td>
<td>.193</td>
<td>.525</td>
</tr>
<tr>
<td>Interest</td>
<td>.697 (.063)</td>
<td>.594 (.801)</td>
<td>.359 (.101)</td>
<td>.193</td>
<td>.525</td>
</tr>
<tr>
<td>Discussion</td>
<td>.051 (.015)</td>
<td>.027 (.075)</td>
<td>.011 (.015)</td>
<td>.036</td>
<td>.015</td>
</tr>
<tr>
<td>Newspaper</td>
<td>.011 (.015)</td>
<td>-.036 (.015)</td>
<td>.011 (.015)</td>
<td>.036</td>
<td>.015</td>
</tr>
<tr>
<td>National TV</td>
<td>-.010 (.034)</td>
<td>-.065 (.045)</td>
<td>-.010 (.034)</td>
<td>-.065</td>
<td>.045</td>
</tr>
<tr>
<td>News</td>
<td>.116 (.053)</td>
<td>.029 (.203)</td>
<td>.116 (.053)</td>
<td>.029</td>
<td>.203</td>
</tr>
<tr>
<td>Gov’t Trust</td>
<td>.260 (.084)</td>
<td>.123 (.398)</td>
<td>.260 (.084)</td>
<td>.123</td>
<td>.398</td>
</tr>
<tr>
<td>Married</td>
<td>1.70 (.438)</td>
<td>.978 (2.42)</td>
<td>.70 (.438)</td>
<td>.978</td>
<td>2.42</td>
</tr>
<tr>
<td>Competitiveness</td>
<td>1.00 (.089)</td>
<td>.855 (1.14)</td>
<td>.100 (.089)</td>
<td>.855</td>
<td>1.14</td>
</tr>
<tr>
<td>Presidential</td>
<td>-.052 (.083)</td>
<td>-.188 (.085)</td>
<td>-.052 (.083)</td>
<td>-.188</td>
<td>.085</td>
</tr>
<tr>
<td>Minority</td>
<td>-.563 (.047)</td>
<td>-.640 (.486)</td>
<td>-.563 (.047)</td>
<td>-.640</td>
<td>.486</td>
</tr>
<tr>
<td>Informed</td>
<td>.914 (.044)</td>
<td>.973 (.145)</td>
<td>.914 (.044)</td>
<td>.973</td>
<td>.145</td>
</tr>
<tr>
<td>Distance</td>
<td>.914 (.044)</td>
<td>.973 (.145)</td>
<td>.914 (.044)</td>
<td>.973</td>
<td>.145</td>
</tr>
<tr>
<td>Constant</td>
<td>-.159 (.376)</td>
<td>-.640 (.486)</td>
<td>-.159 (.376)</td>
<td>-.640</td>
<td>.486</td>
</tr>
<tr>
<td>N=</td>
<td>4970</td>
<td>4970</td>
<td>4970</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4-20: Substantive Effects

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>S.E.</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Change in Prob</td>
<td>-.016</td>
<td>.009</td>
<td>-.031</td>
<td>-.002</td>
</tr>
</tbody>
</table>
Table 4-21: Model Estimates for Partisanship

<table>
<thead>
<tr>
<th></th>
<th>Coef. (S.E.)</th>
<th>Coef. (S.E.)</th>
<th>Coef. (S.E.)</th>
<th>90% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Voting</td>
<td>-.539 (.191)</td>
<td>-.530 (.191)</td>
<td>-.537 (.238)</td>
<td>Lower: -.100 Upper: -.071</td>
</tr>
<tr>
<td>Partisanship</td>
<td>.439 (.046)</td>
<td>.466 (.050)</td>
<td>.279 (.057)</td>
<td>.185 .373</td>
</tr>
<tr>
<td>Partisan x Early Voting</td>
<td>.134 (.066)</td>
<td>.132 (.066)</td>
<td>.144 (.081)</td>
<td>.010 .277</td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>.026 (.003)</td>
<td></td>
<td>.021 .030</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td>.356 (.079)</td>
<td></td>
<td>.226 .487</td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td>.195 (.041)</td>
<td></td>
<td>.128 .262</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td>.218 (.029)</td>
<td></td>
<td>.170 .266</td>
</tr>
<tr>
<td>Interest</td>
<td></td>
<td>.697 (.063)</td>
<td></td>
<td>.594 .801</td>
</tr>
<tr>
<td>Discussion</td>
<td></td>
<td>.362 (.101)</td>
<td></td>
<td>.195 .529</td>
</tr>
<tr>
<td>Newspaper</td>
<td></td>
<td>.051 (.015)</td>
<td></td>
<td>.027 .075</td>
</tr>
<tr>
<td>National TV News</td>
<td></td>
<td>-.011 (.015)</td>
<td></td>
<td>-.036 .015</td>
</tr>
<tr>
<td>Gov’t Trust</td>
<td></td>
<td>-.008 (.034)</td>
<td></td>
<td>-.064 .047</td>
</tr>
<tr>
<td>Efficacy</td>
<td></td>
<td>.116 (.053)</td>
<td></td>
<td>.029 .203</td>
</tr>
<tr>
<td>Married</td>
<td></td>
<td>.260 (.084)</td>
<td></td>
<td>.122 .397</td>
</tr>
<tr>
<td>Competitiveness</td>
<td></td>
<td>1.71 (.439)</td>
<td></td>
<td>.990 2.43</td>
</tr>
<tr>
<td>Presidential</td>
<td></td>
<td>.998 (.089)</td>
<td></td>
<td>.854 1.14</td>
</tr>
<tr>
<td>Minority</td>
<td></td>
<td>-.053 (.083)</td>
<td></td>
<td>-.189 .084</td>
</tr>
<tr>
<td>Informed</td>
<td></td>
<td>-.564 (.047)</td>
<td></td>
<td>-.641 -.487</td>
</tr>
<tr>
<td>Distance</td>
<td></td>
<td>-1.44 (.392)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-.330 (.136)</td>
<td>.124 (.183)</td>
<td>-4.60 (.432)</td>
<td>-5.32 -3.90</td>
</tr>
</tbody>
</table>

N= 4970 4970 4970
Table 4-22: Substantive Effects by Partisanship

<table>
<thead>
<tr>
<th>Partisanship</th>
<th>Mean</th>
<th>S.E.</th>
<th>Lower</th>
<th>Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Independent</td>
<td>-.073</td>
<td>.029</td>
<td>-.132</td>
<td>-.019</td>
</tr>
<tr>
<td>Leaning</td>
<td>-.038</td>
<td>.014</td>
<td>-.067</td>
<td>-.010</td>
</tr>
<tr>
<td>Moderate</td>
<td>-.012</td>
<td>.009</td>
<td>-.029</td>
<td>.005</td>
</tr>
<tr>
<td>Strong</td>
<td>.004</td>
<td>.011</td>
<td>-.017</td>
<td>.026</td>
</tr>
</tbody>
</table>
Table 4-23: Model Estimates for Interest

<table>
<thead>
<tr>
<th></th>
<th>Coef. (S.E.)</th>
<th>Coef. (S.E.)</th>
<th>Coef. (S.E.)</th>
<th>90% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower</td>
</tr>
<tr>
<td>Early Voting</td>
<td>-.486 (.205)</td>
<td>-.479 (.205)</td>
<td>-.321 (.235)</td>
<td>-.708</td>
</tr>
<tr>
<td>Interest</td>
<td>1.19 (.071)</td>
<td>1.19 (.071)</td>
<td>.651 (.084)</td>
<td>.513</td>
</tr>
<tr>
<td>Interest x Early</td>
<td>.164 (.102)</td>
<td>.161 (.102)</td>
<td>.095 (.115)</td>
<td>-.094</td>
</tr>
<tr>
<td>Voting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td>.026 (.003)</td>
<td></td>
<td>.021</td>
</tr>
<tr>
<td>Female</td>
<td></td>
<td>.354 (.079)</td>
<td></td>
<td>.224</td>
</tr>
<tr>
<td>Income</td>
<td></td>
<td>.194 (.041)</td>
<td></td>
<td>.127</td>
</tr>
<tr>
<td>Education</td>
<td>.219 (.029)</td>
<td></td>
<td></td>
<td>.172</td>
</tr>
<tr>
<td>Partisanship</td>
<td></td>
<td>.349 (.041)</td>
<td></td>
<td>.281</td>
</tr>
<tr>
<td>Discussion</td>
<td>.360 (.101)</td>
<td></td>
<td></td>
<td>.194</td>
</tr>
<tr>
<td>Newspaper</td>
<td>.051 (.015)</td>
<td></td>
<td></td>
<td>.027</td>
</tr>
<tr>
<td>National TV News</td>
<td>-.010 (.015)</td>
<td></td>
<td>-.036</td>
<td></td>
</tr>
<tr>
<td>Gov’t Trust</td>
<td>-.001 (.034)</td>
<td></td>
<td>-.065</td>
<td></td>
</tr>
<tr>
<td>Efficacy</td>
<td>.116 (.053)</td>
<td></td>
<td>.029</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>.261 (.084)</td>
<td></td>
<td>.124</td>
<td></td>
</tr>
<tr>
<td>Competitiveness</td>
<td>1.69 (.439)</td>
<td></td>
<td></td>
<td>.966</td>
</tr>
<tr>
<td>Presidential</td>
<td>1.00 (.089)</td>
<td></td>
<td>.855</td>
<td></td>
</tr>
<tr>
<td>Minority</td>
<td>-.053 (.083)</td>
<td></td>
<td>-.189</td>
<td></td>
</tr>
<tr>
<td>Informed</td>
<td>-.562 (.047)</td>
<td></td>
<td>-.638</td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td>.391 (.403)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-1.48 (.144)</td>
<td>-1.34 (.202)</td>
<td>-4.71 (.432)</td>
<td>-5.42</td>
</tr>
</tbody>
</table>

N= 4970 4970 4970
Table 4-24: Substantive Effects by Interest

<table>
<thead>
<tr>
<th>Interest</th>
<th>Mean</th>
<th>S.E.</th>
<th>90% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>-.041</td>
<td>.024</td>
<td>-.082</td>
</tr>
<tr>
<td>Moderate</td>
<td>-.015</td>
<td>.009</td>
<td>-.030</td>
</tr>
<tr>
<td>High</td>
<td>-.002</td>
<td>.010</td>
<td>-.020</td>
</tr>
</tbody>
</table>
Figure 4-1: Priors

Effect of Informative Priors

Figure 4-2: Difference between candidates

Differences Between the Candidates
Figure 4-3: Variation in candidates

![Variability in Candidate Qualities](image)

Figure 4-4: Campaign information

![Voters' Information about the Candidates](image)
Figure 4-5: Q-Q Plot

Propensity Scores for Treatment and Control Samples

- Full Data
- Matched Sample
Chapter 5:
Election Day Vote Centers

5.1 Introduction

This chapter considers whether Election Day vote centers (EDVCs) increase turnout. Vote centers are the most recent election reform investigated in this study. Vote centers are designed to improve the process of election administration for local officials and the balloting experience for voters. The reform does away with the diffuse network of small precinct polling sites in favor of larger, more modern, and better equipped facilities. Vote centers also provide individuals with a choice of where to vote, rather than having the choice made for them when the boundaries of voting districts are drawn, which can be incongruent with neighborhood boundaries.

The next section provides a review of the relevant literature, followed by a definition of vote centers and a brief discussion of their history. The third section discusses the theoretical rationale that suggests a link between vote centers and turnout followed by an empirical investigation of the effects. After presenting the main analyses, I draw from queuing theory to offer an explanation for the long lines that formed at some vote centers (in two counties in particular) in 2006. The final section of this chapter summarizes the results and discusses substantive conclusions.

5.2 Literature Review

Since vote centers are a very recent reform there are relatively few studies that directly examine their effects. Stein and Vonnahme (2008) examine data from the first three
elections in which vote centers were used by comparing voter file data from Larimer and Weld counties. The study uses data from voter registration files to match individual voters across the counties to assess the effects of the reform. Stein and Vonnahme find that the vote centers increase turnout by around 2.5% overall, and as high as 5% for the least engaged voters.

