Five Studies on the Causes and Consequences of Voter Turnout

Share Your Story

Five Studies on the Causes and Consequences of Voter Turnout

A dissertation presented

by

Anthony George Fowler

to

The Department of Government

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Political Science

Harvard University
Cambridge, MA

March 2013

Advisor: James M. Snyder, Jr.   Author: Anthony George Fowler

Abstract

In advanced democracies, many citizens abstain from participating in the political process. Does low and unequal voter turnout influence partisan election results or public policies? If so, how can participation be increased and how can the electorate become more representative of the greater population?

Study #1 examines the adoption of compulsory voting laws in Australia in order to assess the effects of near-universal turnout. When new, working class voters were brought the polls Australia saw significant changes in election results and public policy. Study #2 explores the partisan consequences of marginal changes in voter turnout in the United States by comparing the partisan preferences of regular and marginal voters. Across three independent tests and settings, marginal voters are significantly more supportive of the Democratic Party compared to regular voters.

Studies 3 and 4 examine several potential cures for low turnout. Study #3 examines the common view that social capital and community participation will increase political participation, a claim that has not been rigorously tested. Exploiting the timing of saint's day fiestas in Mexico and a change in the Mexican electoral calendar, this study finds that social capital and community engagement appear to be substitutes rather than complements for political participation. Study #4 tests another widely held view that increased electoral competition will boost electoral participation, but a range of tests finds little support for this hypothesis.

Finally, Study #5 demonstrates that marginal increases in voter turnout will not necessarily make the electorate more representative of the greater population. Get-out-the-vote interventions have been shown to increase turnout significantly and they are widely touted as beneficial for democracy, but surprisingly, these interventions actually increase inequalities in turnout by primarily

mobilizing more of the types of citizens who were already voting. In short, low voter turnout and

inequalities in voter turnout have significant political and policy consequences, but the problem is

hard to fix. Even seemingly benign cures like social capital, electoral competition, and voter

mobilization have little benefit or may even be detrimental.

Table of Contents

After the course, Gabe continued to work with me and provide informal one-on-one training econometrics, statistical computing, data management, and big questions in political behavior and electoral accountability. Gabe taught me many things that I hope to pass along to future students, but there is one simple lesson that will stick with me forever: "always show your data." Marc Meredith calls this "Lenz's Law," and I hope it catches on.

Like many other political scientists I first contacted Jim because I needed data, and over the years, he has given out a lot! More importantly, Jim has been an excellent instructor, mentor, collaborator, and confidant. He loves every aspect of research, and his passion is contagious. Moreover his depth and breadth of knowledge and interest are unmatched in our field. I know that I can go to Jim with a question or idea in any topic from political psychology to social choice to econometrics to political history and I will learn something new. Jim takes his work and his advising responsibilities very seriously, and it shows. I will be sad to leave Cambridge and lose the ability to walk down the hall and knock on Jim's door.

In addition to my committee members, special acknowledgement is owed to Ryan Enos, Matthew Atkinson, and Lynn Vavreck who each co-authored part of this dissertation. I met Ryan and Lynn when I was a prospective student at UCLA which turned out to be a valuable and fortuitous trip. Soon afterward, Ryan connected me with Matthew because we shared some common interests. Despite our 3 year collaboration via e-mail, we have not yet met in person, but I hope to remedy that soon. All three of them have been excellent coauthors. We have agreed on the important things, but they all think about problems differently than I do and offer a fresh perspective. Perhaps most importantly, they are all engaging, passionate, enthusiastic, and curious, so the simple process of talking through our work generates new ideas and produces better work than any of us would create on our own.

My time spent in graduate school and several of my research undertakings would not have been possible without the institutional and financial support of the Graduate School of Arts and Sciences, the Department of Government, the Institute for Quantitative Social Science, the Center for American Political Studies, and the Committee on Australian Studies. I am grateful for the hard work of many administrators and staffers within those organizations who made my life easier including Thom Wall, Lilia Halpern-Smith, Janet Hatch, Craig Lajoie, Lisa Galvin, and Emily Burns. I am also grateful to the Department of Political Science at MIT and the James A. Lash Presidential Graduate Fellowship which supported my first year of graduate studies.

I greatly appreciate my parents, Pete and Diana, for their continual love, support, and encouragement. At every age, they fostered my intellectual curiosity and worked hard to provide me with the best available educational, economic, and career opportunities. Even during times of financial uncertainty and difficulty, they sent me to private schools, sent me to the University of Cambridge to take college-level courses when I was 13, and encouraged me to spend my summers volunteering in labs instead of taking a regular job. Without question, I would not be in this position of pursuing a Ph.D. without their endless support.

My greatest gratitude is owed to my wife, Gloria, although my appreciation for her cannot be justly put into words. She has read and edited every paper, heard me practice every presentation (many, many times in some cases), entered data by hand (the Bendigo data was truly a team effort!), worked out math that I was stuck on, and carefully engaged and critiqued each new research idea. She is a true partner in every respect of my life. In addition to all the direct ways in which she helped me with my work, she let me vent after moments of frustration, celebrated the successes, remained 100% optimistic about my career even when I wasn't, took multiple licensing exams since we didn't know where we would end up after graduate school, and never complained about the uncertainty of our future. Most importantly, she has made me the happiest person in the world

during a time of my life that would have otherwise brought stress and anxiety. This dissertation, along with all of my work, is dedicated to her.

Introduction

The 2008 U.S. presidential election was – by most popular and academic accounts – a transformative moment. Young people and ethnic minorities turned out to vote in record numbers, rallying behind the campaign slogans, "Change" and "Yes We Can." Despite the uniqueness of this election, 38 percent of voting-eligible citizens still failed to turn out, and the voters were systematically unrepresentative of the eligible population. As in every American election where reliable data is available, the voters were significantly wealthier, older, more educated, more likely to be white, and more likely to attend church compared to the eligible non-voters. Even in the extreme case of 2008, overall participation was far from the democratic ideal of universal representation, and young people, ethnic minorities, and the poor were significantly underrepresented in the electorate.

In virtually every election in every advanced democracy in the world, many citizens fail to participate in the democratic process, and those who turn out to vote are systematically unrepresentative of the eligible population. These phenomena pose a significant challenge for democratic government which was designed with the goal of translating public preferences into public policies. How can citizens' preferences be represented if they fail to participate in the simplest and most fundamental part of the democratic process – voting. In his presidential address to the American Political Science Association, Arend Lijphart aptly describes this challenge as "democracy's unresolved dilemma."

This dissertation presents five empirical studies which aim to make progress on this dilemma. The first two studies begin by assessing the extent of the dilemma. Is this really a dilemma at all? Would election results and public policies be different if everyone voted? What are the electoral and policy consequences of exogenous changes to the composition of the electorate? The next two studies test important hypotheses about the causes of voter turnout. What policy prescriptions might potentially increase participation and mitigate this dilemma? The final study

explores the consequences of traditional campaign mobilization strategies on the composition of the electorate. Even when turnout is increased by campaigns, practitioners, and progressive groups, is the electorate becoming more *or less* like the voting-eligible population?

Each study is written as a standalone paper, and the relevant literature for each one is discussed and cited within the papers. For that reason, my introductory remarks are brief. Nonetheless, the studies all originated from same set of motivating questions, and the same themes recur throughout. Moreover, the findings of the papers tell a largely consistent (and grim) story about the problems of low and unequal participation in elections. Below, I discuss some of those common themes, summarize the current state of evidence on these questions, discuss the relationships between the papers, and suggest several paths for future research.

One recurring theme in the studies is the inability of simple, observational analyses to answer difficult causal questions about voter turnout and political representation. Many scholars have previously attempted to answer these questions with cross-sectional correlational analyses. To be blunt, the causes and consequences of voter turnout cannot be assessed through correlational analyses – even sophisticated analyses with many "control" variables. Trustworthy estimates of these causal effects can only arise from randomized experiments or something similar. Typically, these difficult questions do not lend themselves to experimentation because the causal factors of interest are difficult to manipulate (although Study 4 does present a one-of-a-kind field experiment), so this dissertation relies primarily on the identification and analysis of quasi-experiments – naturally occurring processes that closely mimic the ideal randomized experiments I would like to run. Unfortunately, there is no magic bullet or fancy statistical method that can uncover causal effects from observational data. However, progress can be made on these difficult questions with careful experimentation, shoe leather, and the identification of quasi-experimental opportunities.

No individual study in this dissertation can purport to offer the final answer on that particular topic or question. The political world is endlessly complex, and there will always be more progress to be made on each of the specific questions addressed in the study. Even if the results presented are internally valid (a lofty goal in and of itself), there is no guarantee of applicability in other democratic settings. For these reasons, I will continue to study these topics and encourage others to do so as well. In addition to the substantive contributions of these papers, I hope that this dissertation makes several methodological contributions that lead to greater insights in the future. For example, Study 2 provides a simple method that can be used to compare the partisan preferences of regular voters and marginal voters – those whose turnout decisions are sensitive to an exogenous factor. This study applies the test in three different settings, but hopefully I and others will apply the test in many other settings where there exists an exogenous shock to voter turnout. Similarly, Study 5 provides a simple test to determine whether an experimental intervention (or quasi-experimental shock) increases or decreases the participation gap – the disparities between voters and nonvoters. I have applied this test to 26 previous GOTV interventions, but I hope that others will employ this test to develop and test new mobilization methods and policy intervention that may reduce the disparities between voters and nonvoters.

The results of this dissertation can be summarized as follows: low and unequal voter turnout is a serious problem, and it is very difficult to fix! Studies 1 and 2 show that low and unequal participation has significant partisan and policy consequences. Election results and public policies in advanced democracies with voluntary turnout are significantly different from the counterfactual world where everyone turns out to vote. Exogenous factors that influence the composition of the electorate like compulsory voting laws, election-day weather, or the coincidence of lower level elections with a presidential race can significantly change electoral and policy outcomes. Even modest electoral reforms that minimally expand or contract the electorate could have significant

3

consequences because the preferences of marginal voters – those on the cusp of voting or abstaining – are often dramatically different from those of regular voters – those who are sure to turn out. Low and unequal turnout is a serious problem for democracy because election results fail to reflect the preferences of the citizenry.

Given the extent of "democracy's unresolved dilemma," we would like to develop mobilization strategies and public policies that increase turnout and improve the representativeness of the electorate. Studies 3 and 4 show that some of the most commonly proposed cures for low and unequal turnout – when subjected to rigorous testing – fail to produce the hypothesized effects. Social capital, a concept invented by social scientists with the goal of improving the performance of democracy, may actually decrease voter turnout by distracting citizens from the political process, exposing them to conflicting views, and providing alternative routes to civic fulfillment. Matthew Atkinson and I find that patron saint festivals in Mexico – precisely the kind of activities thought to raise social capital and the functioning of democracy – significantly decrease voter turnout when they are held within several weeks of an election. Similarly, electoral competition is often offered as a panacea for the ills of democracy – including low and unequal turnout. The logic is that close elections increase the probability of casting a pivotal vote, increasing the individual incentives to turn out. However, as Tom Schwartz says, close elections only increase the incentives to vote in the same sense that being tall increases the chances of bumping one's head on the moon. After a battery of tests – most notably a field experiment in the aftermath of a tied election – Ryan Enos and I find that increased electoral competition (or the increased perception of electoral competition) does not significantly increase turnout and it may actually decreases the representativeness of the electorate.

Low turnout has important consequences and it is harder than expected to increase participation. Worse yet, higher voter turnout does not guarantee that the electorate will become

more representative of the eligible population. In fact, if a turnout increase is concentrated among rich, educated, white churchgoers, then the electorate will become *less* representative of the greater population. With this concern in mind, Ryan Enos, Lynn Vavreck, and I re-analyzed the most effective collective effort to increase voter turnout in recent history, get-out-the-vote (GOTV) mobilization. Barring extreme legal changes like compulsory voting or consolidating all elections to line up with the presidential cycle, GOTV measures such as direct mail, door-to-door canvassing, and volunteer phone calls are the best-demonstrated methods for mobilizing citizens and bringing them to the polls. Given the widespread application of these methods by campaigns and progressive groups, we wanted to test for their effects on the composition of the electorate. Despite significantly increasing mean levels of turnout, GOTV mobilization significantly increases the participation gap – the extent to which voters differ from the greater population. Even though low turnout is hard enough to combat, unequal turnout appears to be the more significant challenge.

There is an apparent contradiction or inconsistency between Studies 2 and 5 that is worth some discussion. Study 2 shows that several small exogenous increases in voter turnout significantly benefit the Democratic Party relative to the Republican Party. Marginal voters – with respect to weather and election timing – are more Democratic than regular voters. Presumably, they are also poorer, less educated, and more ethnically diverse, although I don't have the data necessary to confirm that. Study 5 shows that the types of citizens mobilized by GOTV efforts tend to be demographically similar to the population of citizens already voting, thereby exacerbating disparities in turnout. How can marginal voters be demographically similar to the voting population and also have significantly different partisan preferences than the regular voters? I can offer several potential explanations but only minimal evidence in favor of one or another:

(1) Perhaps GOTV mobilizes a completely different population than the marginal voters in Study 2. It would be nice to apply the test from Study 2 to the data from Study 5 and vice versa to

see whether the results differ significantly. However, I do not have vote choice data in the case of GOTV experiments (most of them were all done at the individual level, so vote choice is not publicly available) and I do not have individual-level demographic data for the analysis from Study 2, but it could be constructed from a large panel of voter files (old voter files are not usually kept around for more than a few years, so this would be a tall order). The differing nature of GOTV mobilization and weather, for example, could easily lead to significantly different marginal populations. Perhaps poor citizens are more easily demobilized by bad weather than rich citizens, but the rich are more easily mobilized by GOTV interventions than the poor.

(2) Demographics are not the same thing as vote choice. Perhaps marginal voters are "high propensity" in terms of demographics, but more Democratic than regular voters. This explanation can be partially addressed with the GOTV experiments held in states with party registration. One could test whether GOTV mobilization is more likely to mobilize Democrats than Republicans. By my own preliminary analysis, I see little evidence that Democrats are easier to mobilize than Republicans, but party registrants are much more easily mobilized compared to independents. This result suggests that these different factors influencing turnout are tapping into different marginal populations, giving priority to the first explanation. As an aside, this result also suggests that GOTV mobilization increases the polarization of the electorate (a possibility briefly mentioned in the Study 5).