A subsequent study by Juenke and Shephard (2007) finds that vote centers have an ambiguous effect on turnout, and that the effect may be conditioned by party mobilization. Juenke and Shephard analyze county-level data from 2006 to assess the effects of vote centers, although the empirical results on which their conclusions are based may be subject to a number of criticisms. Their list of counties with vote centers was obtained from Fair Vote Colorado which provides a different list than those provided by the Colorado Secretary of State and a survey of county clerks. Furthermore, there are concerns with the analysis of the data.

The authors’ conclusions depend on OLS regression estimates of a linear model of turnout with 15 parameters and 64 data points. Given the large number of parameters to be estimated relative to the number of data points, the large-sample properties of the OLS estimates are less likely to apply, particularly that the sampling distributions of the estimates are normally distributed. The estimates are also normally distributed with a finite sample, but this depends on independent, normally distributed errors which might be problematic for cross-sectional data (Greene 2002). This assumption underlies their hypothesis tests and the results on which their qualitative conclusions are based. This is especially problematic given the lack of clear justification for the included control variables.
Additionally the authors also conduct lagged regressions, including turnout in 1998 and 2002 as control variables in the models. Lagged regressions to account for different baseline rates of turnout may be particularly problematic when studying socially defined units such as counties (Allison 1991). The problem is that lagged regressions under the null hypothesis of no treatment effect implicitly assume between-unit regression towards the mean (Allison 1991). To see this, consider a simple example with only two counties, one receives the treatment the other does not. The observed difference between the counties in the first time period is simply:

$$\Delta_{t-1} = y_{T,t-1} - y_{C,t-1}$$

Applying a lagged regression model, the expected difference in the next time period is:

$$E[\Delta_t] = E[y_T] - E[y_C]$$
$$E[y_T] = \alpha + \beta T + \phi'_{T,t-1}$$
$$E[y_C] = \alpha + \phi'_{C,t-1}$$
$$E[\Delta_t] = \beta T + \phi'_{T,t-1} - \phi'_{C,t-1}$$

Under the null hypothesis of no treatment effect this difference can be expressed as

$$E[\Delta_t] = \phi'_{T,t-1} - \phi'_{C,t-1}$$
$$E[\Delta_t] = \phi(y_{T,t-1} - y_{C,t-1})$$

This implies that the expected difference between the treated and control units will be $\phi$ of the lagged difference (Allison 1991). For our application, if the estimate of $\phi$ was .5 then the difference between the treated and control units under the null hypothesis would be half that of the previous period. In their paper, Juenke and Shephard estimated the coefficient on lagged turnout from 2002 to be .288, suggesting a substantial regression to the mean from the previous midterm election.
There is little apparent justification for this assumption in these models of county turnout. While there is variation in levels of turnout within counties across elections the populations are fairly stable, particularly over short analysis times. It would be unusual for counties to exhibit the type of regression process implied by the lagged regression. Counties with relatively low levels of turnout are more likely to have low levels of turnout in future elections just as counties with high levels of turnout are likely to have high levels of turnout in future elections. It is unlikely that the counties display the volatility in levels of turnout that is implied by the lagged regression results. It should also be mentioned that other models were also considered that exclude the lagged variable. These models are even more problematic as they do not account for previous levels of turnout which show substantial variation across the counties and cannot (assuming a conventional temporal ordering of cause and effect) be attributed to vote centers. An alternative difference-in-difference estimator is considered in the research design which addresses these problems and suggests a different conclusion.

Regarding Juenke and Shephard's conclusion that the effect of vote centers on turnout are conditioned by the efforts of party mobilization, the evidence is derived from an indirect measure of mobilization. To measure mobilization Juenke and Shephard create a measure of the closeness of the county-wide vote in the 2004 presidential election between Bush and Kerry. The use of this variable from the 2004 presidential election to measure party mobilization in 2006 is not well-defended and represents more of a hypothesis to be tested than a strong basis for constructing a measure as the authors argue that mobilization will be higher for closer counties. Yet the statewide offices on the ballot in 2006 aggregate votes at the state rather than county level, suggesting that candidates might focus on areas with large
populations rather than counties with a proportionately close vote between Bush and Kerry in 2004.

Juenke and Shephard then interact the closeness of the county presidential vote in 2004 with the treatment variable to test the hypothesis that the effects of vote centers depend on mobilization. As discussed below, this result might be more dependent on the particular configuration of the data and the implementation of vote centers than a conditional effect of mobilization, but this is difficult to assess given the choice of model specification. That the effects of vote centers might be conditioned by party mobilization efforts is an interesting hypothesis but at this point it would be premature to draw strong conclusions from the available evidence. While there is not a substantial scholarly literature on the effects of vote centers on turnout, there are other related studies of the effects of polling locations more generally.

5.2.1 Effects of Polling Locations

Scholars have also examined the effects of polling locations more generally on electoral participation. One recent area of research is the effects of distance from one's polling location on voter turnout. The general argument from existing research is that individuals' probability of turning out to vote should decrease as the polling location is located farther away, but the empirical findings have varied across studies. One study (Haspel and Knotts 2005) shows the voter turnout tends to decrease as individuals are farther from their polling location, but that this effect is mediated by whether the individual has access to a car. This study uses data from Atlanta, however, and it is not clear how widespread individuals' access to a car is, particularly since the 2005 American Community Survey shows that over 91% of
respondents have one or more cars in their household, and the percentage might be higher for registered voters.

Another study (Gimpel and Schuknecht 2003) suggests a non-linear relationship between distance and turnout, so that individuals that are close, and much farther away are not affected, while individuals that are moderately far from their polling location are less likely to vote. They conjecture that the non-linear relationship might be related to travel times, so that the actual effect is the time spent traveling rather than the distance, and a moderate commute from a suburb might take longer due to traffic congestion than a longer (in terms of distance) rural commute with open roads.

A third (Brady and McNulty 2002) study of precinct consolidations in Los Angeles finds that individuals' likelihood of turning out declines as distance from the polling location increases. Brady and McNulty examine polling locations in the 2002 California recall election in Los Angeles, and find that voter turnout declines by 0.5% with every mile from their polling location. The results found by Brady and McNulty appears to be a general effect, though the context might hold several possible mediating factors constant. Access to a car might be more pervasive in Los Angeles County (according to the 2006 American Community Survey over 90% of households had access to one or more vehicles), which is a more commuter-centric community and is also more geographically homogenous than counties with a mixture of urban and rural populations. The next section provides a brief definition and history of vote centers, followed by the theoretical argument and research design.

5.3 Definition and History
Election Day vote centers provide an alternative means of administering Election Day voting. Two characteristics of vote centers that distinguish them from precinct voting are openness and centralization. Openness refers to polling locations that are open to all voters in a county rather than to a specific voting district. Centralization refers to polling sites that are larger, have more poll workers, and more voting stations. Traditional elections are conducted using a decentralized network of precinct-specific polling locations, with each individual assigned to a specific precinct site based on their residential address in the registration book. Election Day vote centers provide an alternative method of voting by establishing a smaller number of open polling sites throughout the county, so that individuals can vote at any site regardless of their precinct. Election Day vote centers also tend to be larger than precinct sites to accommodate a higher volume of voters, are typically located in larger buildings such as schools, churches, and businesses, and are positioned along major transportation routes. The polling sites also provide more available parking, more voting machines, and more poll workers that are also better trained to assist voters with the electronic voting machines and to address any issues that might arise.

5.3.1 History of Vote Centers

Election Day vote centers (EDVC) were first developed by Scott Doyle, the County Clerk and Recorder in Larimer County, Colorado. The vote center model was developed following the Help America Vote Act which required counties to update their voting technologies and make polling places more accessible for disabled voters. These reforms prompted the development of the vote center model as new voting technologies could accommodate different ballots within counties. In addition, it became more costly to find existing facilities
that would host a polling site on Election Day and meet accessibility requirements, or to modify existing polling place facilities to meet the requirements. The vote center model was adopted in part as a response to reforms from the Help America Vote Act.

The vote center model was first implemented in Larimer County, Colorado in a local election in 2003. County officials have the authority to determine the mode of voting in local elections, and 2003 was the first year in which precinct-specific locations in Larimer were replaced by county-wide vote centers. The first election using vote centers was successful as reported by voters, media accounts, and election officials. Following the first successful implementation of vote centers in 2003, Scott Doyle requested that the Colorado legislature authorize Larimer to continue with the vote center model. Legislation that enabled Larimer to continue using vote centers was passed in May 2004.

Larimer again used vote centers in the 2004 presidential election. Additional sites were added for the general election in anticipation of high turnout. The vote centers again performed well as reported by voters, local media, and election officials. Despite having fewer polling locations and a record high turnout of over 90% of the registered voters, Larimer county election officials were able to efficiently conduct the election. Following the successful implementation of vote centers in the 2004 presidential election, three other counties in Colorado adopted the vote center model for the subsequent election.

By 2005 a total of four counties had adopted vote centers, including Larimer, Weld, Adams, and Otero counties. Also in 2005, the Larimer County elections division won a national award for election administration and Scott Doyle was recognized as Clerk of the Year by the National Association of County Clerks and Recorders for the vote center model. In the 2005 local elections vote centers were again well-received by voters in Weld, Adams,
and Otero counties. By 2006 a total of 19 counties in Colorado and one in Texas had adopted vote centers.

The 2006 midterm elections produced mixed results for the vote center model. Most counties, including previous and new adopters of the vote center model, had positive experiences with vote centers. Denver and Douglas counties, however, experienced long lines at polling locations and there were reports of voters leaving polling sites without voting due to the long wait times. A commission was established by officials in Denver County to review the election process and attempt to diagnose the problems with the 2006 election. The commission reported that problems with the electronic pollbook software slowed the processing of voters at polling locations and led to long lines. In part as a result of problems with the 2006 elections (though also related to previous problems with election administration including losing several file cabinets of voter registration files) Denver's appointed election commission was replaced by an elected county clerk in 2007.

5.4 Theory

From the existing theoretical models of turnout, there are several possible factors that might directly affect whether or not an individual votes. Factors that are directly related to voter turnout are generally thought to be the costs of voting, the benefits of having one's preferred candidate in office, and the probability of casting a decisive ballot (one could also include a voters' marginal contribution, positive or negative, to the mandate of the winning candidate in the benefit term). In this chapter, I focus on how Election Day vote centers might be related to the costs of voting and thereby affect levels of turnout. We will not consider in much detail whether or not vote centers would also be related to the benefits of
voting. While it does not seem obvious how vote centers could be related to the benefits of voting, one might potentially draw from recent research on the effects of polling locations on voter behavior (Berger, Meredith, and Wheeler 2006), or research on voting as a form of social participation to consider whether vote centers might also be related to turnout through the benefits of voting (Edlin, Gelman, and Kaplan 2007). Regarding the costs of voting, we consider how each of the two distinguishing feature of vote centers, centralization and openness, might be related to the costs of voting and thereby indirectly affect voter turnout.

5.4.1 Effects of Centralization

The first characteristic of Election Day vote centers is centralization. There are several possible ways that centralization might affect the costs of voting, and thereby influence the likelihood that an individual votes. One possible effect of centralization is that it might be easier to disseminate information about where an individual can vote. Reducing the number of sites would seem to simplify the information that needs to be conveyed to voters, and might also allow for facilities that are easier for individuals to find. By lowering the total number of sites and making them all available to eligible voters, Election Day vote centers might lower the informational costs of voting.

Centralization might also reduce the number of poll workers that are required to administer an election. This might have several effects. By centralizing polling places, there will be a larger number of workers at the site, allowing workers to focus on more specialized tasks. This might allow for better trained workers for any particular task, could reduce the costs that an individual voter faces at the polling site, and might also lead to a faster processing of voters overall, reducing the time spent in lines. Centralization might also lead
to greater consistency across the polling sites and allow for more consistent expectations of the voters as to how the polling site will operate.

While centralization might tend to increase turnout for the reasons discussed above there might also be concerns that these effects would be offset by having fewer polling locations and perhaps increasing the distance that voters have to travel to polling locations. In this way the positive effects of centralization might be offset by a negative effect on turnout of having fewer polling locations and increasing distances to polling locations. Previous empirical research suggests a link between distance and turnout, although it is not clear if the effects on conditioned by the availability of modes of transportation or if the effects are monotonic (Haspel and Knotts 2005, Gimpel and Schucknect 2003, Brady and McNulty 2005). The possibility that centralization might also have a negative impact suggests that net effect of centralization could be positive, negative, or could neutralize one another. Yet vote centers combine centralization with openness. This might mitigate, to some extent, the potential increase in voters’ distances to polling locations as openness allows them to vote at any site.

While centralization will increase voters’ distances to polling sites it is not clear how centralization combined with openness would affect distance to polling locations. Furthermore the traditional measures of distance from one’s polling location have been calculated using residential addresses. While this is sensible with precinct-based polling locations it might matter less with vote centers as individuals might be voting closer to workplaces, schools, commuting routes, etc. The research design section of the chapter further discusses the issue of centralization and distances to polling locations. Section 5.1 provides evidence that the relationship between vote centers and distances to polling
locations is not strong and that there seems to be little basis for concerns about centralization and vote centers. That is, even if distance affects turnout, vote centers seem to be only weakly related to distance.

5.4.2 Effects of Open Polling Locations

The second characteristic of Election Day vote centers is openness. While both openness and centralization are thought to affect voter turnout by reducing the costs of voting, the specific mechanisms linking the two characteristics of Election Day vote centers to the costs of voting are not identical. One of the main advantages to open polling locations is that it gives voters more choices on Election Day. With precinct locations, an Election Day voter can only go to their precinct location. With Election Day vote centers, individuals can choose to vote at any of the locations throughout the county. While precinct-based locations are assigned on the basis of residential location, individuals that commute to work, school, or shopping might find it more convenient to ballot near one of their destinations. Open polling locations permit individuals to choose the polling location that is the most convenient for them.

Open polling locations not only provide individuals with a choice of where to vote, but might also reduce the informational costs of finding out where to vote. Precinct-based elections require individuals to know the designated polling place for their voting district so as to not go to the wrong place. With open polling locations there is no wrong place to vote, and individuals no longer need to have particular information about the site established for their particular voting district.

More accessible polling locations might also better distribute voter arrival times throughout the day. For individuals that commute to work or school, a residentially-based
polling location might only be accessible in the morning or evenings, before or after their work or school. This could lead to higher loads at the polling places in the morning and evening peak times. Open polling locations allow individuals to vote at places that are more accessible throughout the day, which might help to better distribute voter arrivals throughout the day, reducing the load on the polling places at peak times, and resulting in shorter lines for voters. Not only might open polling locations better distribute voters throughout the day, but they might also be better distributed across polling locations, so that individuals would no longer be bound to a particular site if there was a localized problem at a particular voting site.

By combining centralization and open polling locations, Election Day vote centers might lower the costs of voting and thereby increase the likelihood that an individual turns out to vote. There are several mechanisms by which centralization and openness are thought to affect the costs of voting, and the result of lower costs of voting suggests that Election Day vote centers should be associated with higher voter turnout.

5.4.3 Hypotheses

Hypothesis 1: Vote centers increase the likelihood that an individual turns out to vote.

Hypothesis 2: Vote centers are more effective at increasing turnout among less engaged voters.

5.5 Research Design

5.5.1 Distances from Polling Locations
Before discussing the effects of Election Day vote centers it we should consider how vote centers are implemented. The two shared characteristics of vote centers are openness and centralization. Openness is a binary characteristic not subject to variation in implementation.\(^\text{14}\) Centralization on the other hand can occur to a greater or less extent as counties reduce the number of polling locations to varying degrees. State laws might regulate the extent to which centralization can occur by mandating a minimum number of polling sites, such as one per 10,000 registered voters as has been the minimum statutory requirement in Colorado. While this establishes a legal minimum, counties might establish additional polling sites that could alter the apparent effects of vote centers when comparing counties.

To assess the effects of vote centers on the distance between voter's residential locations and polling places we geocoded the voter registration file from Larimer County, Colorado. This process assigns a decimal latitude and longitude to each record in the registration files. Of the 178,748 records in the registration file we obtained geocodes for 150,079 cases (84\%). We then geocoded the 143 precinct polling locations from the 2002 election, the last year in which Larimer County used precinct-specific sites. We then matched each voter to their designated precinct site. This allows us to calculate the distance between each voter and their precinct-specific polling location.

To obtain the distance measure, I calculated the geographic distance between each voter and their polling location using the haversine formula. Using the latitudes and longitudes (in radians) of each voter and each polling location in the county I calculated the great circle distances using the haversine formula as follows (Sinnott 1984):

\(^{14}\) That is, sites are either open or not. At least conceptually there may be variations in openness if voters in a county were blocked according to zones or had multiple sites available to voters in a particular area without county-wide voting. While conceptually possible no county has implemented such a hybrid vote center model.
\[ dlong = long_2 - long_1 \]
\[ dlat = lat_2 - lat_1 \]
\[ a = (\sin \frac{dlat}{2})^2 + \cos(lat_1) \cdot \cos(lat_2) \cdot (\sin \frac{dlong}{2})^2 \]
\[ c = 2 \cdot \arctan 2(\sqrt{a}, \sqrt{1-a}) \]
\[ d = R \cdot c \]

where \( R \) is the radius of the Earth in miles to convert the distance to miles. This formula was used to account for the curvature of the Earth's surface. Euclidean distances could also be used, but the flat Earth assumption would introduce additional measurement error. The error in Euclidean distances over the area of a county is unlikely to be substantial and would not seem to bias the comparison, but the haversine formula provides a relatively straightforward and more accurate estimate of the geographic distances.

In contrast to precinct locations, voters are not assigned to any one vote center location. Thus to obtain a measure of distance from vote center locations, I calculated the same great sphere distance between voters' residential addresses and vote center locations. The nearest vote center location was used as it seems to be the most relevant comparison to address the concerns that vote centers are geographically remote.

The 22 vote center locations were also geocoded and great sphere distances were calculated between each voter and each vote center location and the shortest distances were recorded for comparison. The results for the full samples are shown in Table 5-1. From these data we can see that vote centers actually decreased the mean distance between voters' residential locations and the nearest available polling location (from 2.46 miles to 1.54). However the results are particularly right-skewed as a few voters living in rural areas are much farther from their polling locations than most. Median distances are less sensitive to
the tails of the distributions, and we can see that vote centers are associated with an increase in median distances of less than .5 miles (from .61 miles to 1.1).

We can also truncate the distributions at five miles which captures most voters and eliminates the long right tails of the distributions. Limiting the distributions to voters within 5 miles of their polling location there is again a less than .5 mile increase in distance (from a mean of .87 miles to 1.33). Figure 5-1 shows a smoothed density estimate of voters' distances from precinct and vote center locations. From these results we can see that despite a dramatic decrease in the number of polling locations (from 143 in 2002 to 22 in 2003) vote centers were associated with an average increase in distance of less than .5 miles from voters' residences to their nearest polling location. This calculation also does not account for the openness of vote centers wherein voters might ballot closer to workplaces, schools, commuting routes, etc. which would further limit the changes in distance.

Maps were also created using the geocoded files to show the location of voters, precinct sites, and vote centers as displayed in Figure 5-2 to visually depict the geographic distributions of voters and polling sites across Larimer County. Even if we ignore every other aspect of vote centers such as better equipped facilities, greater administrative control, and open polling locations and only focus on this distance comparison there is not a substantial difference between precinct sites and vote centers. Applying the Brady and McNulty result, this increase in distance would affect turnout by less than 1.25 percent on average, not including the upper tail of the distribution or accounting for openness.

That the average distance from a voter's residence to the nearest vote center is comparable to the average distance to a voter's precinct location despite the smaller number of vote centers suggest that the concern about distance and vote centers is unlikely to be a
significant factor. Even assuming that distance from one's residence is a significant barrier to turnout, which the existing research suggests might be a conditional effect, the distance is not substantially greater with Election Day vote centers. While distance may be an important consideration for the implementation of Election Day vote centers there is no necessary connection between a smaller number of open polling sites and a substantial increase in residential distances.

5.5.2 Effect on Turnout

With the adoption of vote centers by 19 counties in Colorado in 2006 we can evaluate aggregate turnout data from the counties to assess the potential effects of vote centers. Previous research has considered evidence regarding the effects of vote centers from county-level data (Juenke and Shephard 2007) and found mixed results and a possible conditional effect on turnout.

To further consider the effects of vote centers using the available aggregate data, a difference-in-difference estimate is used. The dependent variable in these analyses is the ratio of the number of ballots cast to the number of registered voters. While previous studies have typically used voting age population or voting eligible population to measure turnout, the proportion of registered voter turnout might be more appropriate in this case for three reasons. The first is that vote centers should affect turnout of the eligible voters but it is not clear how vote centers would affect the number of registered voters. In this way vote centers should operate to increase turnout among those eligible to vote rather than by expanding the pool of eligible voters. The second is that all of the counties are located in Colorado and thus will share statewide regulations as to the maintenance of registration lists making the data
more comparable across counties. The third is that comparisons will also be made within counties so that the analysis also accounts for variation in the implementation of regulations regarding registration lists.

The main independent variable of interest is whether the county used Election Day vote centers in 2006. There is some confusion as to which counties used vote centers with public interest groups publishing different lists from the Colorado Secretary of State's office, which is different from the counties.\textsuperscript{15} Our measure of EDVC is constructed using information from a survey of county clerks conducted in Scott Doyle and the Larimer County Clerk and Recorder’s Office which includes 50 of the 64 total counties. Information provided by the Colorado Secretary of State’s Office and official county websites was used for the other 14 counties. In total there are 19 counties that are identified as using vote centers with the remainder using precinct polling locations. The list of counties is included in the appendix C.

\textbf{Descriptive Summaries} The data span midterm elections from 1994 to 2006 and include all counties in Colorado. In elections prior to 2002 there were 63 counties in Colorado and from 2002 forward there are 64 after the addition of Broomfield county. Turnout rates were calculated for each of the elections and we can see from the summary statistics in Table 5-2 that vote center counties had historically lower turnout in years before 2006. It is important to note that these differences cannot be attributed to vote centers as no county used vote centers in any midterm election prior to 2006. While turnout was lower in vote center

\textsuperscript{15} The main source of confusion seems to be whether precinct consolidation counts as EDVC, which does not fit our definition of vote centers as it lacks the openness characteristic.
counties prior to 2006 we can see that turnout increased in those counties in 2006 relative to any of the four previous midterm elections.

While these summaries can help to account for enduring county-level traits there might be differences in the specific elections. In particular, the campaigns in 2006 might have been more competitive, had higher profiles, greater media attention, more get-out-the-vote efforts, or any number of other factors that are specific to certain elections that might also affect turnout. To attempt to account for this possibility we can compare the increase in turnout in vote centers counties to changes in turnout in other counties. If turnout increases in vote center counties in 2006 but decreases in non-vote center counties we can be more confident that the effect is not attributable to more competitive races (or other factors), which would affect turnout in counties across the state (both treatment and control counties). The data in Table 5-2 shows that while turnout also increased in non-vote center counties it did so at a lower rate. We can see from these data that the gap between turnout rates in vote center and non-vote center counties always favored non-vote center counties (from 1994 to 2006) but the gap was most narrow in 2006.