(3) The tests in these two studies are different from one another, and even if the marginal voting populations were the same and demographics lined up perfectly with vote choice, we might still see these results. Imagine a world consisting of only two equally-sized groups: H's and L's. H's turn out at a rate of 80% while L's turn out at a rate of 20%. Therefore, 80% of regular voters are H's. Imagine an intervention that increased turnout among H's by 10 percentage points and among L's by 5 percentage points. In this scenario, the marginal voter population would have a lower

proportion of H's compared to the regular voter population (two-thirds vs. four-fifths) bringing the election result slightly closer in proportion to the counterfactual world of universal turnout (78% H's in the electorate instead of 80%). However, the intervention was also more effective among the already over-represented population (so the interaction term in Study 5 would be positive), bringing the absolute vote difference between H's and L's further from the counterfactual world of universal turnout (with a population of 100, the experiment caused the R's to win the election by 65 votes instead of 60). This subtle distinction may explain the seemingly contradictory findings arising from the two separate tests. One way to think about this distinction is to say that Study 2 compares regular voters to marginal voters while Study 5 compares marginal voters to non-voters. In this example (and probably in a lot of real-world settings as well), the regular voters, marginal voters, and non-voters can be ordered on a continuum with marginal voters falling somewhere between the regular voters and non-voters in their political preferences and demographic characteristics. To be honest, this distinction could be made more clearly in Study 5, but the paper is already too long and technical for most readers.

Inequality in political participation has significant electoral and policy consequences, many proposed cures for low turnout fail to produce the desired effects, and even when turnout is increased significantly, there is no guarantee that the electorate will become more representative or produce more equal political outcomes. These findings paint a grim picture and a serious challenge for democratic societies. However, many frontiers of participation research are yet to be explored. Before doctors can treat disease, they must first diagnose the problem. This dissertation makes more progress on the front of diagnosis and leaves much of the treatment to future work. However, the development of the diagnostic tests in this dissertation will hopefully provide a toolkit for future researchers to develop and evaluate new treatments.

One obvious treatment for low and unequal turnout – identified and evaluated in Study 1 – is a strong financial incentive. Australia and other nations implement this by fining nonvoters without a valid excuse, but governments could just as easily accomplish the same thing by providing tax subsidies or direct cash payments to voters. Citizens in many advanced democracies may detest the notion of compulsory or government-incentivized voting, but compulsory voting does not impinge on personal freedom any more than compulsory schooling, compulsory tax paying, compulsory jury duty, and compulsion to drive under the speed limit – just to name a few. Therefore, one promising line of future research may explore the conditions under which government-incentivized voting would be supported by the public and elected officials. As a starting point, Study 1 provides some discussion of the motivations behind the passage of compulsory voting in Australia. A related line of research may explore the social costs of such a policy. In addition to forcing citizens to do something that they would otherwise prefer to avoid and also transferring resources from nonvoters to voters, compulsory could also increase the volatility of elections or the susceptibility of the electorate to superficial shortcuts. However, given the current extent of these problems in voluntary voting systems, compulsory voting laws may actually have limited adverse consequences. Careful future research may help democratic societies to weigh the costs and benefits of such a policy.

Assuming that incentivized voting would not be palatable for most advanced democracies, the findings of this dissertation pose a serious challenge to practitioners, policy makers, and academic researchers. After many decades of research on voting and political participation, we have not identified or developed reliable methods for mobilizing under-represented citizens and bringing them to the polls. This body of findings – that low and unequal turnout has significant consequences, and many popular solutions fail to produce the desired effects – may provide the impetus for the development and implementation of more drastic solutions and reforms.

Study #1

Electoral and Policy Consequences of Voter Turnout:

Evidence from Compulsory Voting in Australia

Abstract

Despite extensive research on voting, there is little evidence connecting turnout to tangible outcomes. Would election results and public policy be different if everyone voted? The adoption of compulsory voting in Australia provides a rare opportunity to address this question. First, I collect two novel data sources to assess the extent of turnout inequality in Australia before compulsory voting. Overwhelmingly, wealthy citizens voted more than their working class counterparts. Next, exploiting the differential adoption of compulsory voting across states, I find that the policy increased voter turnout by 24 percentage points which in turn increased the vote shares and seat shares of the Labor Party by 7 to 10 percentage points. Finally, comparing across OECD countries, I find that Australia's adoption of compulsory voting significantly increased turnout and pension spending at the national level. Results suggest that democracies with voluntary voting do not represent the preferences of all citizens. Instead, increased voter turnout can dramatically alter election outcomes and resulting public policies.

In most democracies, voting is voluntary. Citizens can expend a small cost in exchange for a small chance that they will change the outcome of an election. Voters are those who care deeply about politics, have a lot at stake, possess disposable time and resources, or receive utility from the act of voting itself (Riker and Ordeshook 1968). As a result, voters may be unrepresentative of the population as a whole. Typically, they are wealthier and more educated than non-voters which could bias public policy in favor of the few. In this paper, I turn to early 20th century Australia to address a fundamental question of democracy: Would election results and public policy be different if everyone voted?

Previous evidence on this question is conflicted and suffers from methodological limitations. Correlational studies show that the association between turnout and partisan election results can be large (McAllister and Mughan 1986; Nagel 1988; Radcliff 1994; Pacel and Radcliff 1995; Fisher 2007; Hill 2010), small (Nagel and McNulty 1996; Grofman, Owen, and Collet 1999; Martinez and Hill 2007), or non-existent (DeNardo 1980; Erikson 1995). Further studies of survey data conclude that higher turnout may benefit left-wing parties significantly (Herron 1998; Mackerras and McAllister 1999), minimally (Citrin, Schickler, and Sides 2003; Martinez and Gill 2005; Bernhagen and Marsh 2007; Pettersen and Rose 2007), or not at all (Highton and Wolfinger 2001; Rubenson et al. 2007). Finally, studies of small exogenous shocks to turnout suggest that the effects of turnout on election results and public policy can be large (Anzia 2011, 2012; Fowler 2012; Hansford and Gomez 2010) or small (Berry and Gersen 2011; Knack and White 1998; Sled 2008; Stein 1998; van der Eijk and van Egmond 2007).

As a result of methodological limitations, no previous study adequately addresses the primary question of interest. Confounding variables, reverse causation, and model misspecification may bias the correlational and survey studies. The most compelling causal evidence on the effects of turnout comes from studies of small shocks to turnout. These studies address a separate question

10

about the effects of marginal changes to voter turnout, but they do not assess to the effects of near-universal turnout. If marginal voters are unrepresentative of the entire population of non-voters, then these studies do not speak to the counterfactual question at hand. To determine what would happen if everyone voted, we need a policy change that closely mimics the ideal counterfactual – one where almost everyone is brought to the polls.

This paper analyzes the adoption of compulsory voting laws in Australia as a unique opportunity to assess the effects of near-universal turnout. First, I present a brief history of compulsory voting laws in Australia and explain why this policy change provides such a unique opportunity. Next, I analyze two novel data sources to determine which types of citizens voted in Australia before the adoption of compulsory voting. Overwhelmingly, wealthy and property-owning citizens turned out at higher rates leading to the hypothesis that compulsory voting benefits the Labor Party and progressive policies. Then, I exploit the differential adoption of compulsory voting laws across Australian state assembly elections. A difference-in-difference analysis shows that compulsory voting caused a 24 percentage point increase in voter turnout and a 7 to 10 percentage point increase in the vote shares and seat shares of the Labor Party. Finally, I exploit the adoption of compulsory voting at the national level to test for public policy effects. A synthetic control analysis comparing Australia to other comparable OECD nations demonstrates that the national adoption of compulsory voting caused significant increases in voter turnout and pension spending. I conclude by discussing the implications of these results for modern democracies with voluntary voting.

Compulsory Voting in Australia

When it comes to democracy, Australia is an innovator. Australia was one of the first nations to establish universal suffrage, instant-runoff voting, and the secret ballot. Keeping with the

tradition of a fair, expansive democracy, compulsory voting was first advocated by Alfred Deakin at the turn of the 20th century (Evans 2006). Deakin was the second Prime Minister of the Commonwealth and a member of the Protectionist Party. The political system in Australia approximated a two-party system with the Labor Party on the left pitted against several coalition parties on the right. The primary dimension of conflict between the parties was economic, with working class voters and the Labor Party supporting larger government and more progressive policies relative to upper class voters and the coalition parties (Hirst 2002).

Australia's first system of compulsory voting was implemented in the state of Queensland in 1914. The Liberal Party government led by Digby Denham implemented the policy on the belief that the opposing Labor Party was better at "getting out the vote" (Evans 2006). The other Australian states eventually followed suit, implementing compulsory voting for state elections at different times. Victoria was next in 1926, followed by New South Wales and Tasmania in 1928, Western Australia in 1936, and South Australia in 1941. For national elections, compulsory voting was implemented in 1924. Except in Queensland where the Labor Party initially opposed the policy, compulsory voting received unanimous support from all parties at the national level and in each state assembly. Table 1.1 presents the timing of the adoption of compulsory voting for national elections and for each state. This policy was adopted under the control of multiple different parties and with broad roll call support from all the parties.

Compulsory voting worked in the following way. Every voter was expected to show up to the polling place and cast a ballot on Election Day. They were not obligated to support any of the candidates or parties; they could cast a blank ballot if they preferred. Any citizen who did not vote and could not provide a valid excuse would have to pay a fine. For national elections, the fine was two pounds, and traveling or physical illness were explicitly listed as valid excuses. For state

Table 1.1 The Adoption of Compulsory Voting across Federal and State Elections

State	CV Adopted	First Election w/ CV	Roll Call	Controlling Party
Queensland	1914	1915	47/72	Liberal
Federal Elections	1924	1925	75/75	Nationalist/Country
Victoria	1926	1927	65/65	Labor
New South Wales	1928	1930	90/90	Nationalist/Country
Tasmania	1928	1931	30/30	Nationalist
Western Australia	1936	1939	50/50	Labor
South Australia	1941	1944	39/39	Liberal/Country

assembly elections, the fines varied but were never severe. According to anecdotal reports, the government was lenient with excuses and only a fraction of non-voters were asked to pay the fine.

A historical reading of this period reveals four primary reasons that compulsory voting was implemented. First, it was believed that compulsory voting was a natural extension of compulsory registration, which was already in place (Hirst 2002). If everyone voted, it would be easier to administer the election and detect fraud. Second, the fines levied on non-voters could help to defray the cost of administering the election (Mackerras and McAllister 1999). Third, compulsory voting was seen as the only way to ensure a fair election result (Hirst 2002; Evans 2006). During World War I, the issue of conscription was contentious and divisive. Parliament did not want to make a controversial decision on its own, so it proposed a public referendum on the issue. However, since the issue would potentially affect every Australian family, the Parliament felt that the referendum would only be legitimate if everyone voted. The debate surrounding the referendum planted the seed for compulsory voting in all federal elections (Evans 2006).

The final impetus for compulsory voting involved partisan interests. The Labor Party on the left and the coalition parties on the right hoped to reduce a perceived advantage of their opponents (Mackerras and McAllister 1999; Evans 2006). Non-labor voters were more likely to have cars so that they could drive themselves and others to the polling place. On the other hand, the Labor Party had a larger supply of campaign workers and therefore a more extensive effort to bring Labor supporters to the polls. In order to reduce this wasteful competition, both sides felt that compulsory voting would make them better off.

All four reasons are somewhat practical. Compulsory voting would aid in the administration of elections, raise money, increase legitimacy, and reduce costly campaign efforts of the parties. One party did not push for compulsory voting at the expense of another. All parties supported the policy, and to the extent that the parties made a strategic calculation, they all thought that

compulsory voting would benefit them. Therefore, it appears that the adoption of compulsory

voting arose more for practical or principled reasons than it did for strategic reasons.

While the adoption of compulsory voting was not random, the timing of the policy does not

appear to be related to changes in political, economic, or demographic factors – a claim that I

explicitly test later in the paper. Therefore, compulsory voting provides a rare opportunity to test

for the effects of near-universal turnout on election outcomes and public policy. To assess the

effects of compulsory voting on partisan election results, I exploit the differential timing of

compulsory voting laws across Australian states with a simple difference-in-differences design.

Then, in order to assess the policy consequences of compulsory voting, I employ synthetic control

methods to compare changes in Australia's pension spending over time to changes in other

comparable nations. Both of these designs require a parallel trends assumption which is extremely

plausible, justified with data, and significantly weaker than the assumptions required for previous

studies of the effects of near-universal turnout.[1]

Who Voted in Australia before Compulsory Voting?

Education and class biases in turnout have been well documented in advanced democracies

(Gosnell 1927; Leighley and Nagler 1992; Linder 1994; Powell 1986; Tingsten 1937; Topf 1995;

Verba and Nie 1972; Verba, Schlozman, and Brady 1995). In general, wealthier and more educated

citizens vote at higher rates than those of lower socioeconomic status. This effect is greatest for

[1] For example, previous studies of the effects of near-universal turnout (Herron 1998;

MacKerras and McAllister 1999; Highton and Wolfinger 2001; Citrin, Schickler, and Sides 2003;

Martinez and Gill 2005; Bernhagen and Marsh 2007; Pettersen and Rose 2007; Rubenson et al. 2007)

require the stronger assumption that observational differences in voter turnout across individuals are

exogenous.

those nations with low rates of turnout such as the United States and Switzerland but still present in nations with high turnout such as Canada and the United Kingdom. Even in modern day Australia, which achieves 95 percent turnout through compulsory voting, poorer and less educated citizens are more likely to abstain (McAllister 1986). However, due to the lack of data, class biases in turnout have not been previously studied in Australia before the adoption of compulsory voting. Australia does not typically report the names of individuals who turned out to vote, and no survey data was collected on voting before compulsory voting was established.

To my knowledge, only two sources provide data on the types of individuals who turned out in Australia during this period. First, in 1877, the State of Victoria reported voter turnout separately for property-owners and non-property-owners for that year's state assembly election.[2] Second, following Victoria's 1899 referendum on Australian federation, the state honored the event by publishing the names of all who had turned out. Analyzing both sources, I uncover the extent of turnout inequality in Australia before compulsory voting. As in many democracies today, wealthy, property-owning citizens were much more likely to vote than poor, non-property-owning citizens. This analysis does not suggest that property ownership or property values exhibit a causal effect on turnout. Similarly, these factors are not necessarily the most important predictors of voting. The analysis is descriptive, uncovering as best as possible who voted and who abstained from voting before the adoption of compulsory voting.