We can also compare the change in turnout from 2002 to 2006 in vote center counties to the statewide average. The average turnout for all counties in Colorado was 56.04% in 2002 and 67.8% in 2006 for an overall increase of 11.76. Looking only at vote center counties the average turnout was 50.36% in 2002 and 65.11 in 2006, and 58.1% in 2002 and 68.77 in 2006 for non-vote center counties. The ratio of the average change across vote center counties (14.75) over of the average statewide change in turnout (11.76) shows that the average change in vote center counties is 125% of the average statewide change. Looking at non-vote center counties, on the other hand, shows that their average change in
turnout was only 90.8% of the statewide average. By comparing turnout across time and between vote center and non-vote center counties we can see that there appears to be evidence that vote centers led to higher turnout in those counties.

A difference-in-difference estimate is also shown in Table 5-3. The estimate was created by regressing the change in county turnout from 2002 to 2006 on the treatment variable. The results show a point estimate of .038, and a 90% confidence interval from .005 to .071. As the confidence interval does not overlap with zero it is unlikely that these results are attributable to chance alone, and we can conclude that the treatment variable likely has a positive effect on turnout of between .5 and 7.1%.

We might also note that some media reports claimed that due to lines at polls as many as 20,000 voters left polling sites in Denver without having voted (Merritt 2006). If the estimates are accurate, then the effects of vote centers on participation would be even greater. Turnout increased in Denver based on the available turnout data by 13.6 points from 2002 to 2006. If there were 20,000 more voters that left polling locations in 2006 without having voted then the increase would have been 20.98%, one of the highest increases in the state.

The analysis of aggregate turnout data provides evidence regarding the effects of vote centers on turnout but does not explicitly control for other variables that have been found to be related to turnout. While the cross-time and cross-sectional comparisons allow some indirect control for enduring county-level traits and general characteristics of the elections in 2006 they not permit direct control of individual-level potential confounding variables. The aggregate data also do not allow us to test how vote centers might conditionally affect turnout. To supplement the analysis of the aggregate county data we can analyze data from a post-election survey of Colorado residents.
The data analyzed in this section were obtained from post-election telephone interviews of residents of Colorado from November 8th to November 17th, 2006. The subjects were randomly chosen from voter registration lists obtained after the registration deadline in Colorado. A total of 603 interviews were obtained from a sample of registered voters on a number of political variables.

Using the data from the voter registration files, respondents' residential addresses were matched to counties and data were obtained on the treatment variable of interest to determine if the individual was exposed to vote centers. A breakdown of the respondents by treatment condition and county of residence is shown in Table 5-4. One feature of this table is that Denver and Douglas counties are heavily represented in the treatment sample, which may be problematic for the analysis, as these two counties in particular have lower turnout than the others. Looking at 2002, the statewide average turnout rate across counties was 56%, while Denver and Douglas counties had turnout of 44.5% and 52.7% respectively. Similarly in 1998 Denver and Douglas counties had lower rates of turnout (46.9 and 54.6) than the statewide average (55.8). That these two counties are heavily represented in the treatment group may move the estimated effect of vote centers downward.

While the data might be problematic as a result of the over-representation of Denver and Douglas counties in the treatment group we might be able to make adequate adjustments to account for these differences. That is, it is not known why turnout is lower in these counties than others in Colorado, but it could be a result of differences in individual characteristics of the populations in the counties. In this way we might be able to account for individual differences between voters in different counties so that while the population of Denver County does not vote with the same regularity as residents in Boulder, for example,
we might be able to identify individuals that live in each of the counties that would be comparably predisposed to voting.

**Outcome Variable** The main outcome variable of interest is whether an individual voted or not. The first source of data for the outcome variable is taken from the survey, which asked individuals whether or not they voted in the election on November 7th, 2006. The question wording is included in Appendix C. The responses to this question were then verified using official registration records. Verifying the survey responses enables us to adjust for over-reporting of electoral participation. Using this information we create a dichotomous variable that has a value of 1 if the individual voted in the November, 2006 election and 0 if she did not.

**Treatment Variable** The main treatment variable of interest is exposure to Election Day vote centers. The treatment measure is constructed using several sources of information. The first is the list of counties that used vote centers in 2006 which, as discussed above, is constructed using information from a survey of county clerks, information from the Colorado Secretary of State’s office, and county websites. The voters are then matched to counties using the residential addresses in the registration files. Using this information, a dichotomous treatment variable is constructed which takes a value of 1 if the voter is exposed to vote centers and 0 if not.

**Confounding Variables** To assess the effects of vote centers we would ideally be able to conduct a randomized experiment. The main advantage of a randomized experiment is that it
helps to ensure that the treatment and control groups are balances on all other factors. Without randomization the treatment assignment process is occurs in the natural world and is governed by an unknown and likely complex process. This can result in substantial biases in the data as the groups might differ from on another on other factors that are also related to turnout. To attempt to eliminate possible sources of bias we control for a number of factors that have previously been identified as being related to voter turnout. In particular we identify 15 potentially confounding variables that might be related to voting, which are discussed in Appendix C. It specifically provides a list of the control variables and how each are measured.

To adjust for potentially confounding variables we use a matching procedure to construct a sample of matched pairs. The advantages of matched sampling are discussed in earlier chapters. For this analysis, there is a binary treatment variable and thereby different techniques are used to construct the matched sample than those described in Chapter 3. The justification for the matched sample procedure with a binary treatment variable is the same as with a continuous treatment variable.

The matching procedure uses estimated propensity scores to reduce the dimensionality of the matching variables in order to create a sample of matched pairs. There were 176 cases that were exposed to the treatment in the data. These cases were then matched to the nearest available case (which minimizes the difference in propensity scores) out of the 325 available controls. Matching was done so that the treatment cases were exactly matched to the controls on prior voting experience and then the case with the closest propensity score from among those was selected. Following the matching procedure the
sample consisted of 175 matched pairs (1 treated observation could not be matched to a control) and a total sample size of 350 records.

Before analyzing these data it is important to assess the balance in the matched sample. The primary rationale for matching is to help remove the observed possible sources of bias by creating treatment and control groups that have more comparable distributions of the control variables. We can use a number of procedures to compare the full data to the matched sample to assess the improvement in balance. Numerical summaries are included in Table 5-5. This table shows the sample means for the full data set, the matched sample, and the differences in group means on the covariates. From this table we can see that matching generally reduced the difference in the group means on the covariates in the study. We can also graphically assess the distributions across the groups. Figure 5-3 shows a Q-Q plot of the distance measure for the treatment and control groups in the raw data and the matched sample. We can see from this figure that the treatment and control groups in the matched sample are more similarly distributed on the distance measure than with the raw data. Figure 5-4 shows density estimates of the propensity scores for the treatment and control samples, which shows a substantial amount of overlap between the two samples. From both the table and graphs we can see that the covariates tend to be better balanced between the treatment and control groups in the matched sample than before matching.

To examine the effects of vote centers using the matched sample a breakdown of turnout across treatment groups is shown in Table 5-6 along with McNemar's test for the null hypothesis of no effect. McNemar's test of no effect is appropriate for analyzing binary treatment and outcome variables with matched pairs. The p-value was calculated by using
discordant pairs (with one observation in the pair having voted, the other not voting) in the formula:

\[ T = t(Z, R) = \sum A_i q_i \]

For McNemar’s test, \( q \) is a vector of 1’s with a length equal to the number of discordant pairs in the data. \( A \) is a vector of identical length with elements determined by the following equation:

\[ A_i = \begin{cases} 1 & \text{if } (Z_{1i} - Z_{12})(R_{11} - R_{12}) > 0, \\ 0 & \text{if } (Z_{1i} - Z_{12})(R_{11} - R_{12}) \leq 0. \end{cases} \]

\( T \) is then equal to the number of discordant pairs in which the treated case has a higher value of the response (Rosenbaum 2002). This value of the test statistic was then used to find the p-value for the null hypothesis of no effect. The p-value was calculated as:

\[ pr(T \geq k \mid Z, R) = \sum_B I\left( \sum_{i=1} b_i q_i \geq k \right) \left( \frac{1}{2} \right)^I \]

Where \( k \) is the calculated value of the test statistic, \( I(\text{event}) \) is 1 if true, 0 otherwise, and \( B \) is the set of \( 2^I \) distinct vectors of length I (with coordinates 0 and 1). For McNemar’s test of no effect this is equivalent to \( pr(T \geq k) \) of a binomial probability distribution with \( I \) trials and a probability of success of .5 (assuming that all bias has been removed and the treatment is without effect). Based on this calculation, the probability that 43 out of 70 discordant pairs would have the control observation voting if the treatment was without effect is .2, which is greater than conventional levels of statistical significance and does not allow us to rule out the null hypothesis of no treatment effect.

One potential concern with the null hypothesis test reported above is that assumes that all unobserved sources of bias have been eliminated. That is, it assumes that the treatment and control observations are identical to one another on all other factors that are
related to voting. In an observational study, it is difficult to justify this assumption for two reasons. The first is that while matching reduces observed differences between the groups it does not eliminate them entirely. This concern is addressed in the following section. The second, and more serious concern, is that there are unobserved factors that might also be related to voting but for which we cannot explicitly control (Rosenbaum and Rubin 1983b). For example, recent research suggests that a substantial portion of the variation in turnout is attributable to genetic differences that have also been linked to regard for others (Fowler and Dawes 2008). Yet there is no data that are available to directly measure such genetic variables on our sample. There may additionally be unobserved variables that we do not even anticipate that might be related to voting. While matching helps to reduce observed differences between the groups it does not provide any direct adjustment for unobserved differences between the groups.

This concern raises the possibility that the treatment and control samples might systematically differ from one another on other factors that are related to voting. This introduces the possibility that the results might not be attributable to the effects of the treatment but rather might be reflect imbalances of unobserved variables. While there is no way to directly adjust for unobserved variables, we can assess how an unobserved factor might affect the results.

Two types of sensitivity parameters are considered in this analysis. The first was proposed by Rosenbaum (1987) and further discussed in Rosenbaum (2002). This part of the analysis considers how the results of the test would be affected if the treatment and control groups differ on an unobserved binary factor that is an almost perfect predictor of the outcome. By considering various possible differences between the groups, we can assess
how much different they would have to be in order to affect the results. If only very large
differences would affect the results then we can be more confident in our conclusions than if
very small differences could alter the findings.

The second parameter, discussed in Gastwirth, Krieger, and Rosenbaum (1998),
generalizes this procedure. The second approach considers a binary unobserved factor that is
related to the treatment and outcome variables at different levels and how it might affect the
results. Rather than assuming that the unobserved factor is an almost perfect predictor of the
outcome variable this alternative approach allows us to simultaneously consider the strength
of the relationship between the unobserved factor and the outcome variable. This approach
allows us to vary both parameters simultaneously to assess the sensitivity of the results
(Gastwirth, Krieger, and Rosenbaum 1998). As the strength of the association between the
unobserved factor and the outcome variable increases, the dual sensitivity analysis is
identical to Rosenbaum’s original test.

The sensitivity analysis considers an unobserved binary variable scaled to be 0 or 1.
The restriction on the scale of the unobserved variable to the unit interval is imposed so that
the parameters in the model for the sensitivity analysis are interpretable. The restriction of
the factor to 0 or 1 has been discussed in other research that allows for alternative dispersions
of the unobserved factors but that do not have a substantial effect on the sensitivity analysis
(Gastwirth et. al. 1998, Rosenbaum 1987).

The sensitivity analysis considers how the results might change given varying aspects
of hidden bias along two dimensions. The first consideration is for different possible
distributions of the excluded covariate across the treatment and control groups. If the
distribution of the unobserved covariate was the same for the treatment and control groups
the results would be the same as above. We can also consider different distributions of the
covariate were it might be more prevalent in either the treatment or control groups. In this
case the results might be confounded. By considering different possible imbalances for an
unobserved covariate we can assess whether plausible imbalances could account for our
results.

The second factor in the analysis is how strongly the unobserved factor is related to
the outcome. The stronger the relationship the more likely it is that the unobserved variable
might affect the results. If the findings would hold up to even in the presence of an
unobserved variable that was strongly related to the treatment assignment and the outcome
variable then we can be more confident in reaching a conclusion about the causal effects of
the vote centers on turnout from these results.