Turnout by Property Ownership in Victoria's 1877 Election

After his election to the Victoria State Assembly in 1877, Robert Clark, a working-class miner, requested a report of voter turnout by property-ownership. To my knowledge, this is the only Australian electoral return which provides any breakdown of turnout by demographic

---

[2] *Votes and Proceedings of the Legislative Assembly*, 1877-78, Vol. 1, p.767.

characteristics. The report presents turnout data for 37 of Victoria's 55 districts. These 37 districts contained 580,000 residents and 170,000 eligible voters. For most of the missing districts, the election for legislative assembly was uncontested, so no votes were cast.

At that time in Victoria, property-owning males were automatically registered to vote. If a non-property-owner wanted to register, he would have to pay 1 shilling. The property requirement was not severe; 59 percent of voting-age males were automatically registered as property-owners. This group included farmers, masons, shepherds, storekeepers, butchers, and "gentleman." Non-property owners included strictly working-class citizens such as laborers, servants, cooks, and gardeners. The 1877 report presents for each district the approximate number of eligible voters, the number of registered property-owners, the number of registered non-property-owners, the number of voters who were property-owners, and the number of voters who were non-property-owners. From this data, we can back out the proportions of eligible property-owners and eligible non-property-owners that turned out in the state election. Alarmingly, only 18 percent of eligible non-property-owners turned out to vote compared to 66 percent of property-owners. Given the burden of registration, only 32 percent of eligible non-property-owners bothered to pay the shilling and get on the roll. Even conditional on registering, only 57 percent of those individuals turned out. Surprisingly, property-owners were more likely to vote than even the subset of non-property-owners who had paid to become registered. Put another way, property-owners comprised 84 percent of the electorate even though they only comprised 59 percent of the eligible voters.

Turnout by Property Ownership and Property Value in the 1899 Referendum

In 1899, the Australian states held a referendum on forming a federation. The measure was sure to pass in Victoria, but the state government wanted a high turnout to ensure legitimacy. To incentivize voters, the government offered a certificate to all those who turned out and recorded

their names in a commemorative book. This book, held in the Parliamentary Library of Victoria, contains the names, occupations, and locations of the 163,783 men[3] who turned out from the State of Victoria. By this time the registration fee was removed and registration was mandatory, so that all white men should have been registered and eligible to vote in the referendum.

Despite this exciting data source, there is no comparable list of all eligible voters from which we could assess their propensity to turn out in the referendum. However, the city of Bendigo has preserved some historical rolls of all registered voters, and they fortunately have their list from the exact same year, 1899. Bendigo's rolls list the names, occupations, addresses, and property values of all registered voters. For each individual, the roll indicates the value of his residence and whether he is the owner or just an occupier of that residence. The 1899 Bendigo roll contains the names of 3,510 men. For each individual, I have digitized the data and manually searched for that person in the book of referendum voters. 61 percent of Bendigo's registered voters are confirmed as having voted in the referendum, consistent with historical accounts of the turnout rate in Bendigo (Maslunka 1983).

Just as in 1877, we can assess whether property owners in Bendigo were more likely to vote than occupiers. Also, we can assess whether property values, a good proxy for wealth, are related to turnout. Figure 1.1 presents kernel regressions of turnout across property values separately for owners and occupiers. Consistent with expectation, owners were 10 percentage points more likely to vote than occupiers. Also, property values are highly correlated with turnout for both owners and occupiers. A single standard deviation increase in property value is associated with a 7 percentage

---

[3] Women, unfortunately, did not yet have the right to vote in Victoria.

Figure 1.1. Turnout Inequality in Bendigo, 1899
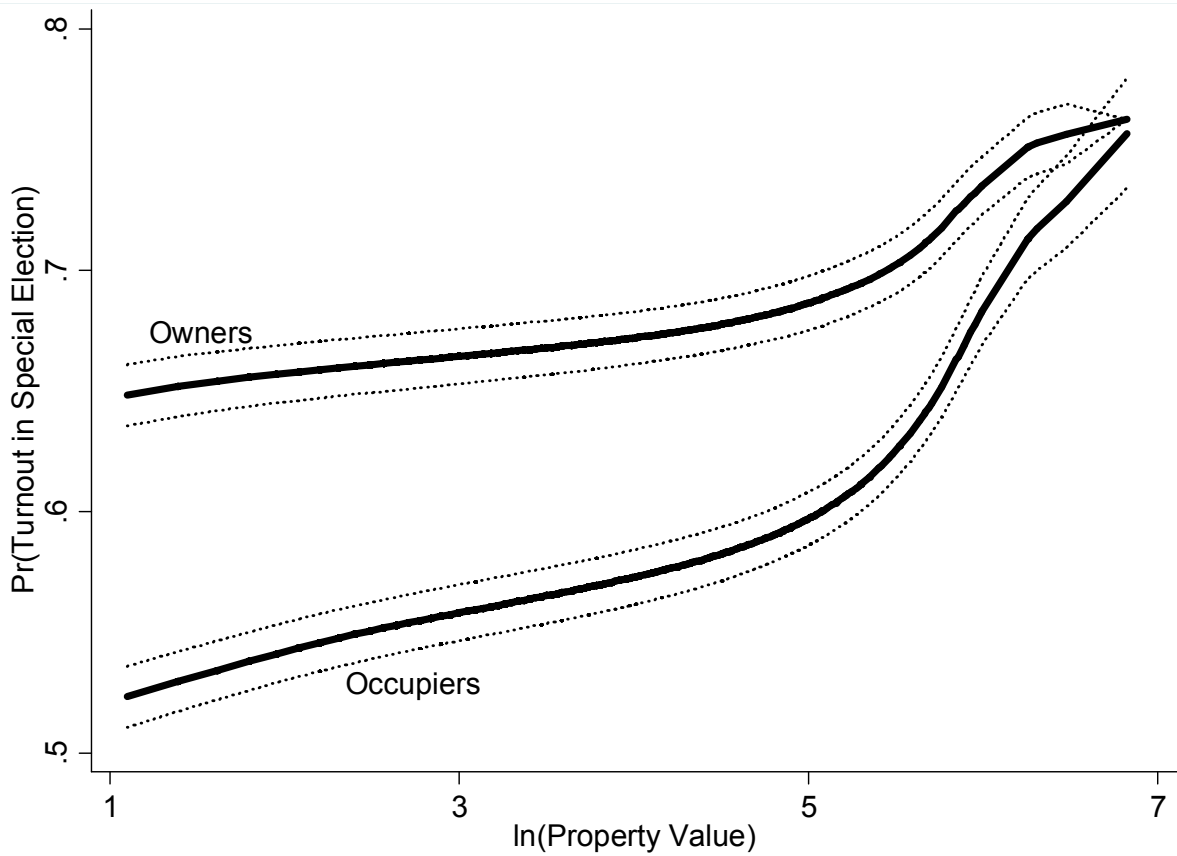
The solid curves represent kernel regressions indicating the probability that an individual citizen turned out to vote in the 1899 special election as a function of the value of his residence. Separate kernel regressions are shown for property-owners and occupiers (non-owners). The dotted lines indicate standard errors. The kernel regression use the epanechnikov distribution with a bandwidth of 1.5 and a p-width of .57, but the general finding is robust to many different specifications.

point increase in an individual's probability of voting for owners and a 10 percentage point increase in an individual's probability of voting for occupiers.[4]

These findings from 1877 and 1899 demonstrate a large degree of turnout inequality in Australia before the adoption of compulsory voting. Due to the lack of available data, these findings cannot be replicated outside of Victoria or Bendigo. However, we have no reason to suspect that turnout inequality was unique to these regions of Australia. Contemporary electoral reports indicate that turnout and registration rates in Bendigo were similar to other municipalities in Victoria and Victoria was comparable to other states in Australia. This analysis constitutes, to my knowledge, the first and only possible analysis of individual level voting behavior in Australia before compulsory voting. Acknowledging the limitations, I take these results as evidence of a larger phenomenon across Australia.

As in many democracies today, Australia's wealthy citizens voted at much higher rates than working class citizens under a system of voluntary voting. For this reason, any lessons drawn from Australia may be applicable to other democratic nations today. Also, since working-class citizens in Australia systematically preferred the Labor Party, these analyses suggest that that compulsory voting and increased voter turnout would benefit the Labor Party and lead to more progressive policies in Australia. The next two sections explicitly test these hypotheses.

The Effects of Compulsory Voting in State Assembly Elections

In order to test the effect of compulsory voting on election results, I have collected the results of every state legislative assembly election from 1910 to 1950.[5] The six Australian states

---

[4] These numbers are estimated through separate OLS regressions of turnout on property value for owners and occupiers.

provide an excellent opportunity to test for the effects of near-universal turnout because they adopted compulsory voting at different times. Employing a differences-in-differences design, I estimate the effect of compulsory voting on three different dependent variables: voter turnout, Labor Party vote share, and Labor Party seat share. As compulsory voting is implemented in one state, we can compare that state's changes in voting behavior to the changes of other states at the same time. This design – in its simplest form – requires a "parallel trends assumption" that voting behavior would, on average, have parallel trends across states in the absence of any changes in compulsory voting laws. Later in the paper, this assumption is discussed in detail, justified with statistical tests, and relaxed by allowing for state-specific trends.

Voter turnout and the electoral success of the Labor Party, measured in two different ways, are the primary outcomes of interest. During this time period, Australia had multiple parties competing for office. However, several conservative parties would typically form a coalition against the Labor Party, so the environment approximates a two-party system. The clear differences between the Labor Party and the coalition parties largely revolved around economic issues; Labor preferred government intervention to protect workers, and coalition leaders preferred free markets. By analyzing the Labor Party's success, we capture the bulk of political competition occurring during this time period. Also, given the economic divide between the Labor and coalition parties, we have a clear prediction about the effect of increased turnout on the Labor Party's success.

---

[5] Data was collected from the Australian Politics and Election Database, hosted by the University of Western Australia (http://elections.uwa.edu.au/electionsearch.lasso). 1910 is the first year for which I can obtain reliable electoral data for every state. Subsequent results are not sensitive to the choice of included years.

Before turning to the explicit econometric tests, I present the raw data for the analysis in Figure 1.2. The top panel of the figure presents the level of voter turnout in every election and the bottom panel presents the Labor Party's vote share. Hollow triangles represent elections before compulsory voting, and solid triangles represent elections after the adoption of compulsory voting in that particular state. The naked eye can readily detect the significant effect of compulsory voting on voter turnout. In each state, turnout jumped dramatically when compulsory voting laws were implemented. However, the effect of compulsory voting on Labor Party vote share is less apparent, because there is so much natural variation in this variable from year to year. (This variation may explain why conservative parties in Australia continued to support compulsory voting laws even though the tests in this study reveal that they suffered from the policy.)

To quantify the effects of compulsory voting, I estimate the following equation using OLS separately for each dependent variable (turnout, Labor vote shares, and Labor seat shares),:

(1) $DV = \alpha(Compulsory\ Voting)_{it} + \gamma_i + \delta_t + \varepsilon_{it}.$

This model includes a dummy variable for compulsory voting, state fixed effects (γ_i), and year fixed effects (δ_t), allowing each state and each year to have their own idiosyncratic effect on the dependent variables. The coefficient α indicates the effect of compulsory voting on the dependent variable. As a robustness test, I also include linear time trends for each state, parametrically relaxing the parallel trends assumptions by allowing each state to have a unique trend over time.

Table 1.2 presents the results of all three difference-in-difference regressions along with the three corresponding models with state-specific trends. Model 1 shows that the implementation of compulsory voting increased turnout by 24 percentage points on average, from 67 to 91 percent. Models 3 and 5 show that compulsory voting increased the Labor Party's vote-share and seat-share by 9 and 7 percentage points, respectively. These results are both substantively and statistically significant. Since the Labor Party's average seat share was 43% before the adoption of compulsory

Figure 1.2. Turnout and Labor Vote Share across State Assembly Elections (1910-1950)
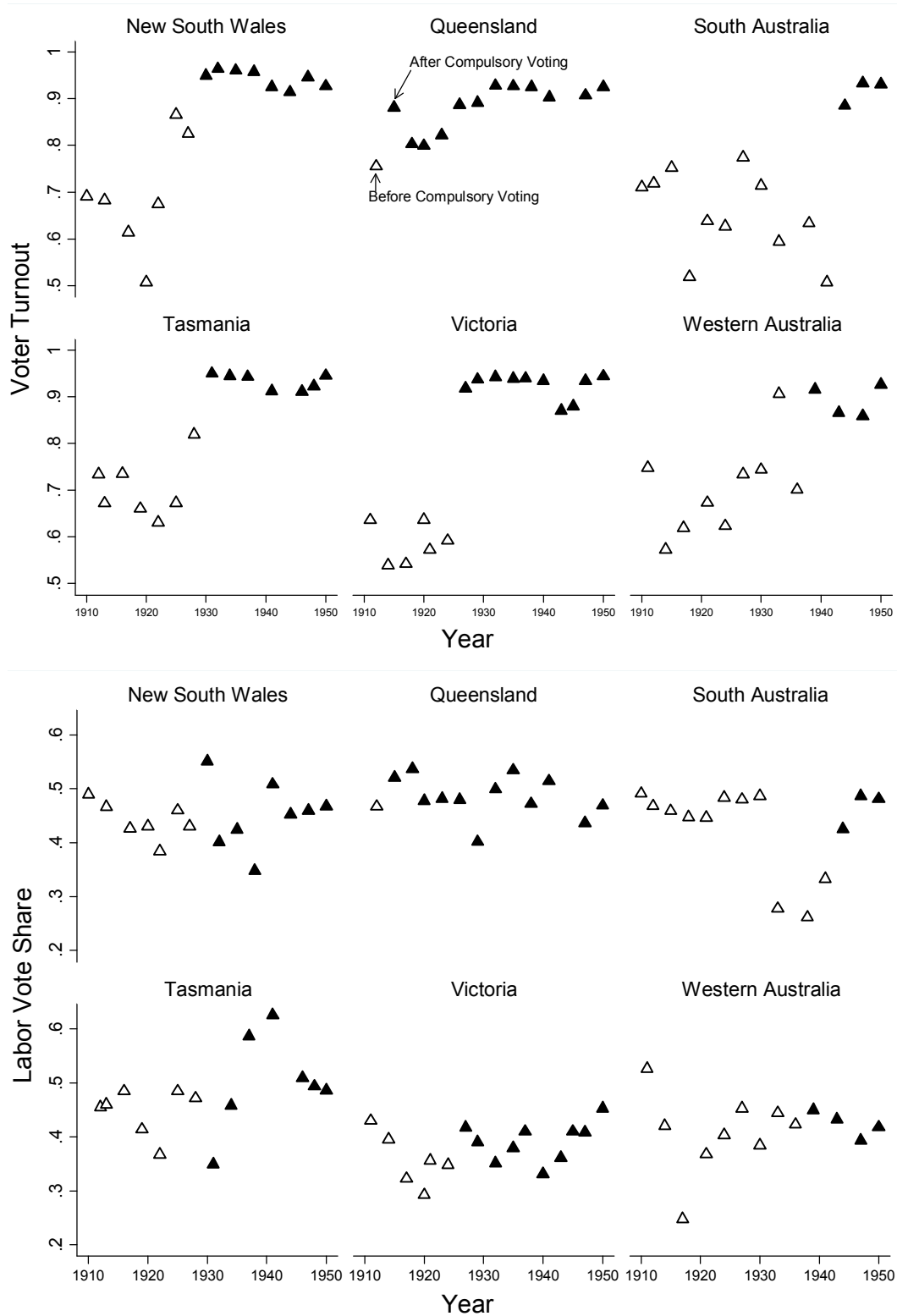
The graph presents the level of voter turnout (top panel) and Labor Party vote share (bottom panel) in every state legislative assembly election between 1910 and 1950. Hollow triangles indicate an election help before the adoption of compulsory voting in that state, and solid triangles indicate an election after the adoption of compulsory voting.

Table 1.2. Effects of Compulsory Voting in State Assembly Elections

	DV = Turnout		Labor Vote Share		Labor Seat Share	
	(1)	(2)	(3)	(4)	(5)	(6)
Compulsory Voting	**.243**	**.245**	**.092**	**.093**	**.069**	**.098**
	(.042)	(.046)	(.033)	(.040)	(.048)	(.065)
State Fixed Effects	X	X	X	X	X	X
Year Fixed Effects	X	X	X	X	X	X
State-Specific Trends		X		X		X
Observations	85	85	85	85	85	85
R-Squared	.925	.933	.759	.790	.660	.721
SER	.056	.052	.049	.045	.110	.098

State-clustered standard errors in parentheses
Each observation is a state assembly election between 1910 and 1950. "Compulsory Voting" is a dummy variable, indicating whether compulsory voting laws were in place. The dependent variables are coded to theoretically range from 0 to 1, so the coefficients in columns 1, 3, and 5 indicate that compulsory voting increased voter turnout, Labor vote share, and Labor seat share by 24, 9, and 7 percentage points, respectively.

voting, these effects were meaningful enough to change the control of power in many legislative assemblies. Models 2, 4, and 6 present the robustness tests allowing for state-specific trends and all three results are virtually unchanged. These additional tests lend credibility to the parallel trends assumption, because relaxing the assumption is inconsequential for the subsequent estimates.

According to these results, average turnout in the absence of compulsory voting during this period would have been 67 percent, and the Labor Party would have received 44 percent of the votes. However, average turnout with compulsory voting would have been 91 percent, and Labor would have received 53 percent of the votes. This suggests that 78 percent of the new (marginal) voters brought to the polls by compulsory voting supported the Labor Party compared to just 44 percent of older (regular) voters.[6] This estimate is substantively large but quite plausible given what we now know about turnout inequality prior to compulsory voting. Among the 91 percent of voters who voted under compulsory voting, only 61 percent of Labor voters turned out to the polls under voluntary voting while 88 percent of non-Labor supporters turned out. These calculations independently confirm the extent of turnout inequality in Australia before compulsory voting. Before compulsory voting, non-Labor supporters were 1.4 times more likely to vote compared to Labor supporters. When compulsory voting caused an exogenous increase in turnout, election results changed dramatically.

Estimating the Standard Errors

While the previous analysis includes 85 different elections, each election does not represent an independent observation. Serial correlation or state-specific clustering could lead traditional approaches to over-estimate the precision of our estimates. Table 1.2 reports state-clustered

---

[6] The 78% figure comes from solving this simple equation for X: .67*.44 + .24*X = .91*.53. See Study #2 for more details on this test.

standard errors which are designed to accommodate these concerns, but these standard errors can be misleading when the number of clusters is small. See Bertrand, Duflo, and Mullainathan (2004) for a more detailed discussion of these issues.

In the Online Appendix, I discuss these issues and demonstrate the statistical significance of these results under numerous different methods for estimating standard errors. For example, the estimated effect of compulsory voting on Labor vote share is statistically significant ($p < .05$) even when calculating the standard errors by non-parametric bootstrap, block bootstrap, or random permutation.

Tests for Confounding Demographic Trends

The parallel trends assumption implies that the timing of the adoption of compulsory voting did not coincide with any state-specific changes in demographic or political factors that could have independently influenced voter turnout or election results. Historical analysis of the adoption of compulsory voting lends credence to this assumption. The insensitivity of the results to the inclusion of state-specific trends provides further support. In this section, I search for further evidence that could potentially falsify or bolster the parallel trends assumption. I find that the adoption of compulsory voting across states was not correlated with any changes in economic or demographic variables.

To conduct these tests, I collected data for each state from the Australian censuses in 1911, 1921, 1933, and 1947.[7] For all six states at each of these four time points, I obtained data on the

[7] One natural robustness test would involve including these time-varying covariates as control variables in the original regressions. However, these tests cannot be conducted because these variables are not available for all election years. As a close substitute, I test whether changes in these covariates coincide with the timing of compulsory voting across states.

state's population and the proportion of the state's population that was under 21, married, born in Australia, identifying with the Church of England, and working in the manufacturing sector. Then, treating each variable as an outcome variable, I regress each variable on a dummy variable for compulsory voting, state fixed effects, and year fixed effects. This procedure mimics the difference-in-difference regressions shown previously. The results of each regression are shown in Table 1.3. For each test, the placebo "effect" of compulsory is statistically and substantively indistinguishable from zero. These results provide further support for the parallel trends assumption and demonstrate that the adoption of compulsory voting across states was not correlated with these demographic or economic changes.

The Effects of Compulsory Voting on Public Policies

The significant effects of compulsory voting on election results suggest that compulsory voting may have influenced public policy as well. The Labor Party held systematically different positions than the other parties over many issues including unions, social spending, immigration policies, and the size of government. Therefore, a greater presence of Labor members in state assemblies and the national Parliament may have had significant effects on the lives of Australians at the time. Moreover, the new wave of working class voters may have led all parties to shift in a progressive direction. A historical account of Australia's national politics around this time period suggests several potential effects of compulsory voting on public policy.

The majority party's platform changed dramatically after the adoption of compulsory voting for federal elections in 1924. In the 1922 election the Nationalist Party maintained power, announcing a conservative domestic policy: "First and foremost, we are against class legislation and class government" (Hughes 1922). In 1925, the first election under compulsory voting, the Nationalist Party maintained power but shifted dramatically on these issues: "It has to be recognized

27

that even under the conditions existing in Australia, the wages of our workers are not sufficient to enable them to safeguard against these evils [sickness, unemployment, and old-age]" (Bruce 1925). In just a three year span, the majority party made a clear turn toward progressive domestic policies, and the introduction of more working class voters into the electorate and increased presence of the Labor Party may have provided impetus for such a change.

Statistical evidence of the effects of compulsory voting comes from data on social spending across various nations over time (Lindert 1994). I focus specifically on pension spending, because this was a key issue of disagreement between the parties, just as there was tension between the left and right in many advanced democracies over pension policies (Baldwin 1990). In Australia, the political parties disagreed strongly regarding the size and structure of the federal pension program, but for various reasons, they did not diverge significantly on other social issues such as education, health, welfare, and housing (Hirst 2002). Because the Labor Party and working class voters systematically preferred a more progressive pension program relative to the coalition parties and upper class voters, I hypothesize that compulsory voting increased pension spending in Australia.

To test this hypothesis, I estimate the effect of compulsory voting at the national level on voter turnout and pension spending by comparing Australia to 20 other developed nations across three different time points: 1910, 1920, and 1930.[8] Again, I rely on a difference-in-differences design. However, in order to maximize the plausibility of the parallel trends assumption, I employ

---

[8] Turnout and pension data are from Lindert (1994) with corrections made to Australia's turnout data. This cross-country design could also be employed to assess the impact of compulsory voting on the electoral success of left-wing parties, but the state assembly design is much cleaner for that purpose and it avoids the challenges of coding left-wing parties across countries and comparing across different electoral systems.

synthetic control methods (Abadie, Diamond, and Hainmueller 2010) to construct synthetic control units that are comparable to Australia before the adoption of compulsory voting.

For each test, the synthetic control unit is the weighted average of other nations that best mirrors Australia in terms of voter turnout or pension spending in 1910 and 1920. New Zealand, France, Canada, and the United Kingdom receive the greatest weights in the turnout analysis, because they closely mirror Australia's level and trend in turnout before compulsory voting, while Denmark and New Zealand receive the greatest weights in the pension analysis. The incorporation of other economic, demographic, and political variables in the weighting algorithm does not change the subsequent results. Moreover, simpler difference-in-difference designs which weigh all comparison units equally yield similar results. More details on these analyses are available in the Online Appendix.

Figure 1.3 presents the trends in voter turnout (top panel) and pension spending (bottom panel) for Australia and the synthetic control units. Their levels of turnout and pension spending are nearly identical in 1900 and 1910. However, after the adoption of compulsory voting in 1924, voter turnout and pension spending increased dramatically relative to the synthetic control units. Difference-in-difference calculations indicate that compulsory voting increased voter turnout by 18.6 percentage points and pension spending by 0.41 percentage points of GDP. In 1930, the pension spending of the control group was just under 1 percent of GDP, so this effect represents a more than 40 percent increase in the number of federal dollars going toward old age pensions – a substantively significant effect.

To assess the statistical significance of these estimates, I conduct placebo tests on the other countries, pretending that a policy change occurred between 1920 and 1930 in each of the comparison units. For each of the 20 countries other than Australia in this analysis, I construct a synthetic control for each country, excluding Australia as a potential control unit. I then calculate

Figure 1.3. The Effects of Compulsory Voting on Turnout and Pension Spending
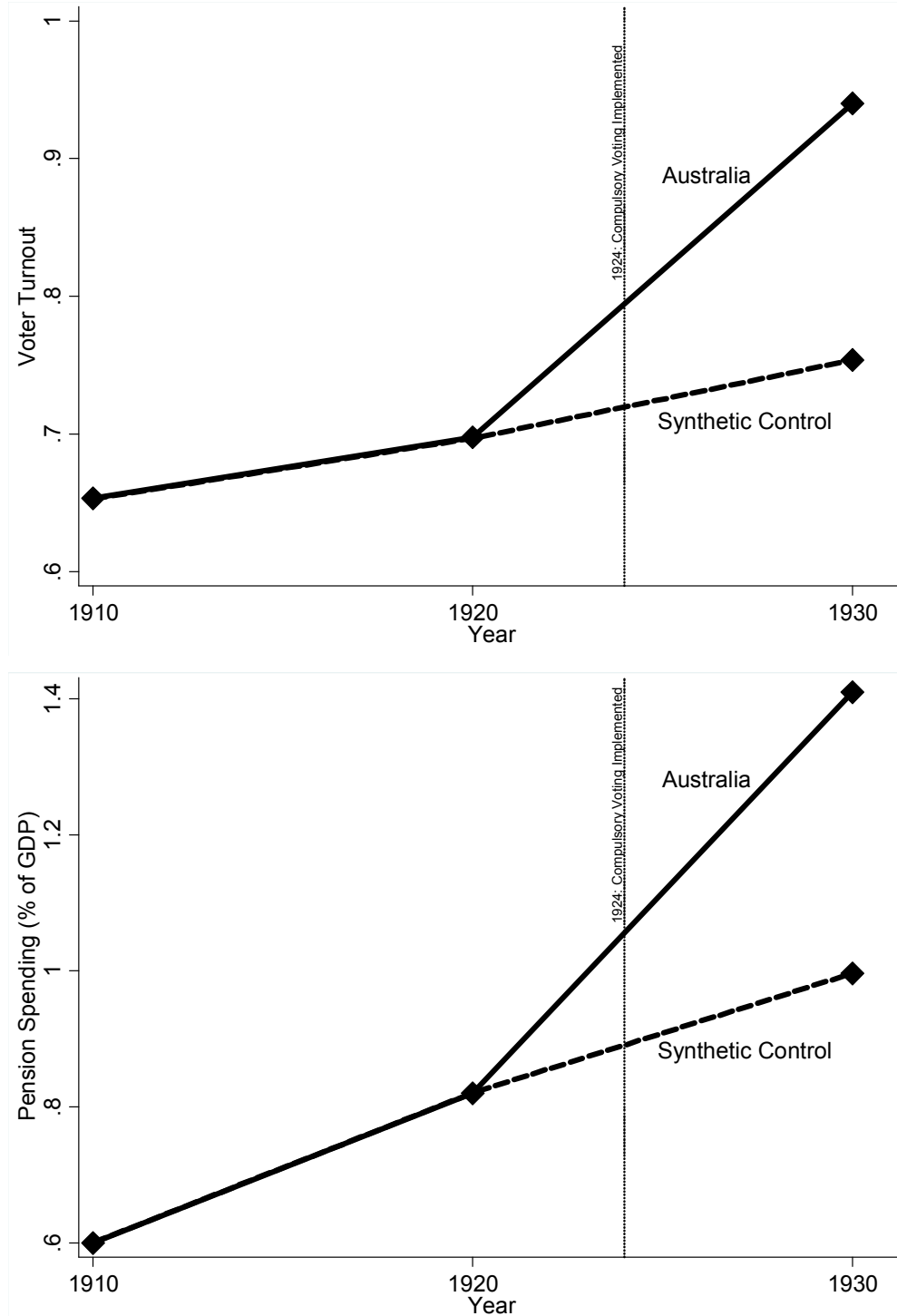
The figure presents the level of voter turnout (top) and pension spending (bottom) in Australia and in a synthetic control nation across three points in time. The synthetic control unit is a weighted average of other OECD countries that closely mirrored Australia before the adoption of compulsory voting for national elections in 1924. Both turnout and pension spending increased dramatically in Australia relative to the synthetic control nation after the adoption of compulsory voting.

the difference-in-difference in pension spending between each country and its synthetic control from 1920 to 1930. The results of these placebo tests are provided in the Online Appendix. For both turnout and pension spending, only one country exhibits a difference-in-difference estimate as large as Australia. The dramatic increases in turnout and pension spending in Australia between 1920 and 1930 are substantively large and much greater than we would expect to see by chance alone. Taken together, the analyses of state assembly elections and federal policy over time provide strong evidence that near-universal turnout can have significant consequences for both partisan election results and public policy.