The results are shown in Figure 5-5. The figure shows the parameter for the
association with the treatment assignment on the horizontal axis, the association with the
outcome on the vertical axis, and the densities with varying p-values on the X-Y plane. From
this figure we can see that the results are not robust to plausible levels of association with
either variable as there is a notable increase the in p-value even at low levels of hidden bias.

A further analysis was done using model-based adjustments to account for remaining
imbalances in the data following matching. A logistic model is estimate and the results are
shown in Table 5-7. The model also includes a measure of distance to further adjust for
remaining imbalances in the distance measure. We can see from these results that further
accounting for imbalances in the observed covariates does not change the sign of the estimate
but the confidence interval does overlap zero and some positive values.
We can use a similar approach to assess whether the effects of vote centers vary depending on individuals' prior frequency of voting. One criticism of previous reforms is that they have only made voting more convenient for habitual voters (Berinsky 2005). By considering individuals' prior vote histories we can assess whether or not vote centers are any more efficacious at turning out less engaged voters. To test this possibility the model from above is amended to include a measure of individuals' vote histories from the past six years and an interaction with the treatment variable. If vote centers are more efficacious for less engaged voters we should see a negative sign on the interaction term. The model estimates are shown in Table 5-8.

Based on these results we can see that vote centers are more efficacious for less engaged voters. We can also see that estimate for the treatment variable is positive in this model. Yet it is not clear what the substantive effects of the results are from the estimates in Table 5-8. In particular the effects of the treatment variable are not clear as positive estimate for the treatment and the negative estimate for the interaction might net out to a different effect for more and less habitual voters. To assess the substantive effects of vote centers on individuals with different levels of engagement we can consider the results shown in Figure 5-6. This figure plots the estimated effects of vote centers over the range of prior vote histories. From this figure we can see that the reform is more efficacious for less engaged individuals than for moderately and highly engaged voters though with smaller samples the confidence bands are particularly wide, suggesting some support for the second hypothesis.

5.5.3 What Happened in Denver?
There could be other areas of implementation where the counties might have varied in how they administered the Election Day vote centers. The two counties from our sample that experienced the most significant problems on Election Day were Denver and Douglas counties, although those issues were not necessarily related to the Election Day vote centers. According to a report from the Colorado Secretary of State’s Office (2007), long lines formed at Denver’s polling locations because the new electronic poll book software they were using broke down and slowed the check-in process. The report found that Denver used a single server with the electronic poll book rather than five, which is used in Larimer county, the software was designed so that sessions stayed open even after the browser had been closed by the poll workers, and that Denver’s election officials had not tested the software to insure that it could handle the volume of voters. The delays from the new poll book software led to long lines at Denver’s polling locations, and anecdotally led some voters to leave the polling locations, which might lead to lower turnout.

Long lines also formed in Douglas County which might have led to lower voter turnout. According to the Colorado Secretary of State’s (2007) report, the problem in Douglas County is that they did not have enough electronic voting machines available at the polling locations. According to the report, officials in Douglas County underestimated the time it would take individuals to vote and thus the number of machines they would need at each polling site. With too few machines at the polling locations, long lines formed at some voting locations in Douglas County and might have also led to lower turnout as some might have left the polling location without voting.

Neither one of these factors, either the breakdown of the electronic poll book software or the equipping of the polling places are characteristic of Election Day vote centers, but
rather raise more general questions about the administration of elections. For our study, these factors raise at least two important questions. The first is whether or not it is plausible that these problems could have led to the long lines at the polling locations, and the second is whether or not Denver and Douglas counties are driving the survey results.

The first question is whether or not the long lines reported in Denver and Douglas counties could have been caused by delays in checking in voters and insufficient equipment at the polling locations. To evaluate these claims, we set up a preliminary simulation of a simplified Election Day voting site. To conduct the first simulation, we used a hypothetical polling location with 5 voting machines that was open for 12 hours. We assumed that voters would arrive at the polling location according to a Poisson process where we varied the parameter from 1 to 5, and would depart the polling location according to a Poisson process with an expectation that the voters would take approximately 3 minutes to cast a ballot.\textsuperscript{16} We then tracked the polling location minute by minute, and stored the average (median) length of the queue. The results are displayed in Figure 5-7.

An interesting feature of these results is the turning point at about 4 arrivals per minute. At that point, we can see that the lines at the simulated polling locations begin growing at a rapid rate. The results suggest that 1 to 4 arrivals per minute led to average wait times between 0 and 20 minutes, which would probably be acceptable to most voters, while between 4 and 5 arrivals per minute the average length of the line goes from about 20 minutes to almost 80 minutes. To assess the possible effects of increasing the resources at the polling location, we reran the simulation with 6 voting booths, and the critical point shifted up by about 1 additional voter arriving per minute, so the upturn in the graph began at

\textsuperscript{16} The departure times are based on observational data that we collected for another project from the time it took actual voters to cast a ballot using electronic voting machines.
about 5 arrivals per minute rather than four, as shown in Figure 5-7. This might suggest that there is a critical point of voter arrivals by minute where the lines really begin to grow, so that voting sites do not necessarily need to make dramatic changes, but rather just need to have the resources to stay below the critical point. It also suggests that these factors could have led to long lines in Denver and Douglas counties if they fell to the right side of the turning point, at which lines might grow at a rapid rate.

The second question for our analysis is whether Denver and Douglas counties might be driving our initial set of results. Denver and Douglas counties are heavily represented in the treatment sample, as 68.79% of our subjects from vote center counties live in Denver or Douglas. To evaluate the possible implementation effects on our analysis, we re-ran the logistic models with a control variable for Denver and Douglas counties. The results are displayed in Table 5-9, which show that the effects were negative for those two counties and the overall effects in the interaction model are signed as expected, but also have a wider confidence interval. These results are consistent with the conjecture that the implementation issues in Denver and Douglas counties might be influencing our results.

5.6 Conclusion

This chapter considers the effects of Election Day vote centers on voter turnout and if the effects of vote centers vary based on individuals’ propensity to vote. The results from the analyses were mixed. Using aggregate data from Colorado counties in 2006 a difference-in-difference estimate suggests that vote centers have a positive effect on turnout. Using data from a post-election survey of registered voters suggests that vote centers may have a negative effect, although the estimate was not strong and might be affected by hidden biases.
The individual results suggest that vote centers may be more efficacious for less engaged voters, as hypothesized, but in light of the overall negative estimated effect from the survey data we should not draw definitive conclusions from these results.

Taken together, the most reliable results are likely the aggregate data. The measures of the treatment and outcome variables are reliable and unlikely to have substantial measurement errors. Furthermore, the difference-in-difference approach allows us to address concerns about potential imbalances in the data on confounding variables. That vote centers seem to be positively related to turnout from the aggregate data suggests that the overall relationship is likely to be positive (increases turnout by about 3% of registered voters).

The results from the individual data are illuminating although somewhat less definitive. There may be issues with the response sets and the over-representation of Denver and Douglas counties in the treatment samples. These counties have historically lower turnout than the rest of the state and while we attempt to control for a number of factors that are related to turnout we cannot be certain that we have eliminated all relevant sources of bias. The results might thus be confounded and could underestimate the effects of vote centers, particularly in light of past research and the aggregate results which suggest a positive effect of the treatment.

The paper additionally addresses the potential problems in Denver and Douglas counties. In particular, those counties experienced long lines at polling locations, which raises questions as to how the reform was implemented. The simulation in Section 5.3 suggests that the implementation of any system of voting does not have to miss by much to lead to substantial lines at polling locations. The simulations demonstrate that there is a clear critical point at which long lines begin to form. Below that critical point and additional
resources do not make much difference. Above the critical point and the lines grow rapidly.

The results suggest that while the outcome might make it appear that the election was poorly administered it might have been close to an effective administration. Based on the results of this simulation, long lines at polling locations in one election does not require a massive investment in new resources, rather sufficient resources must be allocated to shift the polling location beneath the critical point. The simulation suggests that the key concern should be staying below the critical point, but how far below that critical point would seem to be of less significance. This result might be particularly helpful for election administrators seeking to address lines at polling locations as it suggests that a massive investment of resources might not be necessary to alleviate lines, rather a model increase or even a more efficient allocation of existing resources might have a substantial effect on lines.
Table 5-25: Residential Distances to Polling Locations

<table>
<thead>
<tr>
<th></th>
<th>Full Sample</th>
<th></th>
<th>Within 5 Miles</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Median</td>
<td>Mean</td>
<td>Median</td>
</tr>
<tr>
<td>Precincts 2002</td>
<td>2.46</td>
<td>.61</td>
<td>.87</td>
<td>.51</td>
</tr>
<tr>
<td>EDVC 2003</td>
<td>1.54</td>
<td>1.11</td>
<td>1.34</td>
<td>1.08</td>
</tr>
</tbody>
</table>

Table 5-26: Average Turnout Over Counties

<table>
<thead>
<tr>
<th>Year</th>
<th>Precinct</th>
<th>EDVC</th>
<th>Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1994</td>
<td>0.670</td>
<td>0.628</td>
<td>0.042</td>
</tr>
<tr>
<td>1998</td>
<td>0.570</td>
<td>0.535</td>
<td>0.035</td>
</tr>
<tr>
<td>2002</td>
<td>0.578</td>
<td>0.527</td>
<td>0.051</td>
</tr>
<tr>
<td>2006</td>
<td>0.689</td>
<td>0.658</td>
<td>0.030</td>
</tr>
<tr>
<td>N</td>
<td>45</td>
<td>19</td>
<td></td>
</tr>
</tbody>
</table>

Table 5-27: Difference-in-Difference Regression

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>S.E.</th>
<th>90% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>.038</td>
<td>.020</td>
<td>.005 - .071</td>
</tr>
<tr>
<td>Constant</td>
<td>.107</td>
<td>.01</td>
<td>.089 - .124</td>
</tr>
<tr>
<td>R²</td>
<td>.056</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>64</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 5-28: Proportion of Treatment Cases by County

<table>
<thead>
<tr>
<th>County</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Denver</td>
<td>.392</td>
</tr>
<tr>
<td>Adams</td>
<td>.273</td>
</tr>
<tr>
<td>Douglas</td>
<td>.250</td>
</tr>
<tr>
<td>Delta</td>
<td>.028</td>
</tr>
<tr>
<td>Grand</td>
<td>.017</td>
</tr>
<tr>
<td>Chaffee</td>
<td>.017</td>
</tr>
<tr>
<td>Eagle</td>
<td>.011</td>
</tr>
<tr>
<td>Archuleta</td>
<td>.011</td>
</tr>
</tbody>
</table>
Table 5-29: Matched sample