Discussion of Alternative Explanations

Alternative explanations of this study's findings will likely arise in two forms. First, the timing of the adoption of compulsory voting may have been endogenous to changes in turnout or support for the Labor Party. Specifically, there may have been some reason that states adopted compulsory voting precisely when the Labor Party was expected to increase its electoral success. Second, compulsory voting may have influenced election results through some mechanism other than increased turnout. Perhaps the introduction of compulsory voting caused people to change their partisan attitudes in favor of the Labor Party, and the effect of the policy on elections had nothing to do with increased turnout among the working class. In any observational analysis, alternative explanations cannot be ruled out entirely. However, I have made the best possible effort to raise and assess the plausibility of these explanations through historical evidence, previous research, and my own statistical tests.

Was the timing of compulsory voting truly exogenous? The difference-in-difference analyses in this paper assume that election results and public policies would have followed similar trends across states and countries in the absence of any changes in compulsory voting laws. For example,

the results would be biased if compulsory voting was adopted at a time when the Labor Party was expected to improve its electoral success. Perhaps the Labor Party pushed for compulsory voting as it was gaining momentum. Conversely, the other parties may have pushed for compulsory when they expected to lose support in the next election. However, the adoption of compulsory voting received widespread support from all political parties, so these explanations are unlikely. Moreover, previous election results are uncorrelated with the timing of compulsory voting. As discussed previously, compulsory voting was primarily implemented for practical reasons. After numerous discussions with Australian historians, I am unaware of any confounding variable which could have simultaneously led to compulsory voting and an increase in support for the Labor Party. The decision to adopt compulsory voting appears to be orthogonal to the partisan attitudes of the citizenry and any trends in these attitudes. Also, the state-level results are unchanged if the parallel trends assumption is relaxed by allowing for state-specific trends. Lastly, tests for confounding demographic trends in Table 1.3 reveal no evidence that the adoption of compulsory voting was correlated with demographic or economic changes.

Even if compulsory voting increased the electoral success of the Labor Party, it may have done so through some mechanism other than near-universal turnout. This study of the effects of compulsory voting is less interesting for the study of democracy if the mechanism has to do with something other than increased turnout. Could compulsory voting have dramatically changed election results or public policy independent of its effects on turnout? Perhaps voters were upset with the coalition parties for passing compulsory voting and decided to shift their support to the Labor Party. This seems unlikely since all parties supported the policy change. Perhaps the presence of new voters at the polls changed the political attitudes of old voters causing them to shift toward the Labor Party. However, if anything, we might expect that this phenomenon would work against my findings if upper class Australians became lionized against the Labor Party because of the

new presence of working class voters at the polls. A historical analysis suggests no reason that compulsory voting would cause a shift toward the Labor Party through any mechanism other than increasing turnout.

Changes in voter turnout do not take place in a vacuum. The introduction of new voters to the polls will lead candidates to change their campaign strategies and shift their policy platforms. However, these phenomena do not pose a problem for my estimates because they are all part of the downstream effects of compulsory voting and increased turnout. Presumably, the shift in political platforms is one mechanism for the observed changes in pension spending, over and above the increased electoral success of the Labor Party.

Discussion and Conclusion

Advanced democracies expend incredible resources in the administration of elections in the hope that election results and public policies will closely reflect the preferences of the citizenry. However, a significant proportion of citizens often abstain from voting. Worse yet, those who abstain are systematically different from those who vote, meaning that election results may not accurately reflect the preferences of all citizens. Aware of this problem, political scientists have extensively studied voter turnout, typically assessing the correlates and causes of voting. Despite this collective effort to understand turnout, there is little evidence connecting voting to tangible outcomes. How would elections and public policies change if everyone voted?

Previous attempts to address this question suffer from methodological problems. Correlations between turnout and various outcomes lack a causal interpretation because confounding variables influence both turnout and the outcomes of interest and the outcomes of interest may influence voter turnout. Comparisons of voters and non-voters within an electorate also suffer from the possibility of confounding variables or systematic measurement error in surveys.

Previous quasi-experimental approaches lack external validity because the subset of voters influenced by the quasi-experiment is unrepresentative of all non-voters. In short, previous research has failed to answer the counterfactual question of interest. In a democracy with voluntary voting, researchers have been unable to determine the effects of near-universal turnout.

The analysis presented here largely overcomes the problems of previous studies. The timing of the introduction of compulsory voting in different Australian states appears to be exogenous to partisan attitudes and other political events, so difference-in-difference methods can estimate the causal effects of compulsory voting laws. Moreover, because compulsory voting caused a substantial increase in voter turnout, the subset of citizens influenced by the policy is nearly the entire population of non-voters. This analysis brings us closer than ever before to answering the extreme counterfactual question: "what if everyone voted?"

Before the implementation of compulsory voting, Australia was not dissimilar from many advanced democracies today. Less than 70 percent of citizens voted in federal or state elections, and wealthy citizens were much more likely to vote than working class citizens. Given these patterns, the lessons learned from Australia may be applicable to many democracies today. With the extent of turnout inequality in the United States, Switzerland, Poland, and many other countries, the effects of compulsory voting could be just as great in these contexts.

"Democracy's unresolved dilemma" is that elections do not accurately reflect the preferences of the citizenry (Lijphart 1997). Systematic turnout inequality means that some citizens will be better represented than others. In this study, I exploit a rare opportunity to test the extent of this dilemma. Before the introduction of compulsory voting in Australia, election results and public policy were drastically different from the preferences of the citizens. When near-universal turnout was achieved, elections and policy shifted in favor of the working class citizens who had previously failed to participate. While Australia has largely resolved the problem, inequalities in voter turnout remain in

most advanced democracies. Increased turnout has tangible effects on partisan election results and

public policies, and those effects will benefit the disadvantaged subset of citizens who otherwise

would have abstained from the political process.

Study #2

Regular Voters, Marginal Voters, and the Electoral Effects of Higher Turnout

Abstract

How do marginal voters differ from regular voters? I develop a method for comparing the partisan preferences of regular voters to those marginal voters whose turnout decisions are influenced by exogenous factors and apply it to three sources of variation in turnout – weather and the timing of gubernatorial and congressional elections. In each setting, marginal voters are more supportive of the Democratic Party than regular voters, and the substantive size of this divide can be huge – ranging from 5 to 47 percentage points. The findings suggest that electoral reforms and other factors that may expand or contract the electorate can have important electoral consequences. Moreover, the findings suggest that election results do not always reflect the preferences of the citizenry, because those marginal citizens who stay home have systematically different preferences than those who participate.

Many Americans abstain from the political process. In a typical presidential election, at least 40 percent of eligible citizens will fail to turn out (McDonald and Popkin 2001), and this number can be much higher in other elections. This widespread nonparticipation leads to the concern that significant segments of the population are under-represented or even ignored in the making of public policy (Verba, Schlozman, and Brady 1995; Lijphart 1997). If those who abstain are systematically different from to those who participate, then election results will not reflect the preferences of the population as a whole. This paper offers a new method and several new opportunities to assess the difference in preferences between regular voters, those who always vote, and marginal voters whose decisions to turn out are sensitive to exogenous factors. How much do regular voters disagree from those on the margins in terms of their partisan vote choices? The answer will vary with the setting and the particular subset of marginal voters. However, in all cases analyzed in this paper, marginal voters are systematically more supportive of the Democratic Party (and less supportive of the Republican Party) than regular voters. As a result, fluctuations in the size of the electorate can have dramatic consequences for partisan elections.

A significant scholarly literature asks, "what if everyone voted?" (e.g. Highton and Wolfinger 2001; Citrin, Schickler, and Sides 2003; Martinez and Gill 2005). Unfortunately, this question may be impossible to answer convincingly in the U.S. because modern history has not witnessed anything remotely close to universal turnout. For the same reason, the answer to this question may be irrelevant beyond scholarly curiosity. As a result, scholars have sometimes settled for a more tractable question, "what if more people voted?" (e.g. Knack and White 1998; Stein 1998; Sled 2008; Hansford and Gomez 2010; Hill 2010). However, even this question is difficult to answer. Would election results be different if turnout increased by 1 percentage point? Different studies will yield different answers because it depends on *which 1%* of the population is being mobilized. There is no

single, correct answer to the question. Some increases in turnout will help the Democrats, others will help the Republicans, and others will have no effect.

To understand the effects of increasing (or decreasing) voter turnout, we need many sources of exogenous variation in turnout. Some citizens are "regular voters" who will vote in a particular election regardless of the conditions. Others may be "never voters" who are impossible to mobilize. However, the "marginal voters," those whose turnout decisions can be influenced by exogenous factors, electoral policies, or other idiosyncratic features of an election, are of prime importance for this study because they are the easiest citizens to mobilize or de-mobilize, and their decisions to turn out can influence electoral outcomes. In order to understand the consequences of turnout, we must characterize these marginal voters and understand how their attitudes and preferences differ from the regular voters. No single analysis can answer this question. Instead, we must exploit multiple, independent sources of exogenous variation to understand how marginal voters, on average, differ from regular voters.

This paper makes significant progress in characterizing marginal voters in America by developing a method for directly comparing the partisan preferences of regular and marginal voters. Once a researcher has identified an exogenous source of variation in turnout, she can estimate four quantities directly from the data and solve a simple system of equations to compare the voting behavior of regular and marginal voters. This paper reviews the previous literature on the effects of turnout, describes the empirical strategy for comparing regular and marginal voters, implements the test in three different settings, and then benchmarks the results against observational estimates. In every case, marginal voters are systematically more Democratic than regular voters. These findings suggest that voters are typically unrepresentative of the greater population and changes in voter turnout can have significant consequences for partisan election results.

Previous Evidence on the Effects of Turnout

The most common source of evidence on the effects of voter turnout is correlational. Studies sometimes find a positive association between turnout and Democratic electoral success (Radcliff 1994; Hill 2010), no association (DeNardo 1980; Erikson 1995), or an association that varies across time, region, and electoral competitiveness (Nagel and McNulty 1996; Grofman, Owen, and Collet 1999; Martinez and Hill 2007). Of course, correlational analyses, even carefully done, tell us little about the causal relationship between turnout and election outcomes. For a critique of these analyses (in the form of a humorous parody) see Wuffle and Collet (1997). Because turnout (and changes in turnout) could be influenced by previous election results, demographic characteristics, and other unobserved variables and trends, correlational analyses may significantly underestimate (or overestimate) the true effects of higher turnout.

A second strand of evidence comes from survey data. Many studies compare the characteristics and attitudes of voters and non-voters in order to assess the effects of higher turnout. Again, some argue that higher turnout would dramatically help the Democrats (Herron 1998), modestly help the Democrats (Citrin, Schickler, and Sides 2003; Martinez and Gill 2005), or have little effect (Highton and Wolfinger 2001). There is no debate that low socioeconomic-status citizens are less likely to vote (Verba and Nie 1972; Wolfinger and Rosenstone 1980; Leighley and Nagler 1992; Verba, Schlozman, and Brady 1995) and more likely to support the Democratic Party (Ansolabehere, Rodden, and Snyder 2006; Bartels 2006, 2008; McCarty, Poole, and Rosenthal 2006; Gelman et al. 2007, 2008), but studies disagree about the extent to which these disparities might translate into different election results under higher turnout. As with correlational analyses, survey-based analyses are limited in their ability to characterize the preferences of non-voters. Whether estimates of non-voter preferences are obtained from statistical models or direct survey questions, we cannot know what non-voters would do in the hypothetical scenario where they vote. As

Lijphart argues, "Nonvoters who are asked about their opinions on policy and partisan preferences . . . are typically citizens who have not given these questions much thought . . . It is highly likely that, if they were mobilized to vote, their votes would be quite different from their responses in opinion polls" (1997, p. 4). See Jackman (1999) for a related critique.

Given the difficulties of correlational and survey analyses, the most promising avenue for understanding the consequences of higher turnout is the discovery of "natural experiments" or exogenous changes in turnout. Observational studies are limited because the decision to turn out is highly non-random. However, there may be some voters – marginal voters – whose decisions to turn out are influenced by exogenous factors like weather or electoral law. Several studies exploit these factors to assess the effect of increased turnout on election results. Again, the evidence is mixed. Some studies find that higher turnout will greatly benefit the Democratic Party (Gomez, Hansford, and Krause 2007; Hansford and Gomez 2010), modestly benefit the Democratic Party (Knack and White 1998), or have little effect (Stein 1998; Sled 2008).

What can we make of these mixed results? One explanation is that there is no single, correct answer. The effect of increasing turnout will depend on which citizens are mobilized. While experimental and quasi-experimental studies may be the most promising avenue to answer the question at hand, each individual study will yield a different answer because the subset of marginal voters changes in every setting. Electoral reforms like early voting and vote-by-mail (Berinsky 2005) and get-out-the-vote efforts (Enos, Fowler, and Vavreck 2012) tend to mobilize high-SES citizens more so than low-SES individuals, so studies exploiting these exogenous factors will likely find little effect of higher turnout – or even a Republican benefit. Other exogenous factors may have strong

electoral effects if they happen to identify a different subset of marginal voters.[9] Unfortunately, there is no easy solution to this problem. Unless the federal government randomly induces universal turnout through financial incentives (see Fowler 2013 for an analysis of the strong partisan consequences of compulsory voting laws in Australia), we will never fully characterize the preferences of all non-voters. However, the repeated analysis of different, independent exogenous factors will allow us to characterize the preferences of marginal voters – those on the cusp of voting. Moreover, because universal turnout will never be achieved in the U.S., understanding the partisan preferences of different marginal voter populations is a more relevant goal than answering the more elusive question – "what if everyone voted?"