<table>
<thead>
<tr>
<th>Variable</th>
<th>All Data</th>
<th></th>
<th></th>
<th>Matched Sample</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Treated</td>
<td>Control</td>
<td>Difference</td>
<td>Treated</td>
<td>Control</td>
<td>Difference</td>
</tr>
<tr>
<td>Distance</td>
<td>.413</td>
<td>.32</td>
<td>.093</td>
<td>.412</td>
<td>.37</td>
<td>.042</td>
</tr>
<tr>
<td>Attentive</td>
<td>1.5</td>
<td>1.42</td>
<td>.08</td>
<td>1.50</td>
<td>1.47</td>
<td>.03</td>
</tr>
<tr>
<td>Mobilize</td>
<td>1.19</td>
<td>1.13</td>
<td>.06</td>
<td>1.19</td>
<td>1.17</td>
<td>.02</td>
</tr>
<tr>
<td>Negative</td>
<td>2.32</td>
<td>2.09</td>
<td>.23</td>
<td>2.33</td>
<td>2.14</td>
<td>.19</td>
</tr>
<tr>
<td>Preference</td>
<td>1.22</td>
<td>1.19</td>
<td>.03</td>
<td>1.22</td>
<td>1.21</td>
<td>.01</td>
</tr>
<tr>
<td>Group</td>
<td>1.87</td>
<td>1.79</td>
<td>.08</td>
<td>1.86</td>
<td>1.79</td>
<td>.07</td>
</tr>
<tr>
<td>Discuss</td>
<td>2.54</td>
<td>2.55</td>
<td>-.01</td>
<td>2.55</td>
<td>2.54</td>
<td>.01</td>
</tr>
<tr>
<td>GovTrust</td>
<td>2.87</td>
<td>2.94</td>
<td>-.07</td>
<td>2.87</td>
<td>2.93</td>
<td>-.06</td>
</tr>
<tr>
<td>Civic</td>
<td>2.16</td>
<td>2.04</td>
<td>.12</td>
<td>2.16</td>
<td>2.15</td>
<td>.01</td>
</tr>
<tr>
<td>Ideology</td>
<td>1.97</td>
<td>2.17</td>
<td>-.2</td>
<td>1.97</td>
<td>2.01</td>
<td>-.04</td>
</tr>
<tr>
<td>News</td>
<td>6.32</td>
<td>6.32</td>
<td>0.0</td>
<td>6.33</td>
<td>6.15</td>
<td>.18</td>
</tr>
<tr>
<td>Residence</td>
<td>13.16</td>
<td>13.16</td>
<td>0.0</td>
<td>13.16</td>
<td>12.45</td>
<td>.71</td>
</tr>
<tr>
<td>Education</td>
<td>4.52</td>
<td>4.77</td>
<td>-.25</td>
<td>4.51</td>
<td>4.57</td>
<td>-.06</td>
</tr>
<tr>
<td>Female</td>
<td>1.6</td>
<td>1.5</td>
<td>.1</td>
<td>1.6</td>
<td>1.57</td>
<td>.03</td>
</tr>
<tr>
<td>Partisan</td>
<td>2.99</td>
<td>2.86</td>
<td>.13</td>
<td>2.99</td>
<td>2.95</td>
<td>.04</td>
</tr>
<tr>
<td>Anglo</td>
<td>.82</td>
<td>.94</td>
<td>-.12</td>
<td>.83</td>
<td>.91</td>
<td>-.08</td>
</tr>
<tr>
<td>Prior Vote</td>
<td>4.19</td>
<td>4.25</td>
<td>-.06</td>
<td>4.14</td>
<td>4.14</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Table 5-30: Results of Matched Sample

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th>Treatment</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vote</td>
<td>118</td>
<td>113</td>
<td>231</td>
</tr>
<tr>
<td></td>
<td>67.43</td>
<td>64.57</td>
<td></td>
</tr>
<tr>
<td>Did not vote</td>
<td>57</td>
<td>62</td>
<td>119</td>
</tr>
<tr>
<td></td>
<td>32.57</td>
<td>35.43</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>175</td>
<td>175</td>
<td>350</td>
</tr>
<tr>
<td>p= .266</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
### Table 5-31: Logit model estimates

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>-.127 (.226)</td>
<td>-.077 (.229)</td>
<td>-.274 (.328)</td>
</tr>
<tr>
<td>Attentive</td>
<td>-.506 (.273)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mobilize</td>
<td>-.266 (.432)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>.046 (.142)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Preference</td>
<td>-.435 (.394)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td>-.014 (.183)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discuss</td>
<td>.020 (.221)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GovTrust</td>
<td>-.168 (.268)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civic</td>
<td>-.104 (.178)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ideology</td>
<td>.335 (.212)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>News</td>
<td>.033 (.088)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Residence</td>
<td>-.038 (.018)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>.266 (.164)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>.012 (.334)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-.020 (.013)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partisan</td>
<td>-.300 (.148)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anglo</td>
<td>-.824 (.524)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior Vote</td>
<td>.732 (.083)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Distance</td>
<td>-1.12 (.755)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>.728 (.161)</td>
<td>1.14 (.325)</td>
<td>1.90 (2.10)</td>
</tr>
</tbody>
</table>

N= 350  350  350
Table 5-32: Logit estimates with interaction effect

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>.481 (.416)</td>
<td>.514 (.419)</td>
<td>.589 (.460)</td>
</tr>
<tr>
<td>Attentive</td>
<td>-.548 (.285)</td>
<td>-.235 (.435)</td>
<td>.053 (.142)</td>
</tr>
<tr>
<td>Mobilize</td>
<td>-.440 (.402)</td>
<td>-.004 (.224)</td>
<td>-.192 (.272)</td>
</tr>
<tr>
<td>Negative</td>
<td>-.039 (.186)</td>
<td>-.084 (.179)</td>
<td>-.348 (.216)</td>
</tr>
<tr>
<td>Preference</td>
<td>-.024 (.089)</td>
<td>-.192 (.272)</td>
<td>-.348 (.216)</td>
</tr>
<tr>
<td>Group</td>
<td>-.040 (.018)</td>
<td>-.057 (.342)</td>
<td>-.023 (.013)</td>
</tr>
<tr>
<td>Discuss</td>
<td>-.882 (.532)</td>
<td>-.303 (.151)</td>
<td>-.023 (.013)</td>
</tr>
<tr>
<td>GovTrust</td>
<td>-.882 (.532)</td>
<td>-.303 (.151)</td>
<td>-.023 (.013)</td>
</tr>
<tr>
<td>Civic</td>
<td>-.882 (.532)</td>
<td>-.303 (.151)</td>
<td>-.023 (.013)</td>
</tr>
<tr>
<td>Ideology</td>
<td>-.859 (.135)</td>
<td>.856 (.134)</td>
<td>1.01 (.148)</td>
</tr>
<tr>
<td>News</td>
<td>.859 (.135)</td>
<td>.856 (.134)</td>
<td>-1.46 (.316)</td>
</tr>
<tr>
<td>Residence</td>
<td>.859 (.135)</td>
<td>.856 (.134)</td>
<td>-1.15 (.487)</td>
</tr>
<tr>
<td>Education</td>
<td>.263 (.164)</td>
<td>-.366 (.156)</td>
<td>1.90 (2.14)</td>
</tr>
<tr>
<td>Female</td>
<td>.263 (.164)</td>
<td>-.366 (.156)</td>
<td>1.90 (2.14)</td>
</tr>
<tr>
<td>Age</td>
<td>-.057 (.342)</td>
<td>-.366 (.156)</td>
<td>1.90 (2.14)</td>
</tr>
<tr>
<td>Partisan</td>
<td>-.057 (.342)</td>
<td>-.366 (.156)</td>
<td>1.90 (2.14)</td>
</tr>
<tr>
<td>Anglo</td>
<td>-.057 (.342)</td>
<td>-.366 (.156)</td>
<td>1.90 (2.14)</td>
</tr>
<tr>
<td>Prior Vote</td>
<td>-.303 (.151)</td>
<td>-.115 (.487)</td>
<td>1.90 (2.14)</td>
</tr>
<tr>
<td>Treat*Prior Vote</td>
<td>-.416 (.159)</td>
<td>-.416 (.159)</td>
<td>1.90 (2.14)</td>
</tr>
</tbody>
</table>

N = 350
Table 5-33: Controlling for Denver and Douglas counties

<table>
<thead>
<tr>
<th>Variable</th>
<th>Model 1</th>
<th>Model 2</th>
<th>Model 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>-.172 (.307)</td>
<td>.727 (.468)</td>
<td>1.19 (.548)</td>
</tr>
<tr>
<td>Attentive</td>
<td></td>
<td>-.535 (.291)</td>
<td></td>
</tr>
<tr>
<td>Mobilize</td>
<td></td>
<td>-.185 (.445)</td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td></td>
<td>.024 (.144)</td>
<td></td>
</tr>
<tr>
<td>Preference</td>
<td></td>
<td>-.516 (.403)</td>
<td></td>
</tr>
<tr>
<td>Group</td>
<td></td>
<td>-.056 (.189)</td>
<td></td>
</tr>
<tr>
<td>Discuss</td>
<td></td>
<td>.002 (.228)</td>
<td></td>
</tr>
<tr>
<td>GovTrust</td>
<td>-.163 (.274)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Civic</td>
<td></td>
<td>-.086 (.180)</td>
<td></td>
</tr>
<tr>
<td>Ideology</td>
<td></td>
<td>.308 (.220)</td>
<td></td>
</tr>
<tr>
<td>News</td>
<td></td>
<td>.026 (.091)</td>
<td></td>
</tr>
<tr>
<td>Residence</td>
<td>-.043 (.018)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>.314 (.168)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>-.064 (.348)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>-.027 (.014)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Partisan</td>
<td>-.262 (.152)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anglo</td>
<td>-.100 (.548)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prior Vote</td>
<td></td>
<td>.859 (.135)</td>
<td>1.03 (.150)</td>
</tr>
<tr>
<td>Treat*Prior Vote</td>
<td>-.348 (.157)</td>
<td>-.404 (.162)</td>
<td></td>
</tr>
<tr>
<td>Denver / Douglas</td>
<td>.074 (.328)</td>
<td>-.466 (.419)</td>
<td>1.01 (.463)</td>
</tr>
<tr>
<td>Constant</td>
<td>.728 (.161)</td>
<td>-.146 (.316)</td>
<td>1.97 (2.16)</td>
</tr>
</tbody>
</table>

N= 350 350 350
Figure 5-1: Distance to Precinct and Vote Center Locations

Polling Place and Residential Distances

- Precinct 2002
- Vote Centers 2003
- Mean Distances

Distance (miles) vs. Week
Figure 5-2: Maps of Precincts, Vote Centers, and Registered Voter Populations

2002 Larimer County Precinct Locations

2003 Larimer County Vote Center Locations

Larimer County Registered Voter Addresses
Figure 5-3: Q-Q Plots of Distance Measure

Figure 5-4: Propensity Score Densities
Figure 5-5: Results of Sensitivity Analysis

Figure 5-6: Substantive Effects Controlling for Denver and Douglas
Figure 5-7: Polling Place Simulation

Simulated Polling Place (12 Hours, 5 Stations)

Simulated Polling Place (12 Hours, 6 Stations)
Chapter 6: Conclusion

The findings that are reported in this dissertation have a number of policy implications and suggest directions for future research. The recent interest in election reforms can be traced back to the 2000 presidential election and the problems with the casting and counting of ballots in Florida that might have influenced the ultimate outcome of the presidential race. This led to greater interest in election reforms that might improve elections and forgo future problems similar to those that arose in 2000. Recent efforts to improve elections are designed (presumably) to facilitate voter turnout and to assure voters that the outcome of the election is determined by the preferences of the voters rather than procedures of the election such as where the places are located, the types of ballots used, the methods for counting ballots and so on.

While there seems to be a growing interest in issues of election reform, political science research on the effects of electoral procedures has not necessarily kept pace. As a result, there have been policymakers, interest groups, bureaucrats, judges, and commentators that make strong claims about the effects of a particular reform with little systematic evidence. For example, opponents of electronic voting machines have opposed their use generally and some have suggested using a voter verifiable paper audit trail (VVPAT) would assure verifiable elections and preserve public confidence. Yet there is little evidence that a VVPAT would enhance the security of the election or that voters view them as important. Similarly, voter ID laws have drawn spirited support and opposition for those that claim ID laws are integral to preventing voter fraud and those that argue it disenfranchises older,
minority, and lower SES voters. Even issues such as where to locate polling places has been contentious as in the Maryland Supreme Court’s decision to strike down an early voting law.

Despite claims made by different advocates on both sides of reform debates, there is often not very much systematic empirical evidence about the actual effects of the reforms. For example, regarding voter ID laws there is little evidence that polling place election fraud occurs on a large scale or that voter ID laws are an effective way of preventing such fraud. While political calculations as to which party would likely benefit from a particular reform will prevent empirical evidence from being solely determinative, it can still be useful as a means to at least partially inform policy choices and clarify which claims are more or less valid. This project examines three election reforms to help us better understand their effects.

The newest reform investigated in this project is Election Day vote centers. Vote centers have attracted the most recent attention and have spread from one county in 2003 to 19 counties in Colorado in 2006 to counties in 4 states in 2008. Other states and counties are also considering pilot projects for vote centers in upcoming elections. Theoretically, vote centers can be distinguished from precinct voting by two characteristics which are openness and centralization. As discussed above, these two characteristics are thought to make voting more accessible and convenient, particularly for the less engaged. The results of this study provide further evidence that vote centers seem to be associated with an increase in voter turnout. The results showed that analyzing county-level data, there was an increase in turnout in vote centers from 2002 to 2006 that surpassed the increase in the rest of the state. The results from the post-election survey presented a more mixed picture, but controlling for two counties with historically lower turnout than other counties in the state, the results show that vote centers positively affect turnout, particularly among the less engaged.
Two counties that adopted vote centers in 2006, Denver and Douglas counties, had problems with their pollbooks that led to long lines at vote center locations. While not directly related to vote centers, the chapter also considered how such problems might arise. The results showed that while long lines can be a major inconvenience for voters it does not necessarily mean that there was a major administrative failure on the part of election officials. The results of a queuing simulation show that there is a critical point in median wait times at polling locations such that lines can build up rapidly if there are not enough resources at the polling site to remain below the critical point in the efficient processing of voters. The bad news for election officials is that it does not take a major failure to experience long lines at polling locations. The good news is that it might not take a major resource investment to resolve the problem. Based on the simulation the relationship is nonlinear so a small investment of resources can have a large effect on lines at polling locations if it shifts the processing time of voters just below the critical point.

The results from the study of Election Day vote centers provide further evidence that vote centers might make voting more convenient and attract more individuals to the polls, particularly less engaged voters. This has potentially important implications as local election officials have been interested in vote centers as a way of reducing the costs of administering elections over time and enhancing administrative control over the process. In addition to benefiting local election officials, these results also suggest that it is more convenient for voters and might help to reduce the disparities between the population and the electorate. While these results, combined with earlier studies, are very encouraging it would be premature to conclude that vote centers should be universally adopted. The effects of vote centers might vary by local contexts, so for example they might be less effective in large
metropolitan counties, and we should furthermore investigate the mechanisms by which vote centers affect turnout so that some of the lessons of vote centers might be applied to precincts.

This project also investigates the effects of early voting. The empirical results show that more partisan and politically engaged individuals are more likely to vote early than the less partisan and unengaged. This finding provides further evidence that early voting is most attractive to individuals that are already likely to vote. Consistent with previous research and the theoretical arguments, while early voting might make voting more convenient for likely voters it does not necessarily attract new voters to the process. Furthermore, when considering turnout among all eligible voters, early voting seems to have a negative effect. Across sources of data and methods of analysis, turnout seems to decrease by 2-4\% as a result of early voting.

The negative association between early voting and turnout may be the result of a number of factors. Critics of early voting have argued that everyone voting on one day is an important social experience that is undermined by early voting. Additionally, there might be a resource tradeoff between early and Election Day voting. Insofar as there is a fixed allocation of resources for the election, early voting might divert some resources away from Election Day voting. This might hinder Election Day voting as there could be fewer resources available when less engaged individuals are more likely to vote and they might also be more sensitive to the administration of the election.

This possibility might be tested in future research and in particular we might consider whether early versus late adopters of early voting have had different experiences. Innovators of early voting might have been more likely to allocate additional resources for early voting so that it did not take away from Election Day voting. As the reform became more popular
and moved away from the innovators and the early adopters, officials might have been more willing to make tradeoffs between early and Election Day voting. This might lead to different patterns of turnout in early voting states depending on the timing of adoption of early voting.

The analysis of voter registration deadlines finds evidence that deadlines are negatively related to turnout, such that if deadlines are reduced expected turnout goes up. The analysis specifically finds that turnout increases by around .3% for every day of the registration deadline. There is an added effect of Election Day registration of around 1%, but otherwise the effect seems to be fairly linear across the possible lengths of the registration deadline.

One of the findings from the analysis of registration deadlines is that the effect on turnout is greater than the effect on rates of registration. That is, while there is evidence that deadlines affect rates of registration the magnitude of the effect is even larger for turnout. This result is counterintuitive and contrary to theoretical expectations which posit that registration rates should mediate the relationship between deadlines and turnout. While the evidence suggests that registration may mediate this relationship in part there is also evidence that some other mechanisms are also at work.

Several alternative explanations are discussed above, and the possibility of contagion effects of voting is particularly interesting. While there is, to my knowledge, no study of any election reform that incorporates possible contagion effects, both classical and contemporary research suggests that contagion processes may be at work (Lazarsfeld, Berelson, and McPhee 1954, Campbell et al. 1960, Kenney 1992, Huckfeldt and Sprague 1995, Fowler
2005, Nickerson 2008). A simulation was described earlier that showed how a contagion process might change the treatment effect.

There are two issues related to modeling a contagion process. The first is the topology of the network and the second are the dynamics of voting. Network topology has received attention not only from political scientists but also scholars working on the science of networks more generally (Barabasi 2002, Watts 2003, Newman, Barabasi, and Watts 2006). The dynamics of networks is much less well understood. To model network dynamics and recover an estimated treatment effect, I presented a simple simulation model above. In subsequent efforts to model contagion effects, I used two different types of contagion processes.

The first was an SIR model of biological contagion (Ball, Mollison, and Scalia-Tomba 1997). The SIR model stands for Susceptible, Infected, and Recovery which are the three states used to classify units according to their disease status. The SIR model has been used to study the spread of contagious diseases and computer viruses. Applied to the study of voting, susceptible individuals are presently disengaged from the process, but might become committed voters. Infected individuals are the committed voters who are politically engaged in the population. Recovery refers to voters that have for some reason become disengaged, perhaps because of a scandal, negative campaigns, their preferred candidate drops out of the race, or through some other event.

I simulated the treatment effect with an SIR model that was allowed to run on a random network of 200 individuals. In the control condition, individuals had a baseline probability of voting that ranged between .01 and .75 (at 100 different values) at the beginning of the simulation. The transmission probability of voting ranged from .01 to .75
(also at 100 different values) for every voter that an individual was directly connected to. The maximum probability of becoming infected was .95. The contagion process iterated over the network 10 times and the rate of turnout was then recorded. The process was repeated 10 times for each of the 10,000 possible combinations of the parameters and mean rates of turnout were recorded for every unique parameter vector.

The treatment case incorporated a greater baseline probability of voting such that it was 5% higher under treatment. This implies a direct treatment effect of 5% which can then be tracked over the simulation. Repeating the same steps for the treatment condition, mean rates of turnout were recorded and then differenced against the control condition to obtain treatment effects over the full range of the parameter values. The results are shown in Figure 6-1. As we can see in the figure, there is substantial variation in the treatment effect over the range of parameter values, dropping to near zero with high baseline probabilities. With high transmission and low baseline, we can see that the overall treatment effect is more than triple the direct treatment effect. These findings suggest that under certain network conditions the direct treatment effect can have substantial indirect effects as well so that the overall impact of the reform is more than double or even triple what would be expected in the absence of a contagion process. Similarly, some contagion conditions weaken the impact of the treatment likely due to ceiling effects on the probability of voting.

While the SIR model is a useful tool for analyzing the possible treatment effects in the presence of a contagion process, it might not accurately reflect the way in which voting behavior spreads over a population. While voting might be contagious in some sense it might not be contagious in the way that diseases are contagious. A contagious disease can be spread through casual contact, so one might catch the flu, for example, by sitting next to
someone who has the flu. Voting might be less likely to spread in that way as it might require greater interaction between individuals.

I also analyzed a behavioral contagion model which is similar to models that have been used to study the spread of innovations and fads (Gladwell 2000, Watts 2003). These models incorporate behavioral contagion by assuming that individuals are responsive to the behaviors of those around them so if every one of a person’s immediate contacts engage in some behavior the individual will also engage in the behavior. In the model, individuals are assumed to have different thresholds for behavior, so that if someone’s threshold was .5, they would engage in a behavior if more than half of their immediate contacts engaged in the behavior but would refrain if it were less than half. The different thresholds could also be used to classify individuals and identify innovators and early adopters that have low thresholds and are critical to starting trends.

To simulate a behavioral contagion process, I analyzed a hypothetical population of 200 voters on a random network with a mean connectivity that varied from .8 to 8. This range of values spans 1, at which point random networks have been shown to undergo a phase transition where a giant component forms connecting most of the nodes, whereas at values less than one a giant component does not form (Erdos and Renyi 1960). The interval was divided into 100 segments for the simulation. Each unit was also assigned a different threshold which was drawn from a normal distribution with means that varied from .1 to .75 divided into 100 segments. The simulation was then run for each of the 10,000 possible combinations of the parameter values. Under the control condition, the model was seeded with 10% self-actuating committed voters. These are the members of the population that will turnout regardless of the behaviors of those around them. The contagion process was
allowed to operate on the network for 10 iterations and the turnout rate was then recorded. This process was run 10 times for each of the values, and the mean turnout was recorded for every parameter combination.

The simulation was then repeated for the treatment condition for comparison. The simulation was identical under the treatment condition, except that the contagion was initially seeded with 15% committed voters. In this way, the reform has a direct effect of 5% on voters, before the contagion process operates on the network. The simulations were then conducted and the mean rates of turnout were recorded under treatment. Taking the difference between the treatment and control conditions, we can see a plot of the treatment effects under each of the simulation conditions. The results are presented in Figure 6-2. From this figure we can see that the treatment effects vary substantially over a range of plausible values. Starting with a direct treatment effect of 5% we can see that under certain conditions the effect can nearly double. The second feature of the graph is that most of the variation is over the mean connectivity parameter such that higher connectivity is associated with a greater treatment effect. The results suggest that it may be important for future research to account for possible contagion processes and suggests that we might be able to design much more effective election reforms by having a better understanding of possible contagion processes.

While this project contributes to theoretical and empirical questions in political science, one of the aims of the project is also to make a practical contribution to our knowledge of election reforms and hopefully lead to more informed policy choices. These ends are by no means mutually exclusive. As officials adopt different policies, there is an opportunity for scholars to test theoretical propositions and to improve empirical results.
Bringing theoretical perspectives and methodological tools to address a particular policy
question can also improve our understanding its effects if it should be continued, expanded,
revised, applied in some cases, not in others, etc. There are six specific policy
recommendations that follow from this research project.

1. Use of vote centers should proceed and potentially expand to a limited number of
   states and counties. While the results are not definitive and by no means obviate the
   need for future research, there is an accumulating body of research which suggests
   that it might increase turnout in the places where it has been implemented. We might
   recommend caution, however, as the reform has been adopted in a limited number of
   counties and it would seem premature to recommend that all counties move to the
   vote center model.

2. Vote centers might be particularly efficacious in counties with younger and more
   mobile populations that are less likely to be habitual voters. Vote centers might also
   be particularly effective in jurisdictions with short registration deadlines. Vote
   centers are shown to be particularly attractive to less habitual voters and thus might
   be more efficacious in some jurisdictions or when combined with other policies such
   as shorter registration deadlines.

3. If there are long lines at polling places in one election, that does not necessarily imply
   that a dramatic overhaul of the polling places is required for subsequent elections.
   Simulation results suggest that the effect of additional resource allocations on lines is
   non-linear and such that a moderate increase in resources, under particular conditions,
   can have a substantial effect on lines.
4. Early voting should be implemented carefully. It is likely to further skew the electorate towards more partisan and politically engaged individuals but might also lower turnout overall. As less engaged individuals are more likely to be responsive to resource shortages in the administration of elections and more likely to vote on Election Day, officials should be wary not to make substantial resource tradeoffs for the administration of early voting at the expense of Election Day voting.

5. Shortening registration deadlines increases turnout. Results from both county level data and survey data shows that shorter registration deadlines can have an appreciable effect on turnout.