A Method for Comparing Regular and Marginal Voters

Suppose that we could observe an experiment where voters in some places are randomly mobilized while those in other places are not mobilized. Suppose we could also make the assumption that the treatment won't change the voting choices of individuals who were going to vote anyway.[10] Then, a researcher could calculate the effect of the treatment on both turnout and partisan election results with simple difference-in-means tests. Let us call these quantities P_M and ΔV, respectively.[11] P_M, the turnout effect, indicates the proportion of all eligible voters who are

[9] This is a special case of the well-known methodological phenomenon in instrumental variables analysis where two, different, equally-valid instruments can identify two, different, equally-valid local average treatment effects (Angrist and Pischke 2009).

[10] This assumption is similar to the "exclusion restriction" assumption necessary for instrumental variables analyses.

[11] Again, in an instrumental variables setup, these would be the first stage and reduced form estimates, respectively.

marginal voters – individuals whose turnout decision was influenced by the treatment. ΔV, the vote share effect, indicates the extent to which the treatment changed the election result. While these estimates alone may be interesting, they do not allow us to directly compare the regular and marginal voters. We would like to know how the vote choices of marginal voters differ from those of regular voters, but this quantity cannot be calculated without more information.

Following the experiment, the researcher could also calculate the expected level of voter turnout and Democratic vote share in absence of any treatment. Because the treatment was randomly assigned, these quantities are estimated by the mean turnout and mean Democratic vote share of the control group. These quantities indicate the number of regular voters, as a proportion of the voting-eligible population, and the proportion of regular voters supporting the Democratic Party over the Republican Party. Let us call these quantities P_R (proportion of the population occupied by regular voters) and V_R (vote preferences of regular voters).

The crucial unknown quantity is V_M, the partisan preferences of marginal voters. With this quantity we can directly compute $V_M - V_R$, the preference gap between marginal and regular voters and the primary quantity of interest for this study. Unfortunately, this quantity cannot be recovered using only the typical research toolkit (regressions, averages, differences-in-means, etc.). However, we can solve the following equation to back out V_M, and therefore $V_M - V_R$:

$$P_R V_R + P_M V_M = (P_R + P_M)(V_R + \Delta V).$$

Solving for V_M, we get:

$$V_M = [(P_R + P_M)(V_R + \Delta V) - P_R V_R] / P_M,$$

$$= V_R + \Delta V(1 + P_R/P_M).$$

Finally, we can see that the preference gap between marginal and regular voters is

$$V_M - V_R = \Delta V(1 + P_R/P_M).$$

In words, the preference gap is the effect of the treatment on vote share times one plus the ratio of the regular and marginal populations.

Intuitively, we can identify the partisan preference of marginal voters because we know the relative populations of regular and marginal voters, the partisan preferences of regular voters, and the extent to which the introduction of the marginal voters changes the election result. Knowing these quantities, we can then determine the way in which marginal voters must have voted in order to obtain the effects that we see. The average voting behavior of the entire population is just a weighted average of the voting behavior of regular and marginal voters, with weights proportional to the populations of the two groups. Reliable estimates of uncertainty for all quantities can be obtained through bootstrap simulations.

To clarify this procedure, let us consider a numerical example. Suppose an experiment raised voter turnout from 40 percent or .40 in the control group to 60 percent or .60 in the treatment group. Suppose also that this treatment increased the Democratic vote share from .50 in the control group to .55 in the treatment group. If we can safely assume that this experimental intervention did not change the vote choices of regular voters and only changed the election result through its effect on the composition of the electorate, then we can apply the test to calculate the preference gap between the regular and marginal voters. In this example, regular voters comprise 40 percent of the population (P_R = .4), as indicated by the turnout level in the control group. Marginal voters comprise 20 percent of the population (P_M = .2), as indicated by the difference in turnout between treatment and control. Regular voters are evenly split between the two parties (V_R = .5), as indicated by the Democratic vote share in the control group, but we do not yet know how the marginal voters are divided. The treatment increased the Democratic vote share by 5 percentage points (ΔV = .05), so we can input these figures into the equation above to back out the partisan preferences of marginal voters:

$$P_R V_R + P_M V_M = (P_R + P_M)(V_R + \Delta V),$$

$$.40*.50 + .20*V_M = (.40 + .20)(.50 + .05),$$

$$V_M = [.60*.55 - .40*.50]/.20,$$

$$V_M = .65,$$

$$V_M - V_R = .65 - .50 = .15.$$

In this numerical example, 65 percent of the marginal voters must have supported the Democratic Party, relative to just 50 percent of the regular voters, in order to obtain the results of the experiment. Therefore, the partisan preference gap between marginal and regular voters is .15 or 15 percentage points. Hopefully this hypothetical example clarifies the intuition and simplicity of this test.

Most studies examining the electoral effects of turnout would simply calculate the effect of a treatment on turnout and vote share, and stop there without ever calculating the preference gap. We can see from the equations that ΔV indicates the sign of the preference gap: if ΔV is positive, then the preference gap ($V_M - V_R$) will also be positive (marginal voters are more supportive of the Democratic Party than regular voters) and vice versa. However, the value of ΔV alone provides no substantive meaning about the differences between regular and marginal voters. For the purposes of understanding political representation, this substantive quantity is crucial. In the next sections, I apply this test to three different sources of exogenous variation in turnout – weather, the timing of gubernatorial elections, and the timing of congressional elections. The population of marginal voters is different in each setting, so the preference gap is different as well. Nonetheless, across all three settings, I find that marginal voters are systematically more supportive of the Democratic Party than regular voters. The repeated application of this test in different settings improves our understanding of the marginal voting population, the extent to which election results reflect public preferences, and the consequences of electoral reforms.

Marginal Voters Sensitive to Weather in Presidential Elections

To present an example of the descriptive power of the method, I first apply it to a known quasi-experiment previously analyzed by Gomez, Hansford, and Krause (2007) and Hansford and Gomez (2010). The turnout decisions of some citizens are sensitive to weather. Specifically, these marginal voters are less likely to vote if it rains or snows, presumably due to the difficulties of getting to the polling place in these conditions. Gomez, Hansford, and Krause (2007) demonstrate that rain and snow decrease turnout and also benefit the Republican Party in presidential elections, and Hansford and Gomez (2010) extend this analysis with an instrumental variables study of the effects of turnout. However, neither paper directly leverages the data to compare the partisan preferences of regular voters and those marginal voters who are sensitive to weather.

In Table 2.1, I reanalyze their data set of county-level voting in presidential elections from 1946 to 2000, made publicly available by Fraga and Hersh (2011). Each observation is a county-election. I code a single dummy variable which indicates whether there was any precipitation in a county on Election Day. Then, conducting the simplest possible tests, I estimate the effect of precipitation on turnout and Democratic vote share by regressing these outcomes on the precipitation dummy variable, county fixed effects, and election fixed effects. Because of the fixed effects, these are difference-in-difference estimates which indicate the extent to which turnout or vote share vary *within* counties as the presence of precipitation changes. The top panel of Table 2.1 shows that the lack of precipitation, on average, increases a county's turnout by 0.6 percentage points which in turn increases their Democratic vote share by 0.5 percentage points.

Having estimated the effect of some factor on turnout and vote shares, most studies would stop here, noting that these effects are "statistically significant." However, the coefficients above, 0.6 and 0.5, do not convey the information that we care about most. Digging deeper into the data, we can extract more meaningful quantities. First, the 0.6 percentage point effect of precipitation on

Table 2.1. Exploiting Weather in Presidential Elections (1948-2000)

	DV = Turnout	Democratic Vote Share	
No Precipitation	.006**	.005**	
	[.004,.008]	[.003,.007]	
County Fixed Effects	X	X	
Election Fixed Effect	X	X	
R-squared	.72	.63	
Observations	43,340	43,340	
	Regular Voters	Marginal Voters	Difference
Proportion of Population	.576**	.006**	
	[.572,.581]	[.004,.008]	
Support for Democratic Party	.489**	.955**	**.467**
	[.486,.492]	[.732,1.311]	**[.242,.823]**

*Block bootstrap 95% confidence intervals in brackets; **p < .01.*

No Precipitation is a dummy variable indicating whether there was no precipitation in a particular county on election day. The top panel shows that good weather (relative to some rain or snow) increases turnout by 0.6 percentage points which in turn increases the Democratic candidate's vote sharein the county by 0.5 percentage points. The bottom panel shows that these marginal voters (0.6 percent of the population) who are sensitive to the weather are 46.7 percentage points more likely to support the Democratic candidate compared to the regular voters (57.6 percent of the population) who would have voted regardless of the weather.

turnout tells us that 0.6 percent of voting-eligible citizens are marginal voters with respect to weather. They will vote if the sky is clear, but if it rains or snows, they will abstain. Obviously, we do not know the names of these individuals, and they themselves may not know that they are part of this marginal population. However, in any given election, about 1 in 200 people will have either voted only because the weather was nice or abstained only because it rained or snowed.

Next, I quantify the number of regular voters – those who would have voted regardless of the weather. Because this wasn't an ordinary experiment, we cannot simply calculate the mean turnout of the control group, but the intuition is similar. After running the regression, I calculate each county's predicted level of voter turnout under the counterfactual scenario of precipitation. Averaging across these predicted values, I obtain the expected level of turnout with rain everywhere: .576, the proportion of the voting-eligible population made up of regular voters. Applying the same logic to the vote share regression, I also calculate the expected level of Democratic vote share with rain or snow everywhere: .489, the proportion of regular voters supporting the Democratic Party.

We have now computed all of the quantities necessary to estimate the partisan preferences of the marginal voters and the partisan gap between marginal and regular voters. Solving the equation provided in the previous section, I estimate that 95.5 percent of marginal voters supported the Democratic Party. Compared to the regular voters, marginal voters are 46.7 percentage points more likely to support the Democratic Party – a massive divide. 95% confidence intervals are calculated for each of these quantities through block-bootstrap simulation (where counties are sampled as blocks), and all estimated quantities including the preference gap are statistically significant ($p < .01$).

If we had simply looked at the top panel of Table 2.1, the traditional regression results, we would have missed the bigger picture. Those marginal voters whose turnout decisions are influenced by weather support the Democratic Party at an astounding rate. Obviously, these

individuals are a unique subset of Americans, only 0.6 percent of the voting-eligible population. Nonetheless, the fact that these marginal citizens on the cusp of voting or abstain are so different from regular voters is alarming for democracy. When trivial factors like rain can dramatically change the composition of the electorate, we should have little faith that the regular voters are adequately representative of the population as a whole.

Marginal Voters Sensitive to Timing of Gubernatorial Elections

The method for comparing regular and marginal voters is generally applicable to any exogenous factor that influences voter turnout and brings marginal voters to the polls (or keeps them away). Having demonstrated the power of the method with weather, a previously studied factor, I introduce several new opportunities to compare regular and marginal voters. Several recent studies have exploited the timing of elections to understand the effects of mass participation on interest group power and policy outcomes. When local elections are held "on cycle" – meaning that they coincide with national elections – turnout is significantly higher than when they are held "off cycle." This uptick in participation in school board elections decreases the chances of interest group capture and leads to lower teacher pay (Anzia 2011, 2012; Berry and Gersen 2011). Beyond local elections, turnout in federal and statewide elections is also highly dependent on election timing. When gubernatorial and congressional elections happen to coincide with a presidential race, turnout is significantly higher. In other words, a large population of marginal voters turns out simply because of the more salient race at the top of the ticket. Even though this phenomenon is well-known, no previous study has exploited election timing to compare the partisan preferences of regular and marginal voters.

States vary idiosyncratically in the timing of their gubernatorial elections. Two states – NH and RI – hold their elections every two years (on even years – i.e. 2006, 2008, 2010). Two other

states – NJ and VA – hold their elections on odd years following presidential election years (e.g. 2005, 2009, 2013). Three states – KY, LA, and MS – hold their elections on odd years preceding presidential election years (e.g. 2003, 2007, 2011). Nine states – DE, IN, MO, MT, NC, ND, UT, WA, and WV – coincide their elections with presidential elections (i.e. 2004, 2008, 2012), while the remaining 34 states hold their gubernatorial elections in midterm years (i.e. 2002, 2006, 2010).

The timing of each state's elections was typically established long ago for arbitrary reasons unrelated to the political orientation of the state. Conducting a thorough review of the history of each state's electoral calendar, I found that 12 states have maintained the same electoral calendar since the founding of the state. In each of these cases, the timing was determined by the first year that each state could hold a gubernatorial election, and that timing has remained unchanged. In three other states, the timing was influenced by the death of the sitting governor (DE and KY) or a governor who ascended to the Senate in the middle of his term (CA). In each of these three cases, the state had no established law regarding the replacement of a sitting governor, so they established that law at the time and kept their electoral calendar in place from that point on. In 31 states, the vast majority, the timing of gubernatorial elections was set when, at some point, the state decided to switch from 2 or 3 year terms to 4 year terms. The particular timing of the elections was typically set for whenever the next election could be held, and this decision was probably unrelated to the political leanings of the state.

Finally, there are 5 remaining states (LA, MD, MN, FL, IL) which actively changed their electoral calendar through a state law. Louisiana switched from presidential years to odd years in 1879, Maryland switched from odd years to midterm years in 1926, and Minnesota, Florida, and Illinois switched from presidential to midterm years in 1930, 1966, and 1978, respectively. For the purposes of this study, these 5 states could be problematic because they may have strategically changed their electoral calendar for partisan, political reasons. Looking into these legislative debates,

the primary arguments given for switching are non-partisan. Legislators on both sides of the aisle in Florida, for example, wanted to move away from presidential years so that the gubernatorial race would not be drowned out by the presidential campaign. Importantly, concerns about strategic timing in these 5 states do not influence the subsequent empirical results. The findings are unchanged if these 5 states are removed or if they are treated as non-compliers in an instrumental variables framework.

This history of gubernatorial election calendars suggests that those states whose calendar happens to coincide with the presidential calendar are, on average, no different from those with another calendar. As a result of various historical accidents, these two groups of states, despite being comparable in terms of their governments, partisan leanings, and demographic makeup, happen to elect their governors at a time when many more people turn out to vote. Figure 2.1 compares the two groups of states that I leverage in my analysis – "on cycle" states whose gubernatorial elections coincide with presidential elections and "off cycle" states whose elections occur at other times. Consistent with the claim that these electoral calendars are idiosyncratically determined, these two groups of states are nearly identical in their average voting behavior in presidential elections. They turn out and support Democratic candidates at the same rates. However, the "on cycle" states turn out at much higher rates in gubernatorial elections, providing the necessary leverage to compare regular voters and marginal voters who only turn out because of the presidential race. Statistical tests of balance reveal no substantively or statistically significant differences in turnout (difference = .012; p = .562) or vote share (difference = .026; p = .353) between on cycle and off cycle states in presidential elections. However, the turnout differential in gubernatorial elections is substantively large (15.5 percentage points) and statistically significant (p < .001).

Figure 2.1. Comparing States with On and Off Cycle Gubernatorial Elections
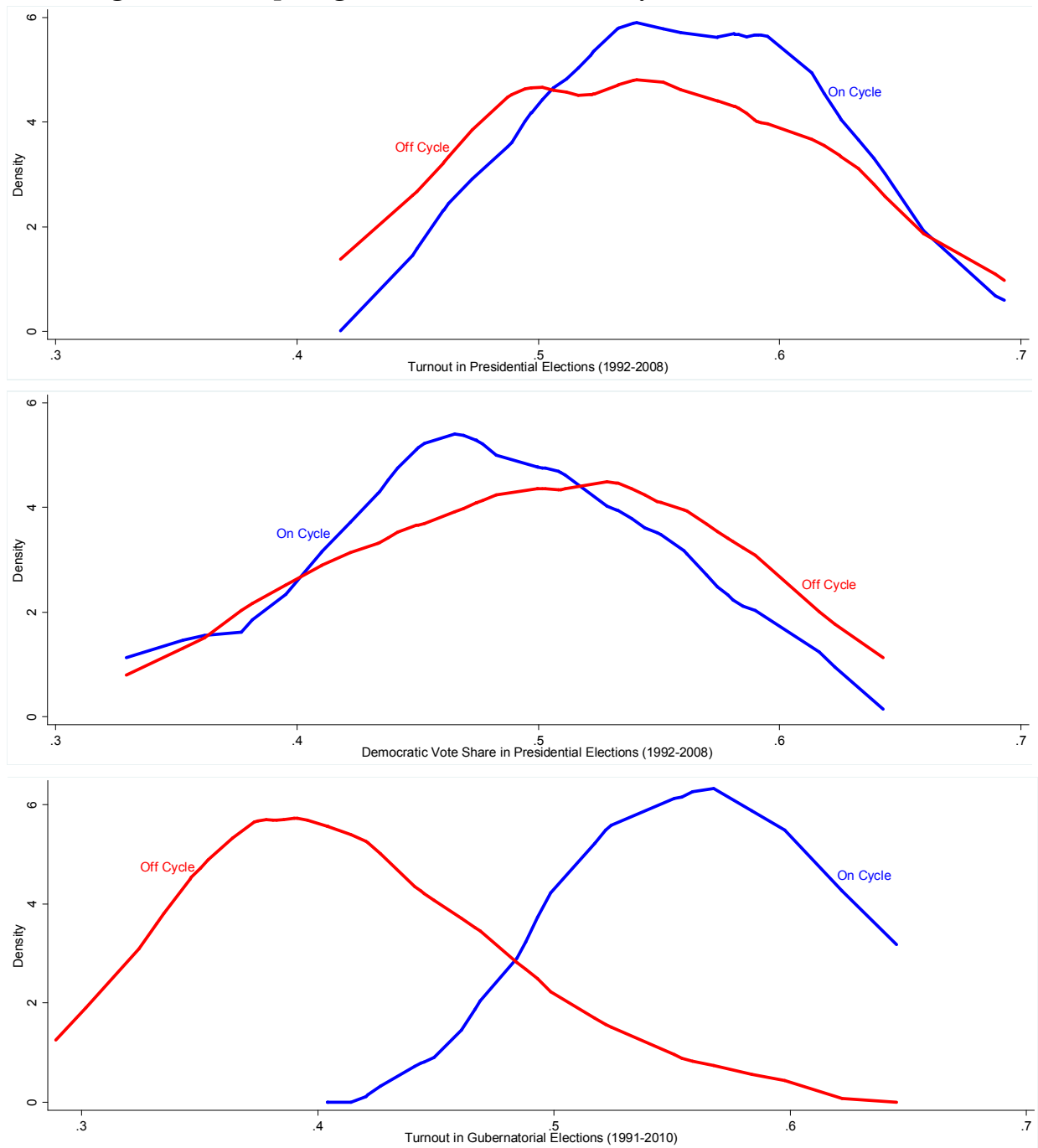
The figure compares "on cycle" states – those with gubernatorial elections coinciding with presidential elections – to "off cycle" states. New Hampshire and Vermont are excluded from the figure because they have two-year instead of four-year terms and therefore alternate between being on and off cycle. Kernel density plots show that on and off cycle states are nearly identical in terms of their voting behavior in presidential elections. They turn out at the same rates and support Democratic candidates at the same rates. However, in gubernatorial elections, on cycle states turn out at much higher rates, providing the opportunity to compare the preferences of regular voters to those of marginal voters – citizens who are mobilized to vote in the gubernatorial election because it happens to coincide with a presidential race.

Table 2.2. Exploiting the Timing of Gubernatorial Elections (1991-2010)

	DV = Turnout	Democratic Vote Share	Democratic Victory
On-Cycle	.174**	.064**	.454**
	[.139,.197]	[.027,.090]	[.251,.610]
Normal Presidential Vote	-.225	.401**	2.180**
	[-.506,.056]	[.193,.608]	[.640,3.720]
Normal Presidential Turnout	.742**	.022	-.072
	[.533,.951]	[-.353,.397]	[-1.793,1.937]
Constant	.010**	.264**	-.754
	[-.047,.246]	[.099,.428]	[-.1.533,.024]
R-squared	.57	.09	.14
Observations	261	261	261

	Regular Voters	Marginal Voters	Difference
Proportion of Population	.395**	.174**	
	[.377,.414]	[.139,.197]	
Support for Democratic Party	.472**	.682**	**.210****
	[.458,.488]	[.569,.778]	**[.093,.308]**

Block bootstrap 95% confidence intervals in brackets; **p < .01.
On-Cycle is a dummy variable indicating whether an election coincided with a presidential election. Normal Presidential Vote is a control variable indicating the state's average Democratic vote share in the past three presidential elections. Normal Presidential Turnout is a control variable indicating the state's average turnout in the past three presidential elections. The top panel shows the on-cycle elections increase turnout by 17.4 percentage points which in turn increases the Democratic candidates' vote shares by 6.4 percentage points. Because most gubernatorial elections are competitive, this is enough to increase the Democratic candidate's probability of winning by 45.4 percentage points. All 3 regressions are weighted by the voting-eligible population of the state, because the appropriate unit of analysis is the individual citizen. The bottom panel shows that marginal voters (17.4 percent of the population) who are mobilized by the occurrence of a presidential election are 21.0 percentage points more likely to support the Democratic Party compared to the regular voters (39.5 percent of the population) who would have voted regardless of the presidential race.

In Table 2.2, I estimate the effect of "on cycle" gubernatorial elections on voter turnout, Democratic vote share, and the probability of Democratic victory. Each observation is a state-election from 1991 to 2010. Each observation is weighted by the voting-eligible population in each state (see McDonald and Popkin 2001), because the primary unit of analysis is the voter. However, unweighted estimates yield similar results. These effects are estimated from regressions of each dependent variable on a dummy variable for an "on cycle" election – meaning that the election coincided with a presidential race – and two control variables – each state's average level of turnout and Democratic vote share in the past three presidential elections. Because on and off cycle states are balanced in terms of these variables, controls are unnecessary for unbiased estimates, but they are included for statistical precision. These regressions indicate that on cycle gubernatorial elections, on average, increase turnout by 17.4 percentage points and the Democratic candidate's vote share by 6.4 percentage points. Moreover, because most gubernatorial races are competitive, this increase in Democratic vote share is enough to increase the probability of Democratic victory by 45.4 percentage points. While weather is unlikely to change many election results because the small effects are attenuated in close races (Fraga and Hersh 2011), election timing determines the winner and loser of many elections.

Applying the method for comparing marginal and regular voters, I find that that regular voters – 39.5 percent of the voting-eligible population – support Democratic gubernatorial candidates 47.2 percent of the time while marginal voters – 17.4 percent of the voting-eligible population – support Democratic candidates at a rate of 68.2 percent. Again, with a different, independent sample of marginal voters, I find a massive preference gap between regular and marginal voters. The citizens whose turnout decision are influenced by gubernatorial election timing are 21 percentage points more likely to support the Democratic Party than regular voters. As before, I estimate 95% confidence intervals for all quantities with block-bootstrap simulations (in

this case states are sampled as blocks), and all estimates are statistically different from zero ($p < .01$). Those voters comprising the regular electorate are systematically different from those on the margins, and when marginal voters are introduced into the electorate, partisan election results can change dramatically.

Marginal Voters Sensitive to Timing of Congressional Elections

In addition to the idiosyncratic timing of gubernatorial elections, I exploit the timing of congressional elections as a third opportunity to compare regular and marginal voters. The U.S. House of Representatives presents a particularly advantageous opportunity because there are many districts (435) and each seat comes up for reelection every two years. Half the time, elections take place in midterm years and half the time elections coincide with presidential elections where turnout is significantly higher. Also, because each district experiences both on and off cycle elections, we can focus on within-district variation in election timing and avoid concerns about heterogeneity between districts.

Theoretically, we should expect the preference gap between regular and marginal voters to be smaller in congressional elections, because most members of the House seek reelection (Jacobson 1997) and incumbents receive a large electoral benefit simply by virtue of being the incumbent (Gelman and King 1990; Ansolabehere and Snyder 2002, 2004; Fowler and Hall 2012). If this incumbency advantage influences both regular voters and marginal voters, then we will see a smaller gap in their partisan preferences.[12]

_____

[12] We may also like to know how regular and marginal voters differ in their preference for incumbents. However, this research design does not allow us to address this question because the party of the incumbent is causally influenced by the timing of the past election which is strongly correlated with the timing of the current election.

Table 2.3. Exploiting the Timing of U.S. House Elections (1982-2010)

	DV = Turnout	Democratic Vote Share	Democratic Victory
On-Cycle	.134**	.015**	.015**
	[.130,.138]	[.009,.021]	[.005,.025]
District Fixed Effects	X	X	X
R-squared	.78	.76	.80
Observations	6388	6388	6388

	Regular Voters	Marginal Voters	Difference
Proportion of Population	.342**	.134**	
	[.336,.346]	[.130,.138]	
Support for Democratic Party	.517**	.571**	**.053****
	[.506,.530]	[.549,.592]	**[.032,.074]**

*Block bootstrap 95% confidence intervals in brackets; **p < .01.*
On-Cycle is a dummy variable indicating whether an election coincided with a presidential election. The top panel shows the on-cycle elections increase turnout by 13.4 percentage points which in turn increases the Democratic candidates' vote shares and probabilities of victory by 1.5 percentage points. The bottom panel shows that these marginal voters (13.4 percent of the population) who are mobilized by the occurrence of a presidential election are 5.3 percentage points more likely to support Democratic candidates than the regular voters (34.2 percent of the population) who would have voted regardless of the presidential race.

One concern in comparing presidential and midterm elections is that idiosyncratic effects of time and national mood could cause regular voters to change their voting behavior between these two types of elections. This concern applies to both the analysis of gubernatorial and congressional elections. Multiple tests suggest that this concern does not plague the subsequent estimates. The inclusion of numerous time-varying covariates such as presidential approval, national party identification, and the interaction of the president's party with these variables does not change the results of either the gubernatorial or congressional analyses. With enough elections in the sample, idiosyncratic events that may lead midterms to be different from presidential elections appear to have canceled out.

Closely mirroring the gubernatorial analysis, Table 2.3 presents the effect of congressional election timing on turnout, Democratic vote share, and the probability of a Democratic victory. Each observation is a district-election between 1982 and 2010. These estimates are obtained from regressions of each dependent variable on a dummy variable for an "on cycle" election and district fixed effects. Presidential election years increase turnout in House elections by 13.4 percentage points, which in turn increases the Democratic candidate's vote share and probability of electoral victory by 1.5 percentage points. Solving for the preferences of regular and marginal voters, I estimate that regular voters – 34.2 percent of the voting-eligible population – support Democratic congressional candidates at a rate of 51.7 percent while marginal voters – 13.4 percent of the voting-eligible population – support Democratic candidates at a rate of 57.1 percent. Because of the electoral setting and the different subset of marginal voters, this 5.3 percentage point preference gap is smaller than the previous estimates. Nonetheless, this gap is still substantively meaningful and statistically significant (p < .01; 95% confidence intervals estimated from block-bootstrap simulations where districts are samples as blocks). Even in a setting where elections are less

56

competitive and incumbents are extremely popular, regular and marginal voters still diverge

significantly in their preferences.

Observational Estimates of Partisan Preference Gaps

Across three different settings, I have identified preference gaps of 5, 21, and 47 percentage

points between marginal and regular voters. In each case, the marginal voters are systematically

more supportive of the Democratic Party. Also, in each case, the estimates have arisen from

exogenous changes to turnout in real-world elections, so concerns about confounding variables and

survey misreporting are mitigated. How different would our results be if we relied entirely on

observational or survey evidence? In this section, I benchmark the quasi-experimental estimates

against several observational estimates – some familiar and others novel. Results are shown in Table

2.4.

One obvious method for estimating the partisan preference gap is to directly ask non-voters

how they would have voted. Averaging across all ANES surveys in presidential years from 1952 to

2008 (all years when non-voters were asked about their presidential preferences) and focusing on

only those individuals who expressed a preference for one of the two major parties, I find that

50.0% of voters supported the Democratic Party, while 55.1% of non-voters supported the

Democratic Party. Previous studies making use of these survey questions (e.g. Wolfinger and

Rosenstone 1980; Highton and Wolfinger 2001) failed to detect this gap because they never

calculated these crucial quantity of interest. Interestingly, non-voters report similar preferences to

those of marginal voters who abstain in midterms and participate in presidential elections – with

both groups supporting Democratic candidates at higher rates than regular voters.