6. The effects of registration deadlines are largely linear. The analysis of registration deadlines shows that they seem to have a linear effect on turnout. There is an added effect of Election Day registration, but it composes a relatively small amount of the total effect of deadlines. This is potentially important as some states have encountered political and administrative obstacles to Election Day registration. This result suggests that most of the effects of Election Day registration can be obtained by reducing registration deadlines, so even if the elimination of deadlines is not feasible participation could still be increased by shortening deadlines to a more limited extent.

The research reported in this project not only suggests these policy recommendations, but also help us to better understand theoretical issues related to voter turnout and suggests directions for future research. The findings suggest that future research might explore possible contagion effects of voting and resource tradeoffs that might result from early voting. These and other related issues might be the focus of future research projects, and hopefully
will continue to refine our scholarly understanding of elections and voter participation as well as suggesting ways to improve the conduct of elections.
Figure 6-1: Biological Contagion Results

Figure 6-2: Behavioral Contagion Results
Bibliography


Green, Donald; Alan Gerber; and David Nickerson. 2003. Getting Out the Vote in Local Elections: Results from Six Door-to-Door Canvassing Experiments. *Journal of Politics* 65 (4): 1083-1096.


Appendix A

Figure A1: Estimated County Turnout by State

<table>
<thead>
<tr>
<th>Alabama</th>
<th>Arizona</th>
<th>Arkansas</th>
</tr>
</thead>
<tbody>
<tr>
<td>California</td>
<td>Colorado</td>
<td>Connecticut</td>
</tr>
<tr>
<td>Delaware</td>
<td>Florida</td>
<td>Georgia</td>
</tr>
</tbody>
</table>

Figure A2: Estimated County Turnout by State

<table>
<thead>
<tr>
<th>Hawaii</th>
<th>Idaho</th>
<th>Illinois</th>
</tr>
</thead>
<tbody>
<tr>
<td>Indiana</td>
<td>Iowa</td>
<td>Kansas</td>
</tr>
<tr>
<td>Kentucky</td>
<td>Louisiana</td>
<td>Maine</td>
</tr>
</tbody>
</table>
Figure A3: Estimated County Turnout by State

Figure A4: Estimated County Turnout by State
Figure A5: Estimated County Turnout by State

Figure A6: Estimated County Turnout by State
Figures A7- A13: Matched Sample for Turnout

**Mean Age by Registration Deadline**

**Proportion Female by Registration Deadline**
Mean Education by Registration Deadline

Proportion Black by Registration Deadline
Proportion Other by Registration Deadline

Proportion Other

0 0.00 0.01 0.02 0.03 0.04
Figures A14- A20: Graphs for Registration Matched Sample

**Mean Age by Registration Deadline**

![Graph](image1)

**Proportion Female by Registration Deadline**

![Graph](image2)
Appendix B

Table B1: Turnout as a ratio of VAP

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>S.E.</th>
<th>90% Confidence Interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Voting</td>
<td>-.029</td>
<td>.002</td>
<td>-.032</td>
</tr>
<tr>
<td>Constant</td>
<td>.515</td>
<td>.001</td>
<td>.513</td>
</tr>
</tbody>
</table>

N= 21759

Table B2: Turnout as a ratio of VAP (Early Voting counties only)

<table>
<thead>
<tr>
<th></th>
<th>Coef.</th>
<th>S.E.</th>
<th>Coef.</th>
<th>S.E.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Early Voting</td>
<td>-.029</td>
<td>.002</td>
<td>-.023</td>
<td>.002</td>
</tr>
<tr>
<td>Age</td>
<td>.006</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>.000</td>
<td>.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anglo</td>
<td>.175</td>
<td>.005</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Competitiveness</td>
<td>.127</td>
<td>.009</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Presidential</td>
<td>.130</td>
<td>.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>.515</td>
<td>.002</td>
<td>-.021</td>
<td>.008</td>
</tr>
</tbody>
</table>

N= 14745


Table B3: Turnout as a ratio of CVAP

<table>
<thead>
<tr>
<th></th>
<th>All Counties</th>
<th></th>
<th>EV Only Counties</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef. (S.E.)</td>
<td>Coef. (S.E.)</td>
<td>Coef. (S.E.)</td>
</tr>
<tr>
<td>Early Voting</td>
<td>-.022 (.002)</td>
<td>-.018 (.001)</td>
<td>-.022 (.002)</td>
</tr>
<tr>
<td>Age</td>
<td>.006 (.000)</td>
<td>.006 (.000)</td>
<td>.000 (.000)</td>
</tr>
<tr>
<td>Income</td>
<td>.000 (.000)</td>
<td>.000 (.000)</td>
<td>.000 (.000)</td>
</tr>
<tr>
<td>Anglo</td>
<td>.130 (.004)</td>
<td>.144 (.005)</td>
<td></td>
</tr>
<tr>
<td>Competitiveness</td>
<td>.178 (.007)</td>
<td>.134 (.009)</td>
<td></td>
</tr>
<tr>
<td>Presidential</td>
<td>.128 (.001)</td>
<td>.132 (.002)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>.522 (.001)</td>
<td>-.008 (.006)</td>
<td>.522 (.002)</td>
</tr>
</tbody>
</table>

N= 21750  21071           14743  14267

Table B4: Turnout as a ratio of Registered Voters

<table>
<thead>
<tr>
<th></th>
<th>All Counties</th>
<th></th>
<th>EV Only Counties</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coef. (S.E.)</td>
<td>Coef. (S.E.)</td>
<td>Coef. (S.E.)</td>
</tr>
<tr>
<td>Early Voting</td>
<td>-.036 (.002)</td>
<td>-.024 (.002)</td>
<td>-.045 (.002)</td>
</tr>
<tr>
<td>Age</td>
<td>.001 (.000)</td>
<td>.001 (.000)</td>
<td>.000 (.000)</td>
</tr>
<tr>
<td>Income</td>
<td>.000 (.000)</td>
<td>.000 (.000)</td>
<td>.000 (.000)</td>
</tr>
<tr>
<td>Anglo</td>
<td>.123 (.005)</td>
<td>.150 (.005)</td>
<td></td>
</tr>
<tr>
<td>Competitiveness</td>
<td>.253 (.009)</td>
<td>.212 (.010)</td>
<td></td>
</tr>
<tr>
<td>Presidential</td>
<td>.141 (.002)</td>
<td>.146 (.002)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>.637 (.001)</td>
<td>.259 (.008)</td>
<td>.647 (.002)</td>
</tr>
</tbody>
</table>

N= 19879  19326           13468  13073
Appendix C

C1: Table of Vote Center Counties in 2006

<table>
<thead>
<tr>
<th>EDVC Counties in 2006</th>
</tr>
</thead>
<tbody>
<tr>
<td>Adams</td>
</tr>
<tr>
<td>Archuleta</td>
</tr>
<tr>
<td>Broomfield</td>
</tr>
<tr>
<td>Chaffee</td>
</tr>
<tr>
<td>Conejos</td>
</tr>
<tr>
<td>Denver</td>
</tr>
<tr>
<td>Douglas</td>
</tr>
<tr>
<td>Gilpin</td>
</tr>
<tr>
<td>Grand</td>
</tr>
<tr>
<td>La Plata</td>
</tr>
</tbody>
</table>

C2: Control Variables

Attentive

We would like to begin by asking you a few questions about this month's elections (that is, the elections on November 7, 2006). Some people don't pay much attention to political campaigns. How about you? Would you say that you were...

Very Interested .......................................................... 1
Somewhat Interested ...................................................... 2
Not very interested ........................................................ 3
DON'T KNOW (Don't read) .................................................. 8
REFUSED (Don't read) ...................................................... 9

Mobilize

As you know, the parties and candidates often try to get as many people to vote for them as possible. Did anyone call you up or come around and talk to you about the campaigns this year?

Yes .................................................................................... 1
No ..................................................................................... 2
DON'T KNOW (Don't read) .................................................. 8
REFUSED (Don't read) ...................................................... 9

Negative
Sometimes candidates for political offices run negative campaigns that target the other party, while sometimes they focus more on their own qualities. Thinking back to the campaigns for governor, how would you rate the overall tone on a 1 to 5 scale, 1 being MOSTLY NEGATIVE and 5 being MOSTLY POSITIVE?

Mostly negative.................................................................................................1
Somewhat negative............................................................................................2
Both negative and positive..................................................................................3
Somewhat positive...............................................................................................4
Mostly positive....................................................................................................5
DON'T KNOW (Don't read)..................................................................................8
REFUSED (Don't read)............................................................................................9

Preference

In this election, did you prefer one of the candidates for governor more than the others?

Yes .......................................................................................................................1
No.........................................................................................................................2
DON'T KNOW (Don't read)..................................................................................8
REFUSED (Don't read)............................................................................................9

Group

There are many different types of groups that people can belong to, such as labor unions, associations of people who do the same kind of work, hobby clubs, community groups, and school groups. Of course, there are lots of other types of organizations too. Not counting membership in a local church or synagogue, are you a member of any of these organizations? [IF YES, ASK:] "Are you active in any of the groups?"

Not a member of any group.................................................................................1
Member, inactive..................................................................................................2
Member, active......................................................................................................3
DON'T KNOW (DON'T READ)..............................................................................8
REFUSED (DON'T READ)....................................................................................9

Discuss

When you talk to people you know, about how often do you discuss politics? Would you say...

Never....................................................................................................................1
Only once in a while.............................................................................................2
Fairly often............................................................................................................3
Most times............................................................................................................4
DON'T KNOW (DON'T READ)..............................................................................8
REFUSED (DON'T READ)....................................................................................9

GovTrust (Trust in Government)
How much of the time do you think you can trust the government in Washington to do what is right? Would you say...
Just about always ......................................................... 1
Most of the time .......................................................... 2
Only some of the time .................................................... 3
Never (Volunteered) ..................................................... 4
DON'T KNOW (DON'T READ) ........................................... 8
REFUSED (DON'T READ) 9

Civic

Would you generally AGREE, DISAGREE, or NEITHER AGREE NOR DISAGREE with the following statement [READ ITEM ROTATED]? (And what about...) If a person doesn't care how an election comes out he shouldn't vote in it.
Agree ............................................................................. 1
Neither agree nor disagree .............................................. 2
Disagree ........................................................................... 3
DON'T KNOW (DON'T READ) ........................................... 8
REFUSED (DON'T READ) 9

Ideology

On most political matters, would you consider yourself to be...
Liberal ............................................................................ 1
Somewhat liberal .............................................................. 2
Moderate (middle of the road) ......................................... 3
Somewhat conservative ................................................... 4
Conservative .................................................................... 5
DON'T KNOW (DON'T READ) ........................................... 8
REFUSED (DON'T READ) 9

News

During the past week, about how many days did you read a daily newspaper, consult an online news source, or watch a national news network program?  [ENTER NUMBER OF DAYS]
*
DON'T KNOW (DON'T READ) ........................................... 8
REFUSED (DON'T READ) 9

Residence

About how long have you lived at your CURRENT ADDRESS?  [RECORD NUMBER OF YEARS]
**
DON'T KNOW (DON'T READ) ........................................... 88
REFUSED (DON'T READ) 99
Education

What is the most EDUCATION that you've completed: Grade school, some high school, graduated high school, some college, bachelor's degree, or a professional degree?
Grade school.................................................................1
Some high school.........................................................2
Graduated high school....................................................3
Some college........................................................................4
Bachelor's degree ...........................................................5
Professional degree ..........................................................6
DON'T KNOW (DON'T READ) .............................................8
REFUSED (DON'T READ) ................................................9

Female

Female 0
Male 1

Age

Age in years

Partisan

Four point scale of partisan strength:
Independent 1
Leaning 2
Moderate 3
Strong 4

Anglo

Non-anglo 0
Anglo 1

Prior Vote

Count of prior vote history from Colorado voter files (0-13)