Instead of directly asking respondents how they would have voted, we can statistically model

their likely vote choices similar to Herron (1998), Citrin, Shickler, and Sides (2003), and Martinez

Table 2.4. Observational and Quasi-experimental Estimates of Preference Gaps between Regular and Marginal Voters

	Regular	Marginal	Difference
Reported presidential preferences of voters and non-voters (ANES, 1952-2008)	.500	.551	.051
Predicted presidential preferences of voters and non-voters (ANES, 1952-2008)	.500	.554	.054
Presidential preferences of voters and those who intended to vote (ANES, 1952-2008)	.500	.581	.081
Party registration of voters and non-voters in 2010	.536	.631	.095
Rain and snow in presidential elections (1948-2000)	.489	.955	.467
Gubernatorial election timing (1991-2010)	.472	.682	.210
Congressional election timing (1982-2010)	.517	.571	.053

The table presents 7 different comparisons of regular and marginal voters. Row 1 shows the reported presidential preferences of voters and non-voters in ANES surveys from 1952 to 2008. Only those individuals who expressed a preference for one of the two major-party candidates are included in the analysis. Row 2 shows the predicted preferences of these individuals generated from a multinomial logit model using age, race, and education to predict vote choice. Both approaches yield nearly identical results, a 5 percentage point preference gap between voters and non-voters. Row 3 compares the stated preferences of voters to that of non-voters who intended to vote before the election. Row 4 compares the party registration of voters and non-voters in 2010 who were registered with one of the two major parties. Non-voters are 9.5 percentage points more likely to side with the Democratic Party. The final 3 rows re-print the results from tables 2.1, 2.2, and 2.3 where those voters sensitive to weather and election timing are compared to regular voters. Across all tests, non-voters or marginal voters are systematically more supportive of the Democratic Party than regular voters.

and Gill (2005). Examining the same ANES sample from 1952 to 2008, I conduct a multinomial logit analysis (McFadden 1974) where each respondent's probability of voting for a Democrat, voting for a Republican, or abstaining is predicted using her age, gender, race, and education level. Unlike Martinez and Gill, I do not include attitudinal variables such as party identification because those preferences are endogenous to the decision to turn out (see the previous Lijphart quote). With these predicted probabilities, I calculate the conditional probability that each respondent would support the Democratic candidate if she voted – Pr(D)/[Pr(D) + Pr(R)]. As expected, the average conditional probability for voters is nearly identical to the average level of Democratic support – 50.0%. Remarkably, the average predicted probability for non-voters is very similar to the average level of reported preferences – 55.4%. With both methods, I estimate a 5 percentage point preference gap between voters and non-voters, on average, across all presidential elections from 1952 to 2008.

One particularly interesting group of ANES respondents is those intended to vote before the election but did not. These individuals – 3.1% of all respondents – are similar to other groups of marginal voters in this study because they may have been on the cusp of voting but failed to turn out for one reason or another. Perhaps they were busier than expected on Election Day, couldn't find their polling place, or were deterred by bad weather. According to their survey responses, 58.1% of these individuals preferred Democratic candidates, creating an 8 percentage point gap between voters and those who intended to vote but did not.

State records of registered voters represent another valuable source of information on marginal voters. All states maintain records of which individuals turned out in recent elections, and 31 states keep records of party registration. Analyzing these voter lists, as compiled by Catalist Inc., following the 2010 election and focusing exclusively on those states with party registration, we can assess the partisan difference between voters and non-voters. Among the 42 million individuals

who are registered with a major party and voted in 2010, 53.6 percent of them side with the Democrats. Alternatively, among the 34 million individuals registered with a party who did not vote, 63.1 percent of them side with the Democrats. Between these two groups of citizens, we see a 9.5 percentage point preference gap.

The final three rows of Table 2.4 present the quasi-experimental comparisons of regular and marginal voters that represent the crux of this study. Across all 7 different comparisons, non-voters or marginal voters are systematically more Democratic than regular voters. The estimated gap between regular and marginal voters varies from 5 to 47 percentage points depending on the setting or the particular subset of marginal voters. Those marginal voters sensitive to weather appear to be unique and not representative of the entire population of non-voters, while those marginal voters sensitive to election timing may be more representative of the non-voting population. Also, despite concerns that observational methods may under-report the true preference gap between regular and marginal voters, substantively large gaps appear with all methods.

Conclusion

The academic literature on the effects of turnout can be confusing, because different studies ask different questions (Grofman, Guillermo, and Collet 1999). What if everyone voted? How different are voters and non-voters? Does higher turnout benefit one party or the other? The first two questions may be impossible to answer in the U.S. and the third question has no definite answer. Higher turnout could help either party depending on who is mobilized. This paper focuses on the more relevant and reachable goal of comparing regular voters to those marginal voters whose decisions to turn out are sensitive to exogenous factors. For the purposes of understanding political representation and the effects of electoral reforms, we may care most about the attitudes of those citizens right at the margins. No single regression can resolve the debate. However, the repeated

analysis of multiple, independent, exogenous factors that influence turnout will help us to understand the preference of these marginal voters.

This paper presents a simple method for comparing the preferences of regular and marginal voters and applies it in three different settings. The major contribution of the method is that it translates opaque regression coefficients into a valuable quantity of interest – the partisan preference gap between regular and marginal voters. However, the method is not a magic bullet. Researchers must still work hard to identify new, exogenous sources of variation in turnout which provide opportunities to understand marginal voters.

Every test in this paper points in the same direction. Regardless of the setting or the particular sample of marginal voters, regular voters are not representative of the larger pool of possible voters. Citizens on the margins are systematically more supportive of the Democratic Party than regular voters, and this gap can have significant electoral consequences. For example, the party of many states' governors would be different if their elections were held in different years when a different subset of citizens turns out. Even if the introduction of marginal voters in the electorate does not change discrete electoral outcomes, the new composition of voters may still influence the platforms of candidates and the distribution of public services. Currently, American elections fail to reflect the preferences of all citizens because those on the margins are systematically different from those regularly participating. The repeated testing of preference gaps between marginal and regular voters may improve our understanding of this phenomenon and identify solutions for the mitigation of this participatory inequality.

Study #3

Social Capital and Voter Turnout: Evidence from Saint's Day Fiestas in Mexico

Co-authored with Matthew D. Atkinson

Abstract

Social capital and community activity are thought to increase voter turnout, but reverse causation and omitted variables may bias the results of previous studies. We exploit saint's day fiestas in Mexico as a natural experiment to test this causal relationship. Saint's day fiestas provide temporary but large shocks to the connectedness and trust within a community, and the timing of these fiestas is quasi-random. Employing both cross-municipality and within-municipality estimates, we find that saint's day fiestas occurring near an election decrease turnout by 2.5 to 3.5 percentage points. Community activities which generate social capital can inhibit political participation, giving pause to scholars and policymakers who assume that such activity will improve the performance of democracy.

Collective action problems are an enduring concern in political science and all of social science. Activities that are vital to a community may be under-provided by self-interested individuals (Olson 1965). Voter turnout and political participation represent one such collective action problem because the community benefits are significant (Lijphart 1997), but the individual incentives are low (Downs 1957). Social capital has been offered as one potential solution to low turnout and other collective action problems (Putnam 1993, 1995, 2000). This paper explicitly tests the validity of this claim.

Social capital refers to the degree of connectedness and trust within a community which are built through interactions among community members.[1] In this paper, we exploit a natural experiment to test the effect of social capital on voter turnout. Does community activity which increases trust and connectedness lead to collectively beneficial behavior and improve the quality of democracy? In particular, we evaluate whether voter turnout is affected by the timing of large-scale community interactions which create and reinforce social capital.

Previous scholars have proposed several ways in which community activity and subsequent social capital can increase collective actions such as voter turnout. First, well-connected communities can more easily enforce cooperative behavior (Stolle 1998), and recent field experiments demonstrate the potential role of social pressure in voter mobilization (Gerber, Green, and Larimer 2008; Nickerson 2008). Second, social capital can increase turnout by increasing the flow of information. While the direct costs of voting are a significant barrier to turnout, social

---

[1] While numerous and conflicting definitions of *social capital* exist in the literature, we restrict ourselves to this specific definition. Previous scholars make distinctions between bridging vs. bonding social capital as well as the relative importance of connectedness vs. trust. We are agnostic in regard to these debates, and our identification strategy exploits variation in both connectedness and trust along with both bridging and bonding social capital.

capital may alleviate the additional costs of becoming engaged and informed about politics (Fiorina 1990; Berinsky 2005). Lastly, social capital may expose citizens to potential benefits for others and increase their incentive to vote (Duch and Palmer 2004; Edlin, Gelman, and Kaplan 2007; Fowler 2006; Rotemberg 2009).

Despite these theoretical predictions that social capital will increase democratic participation, there are competing possibilities. First, time is an essential resource for voting and other collective behaviors (Verba, Schlozman, and Brady 1995), and community activities can consume much of an individual's available time (Rupasingha, Goetz, and Freshwater 2006). Second, social capital exposes citizens to conflicting views, which may create uncertainty and depress participation (Mutz 2002). Lastly, social capital could decrease voter turnout through personal satisfaction. Previous theories suggest that voters derive satisfaction from the act of voting itself (Riker and Ordeshook 1968), and social capital provides an alternative way in which citizens can achieve this personal fulfillment. Furthermore, even if connectedness and trust themselves are not detrimental for participation, the community activities recommended by scholars to increase social capital may have adverse consequences, distracting citizens and depressing political participation. In fact, governments and development organizations are currently implementing policies to increase community activity on the assumption that this will foster and maintain democracy (Nelson 1995; Paxton 2002; Krishna 2007), so an empirical assessment of these effects is needed.

Putting the theoretical arguments together, social capital and the community activities which generate social capital may have positive or negative effects on collective action, and they could serve as either a complement or a substitute for political participation. In this paper, we explicitly assess the relationship between social capital and voter turnout by exploiting exogenous increases in community interactions which build and reinforce social capital. In doing so, we test a fundamental question in the literatures on social capital and political participation: can increased community

64

activity lead to increased participation thereby improving the health of democracy? In the ideal experiment, we would randomly treat some individuals or communities with high doses of community engagement right around the time of an election. While this experiment is practically unfeasible, we identify a naturally occurring process which closely mimics this ideal experiment: saint's day fiestas in Mexico.

To assess the relationship between social capital and voter turnout, we exploit two sources of quasi-random variation. First, the fiesta date of each community is quasi-random, such that a community with a fiesta date near an election is on average no different than a community with a fiesta date far from the election, aside from their differing fiesta dates. Comparing across these groups of communities, we can estimate the causal effect of heightened community interaction right around the time of an election. Second, the timing of Mexico's federal elections switched from mid-August in 1994 to early July in 1997. As a result, some communities which previously had fiestas near the election no longer do and vice versa. Exploiting this quasi-random shock to the electoral calendar, we test whether voter turnout within a community changes as the election date move close to or further from the fiesta date. With both tests, we estimate a significant, negative effect of social capital on voter turnout. When a saint's day occurs within two weeks of an election, turnout in that community is depressed by 2.5 to 3.5 percentage points.

In the next section, we discuss previous attempts to assess the effect of social capital, and we explain why these studies systematically overestimate the true causal effect. Next, we provide a brief history of saint's day fiestas, demonstrate that the fiesta dates are quasi-random, and show that the fiestas successfully generate social capital. Then, we provide the main empirical results, analyzing the voting behavior of 325 Mexican municipalities with only one Catholic Church in 7 federal elections between 1991 and 2009. Then, we extend the analysis to present three additional sets of findings. (1) Because turnout is persistent across elections, a fiesta which demobilizes a community

in one election will depress turnout in subsequent elections as well. (2) The size of the effect of saint's day fiestas varies in important ways. The effect is stronger (more negative) in smaller, more Catholic communities where the fiestas are more intense. Also, the effect is equally strong when local elections coincide with the federal election even though overall participation is greater during local elections. (3) Having analyzed rural municipalities, we replicate the same effect in Monterrey, a large urban setting. Then, we discuss and rule out several alternative explanations of our findings. Finally, we conclude by discussing the implications of our results for collective action and democratic governance.

In numerous academic fields, social capital is posited to boost the economic, social, and political health of society (Putnam 1993; Knack and Keefer 1997). Despite the importance of these causal claims, they have not been rigorously tested. Not surprisingly, individuals who are connected to their community or trust their community are more likely to participate in politics. However, it remains to be determined whether social capital as a policy prescription could causally increase democratic participation. To our knowledge, this paper offers the best available evidence on the causal effects of social capital and the types of community activity which generate social capital. We find no evidence that social capital causally increases collective action; in fact, an exogenous increase in community activity actually decreases voter turnout.

Previous Estimates and the Problem of Endogeneity

Existing empirical evidence for the effects of social capital is correlational: trends in measures of social activity over time correspond with trends in turnout and other collective actions (Skocpol 1997; Putnam 2000), regions with high measures of social activity tend to have higher turnout (Putnam 2000), and individuals who are socially connected are more likely to vote (Knack 1992; Lake and Huckfeldt 1998; Fowler 2005). Moreover, citizens who participate in their

community are more likely to participate in politics (Verba, Schlozman, and Brady 1995). However, these correlations lack a causal interpretation due to reverse causation and confounding variables. Social capital is not randomly assigned but rather an endogenous characteristic of generations, regions, and individuals. For instance, the act of participating in politics may generate social capital. Additionally, the types of individuals who are socially connected are probably the types of citizens who would vote regardless of their social situation.

Even in 1840, Alexis de Tocqueville acknowledged the possibility that social capital is endogenous to political activity: "Civil associations, therefore, pave the way for political associations; on the other hand, political associations develop and improve in some strange way civil associations" (De Tocqueville 1840). This type of reverse causation would lead any correlational finding to overestimate the true causal effect of community engagement on political participation. Several studies suggest that good political institutions foster community connectedness but the reverse relationship does not hold (Booth and Richard 1998; Letki and Evans 2005).

One important confounding variable in these studies is an individual's underlying level of sociability. Researchers have shown that personality traits such as extroversion, social aggression, and self-confidence have direct effects on both measures of social capital (Scheufele and Shah 2000) and voter turnout (Gerber et al. 2011). Individuals predisposed to be social are more likely to be involved in their communities and are more likely to vote. However, forcing a non-social person to connect with others may have no impact on her decision to vote. Numerous other omitted variables may further cloud the interpretation of the generational, regional, and individual correlations between social capital and turnout.

Ideally, we could obtain an unbiased estimate of the effect of community participation on turnout through a randomized, controlled experiment. In the correlational observations, socially connected individuals or groups are significantly different from those who are unconnected.

Randomization would remove endogeneity because we could be sure that the comparison groups are truly comparable to one another. We would randomly assign some individuals or groups to be socially connected and others to be disconnected. Unfortunately, such an experiment would be practically, financially, and ethically unfeasible. Therefore, we exploit a natural experiment in which social capital is assigned in a quasi-random manner: saint's day fiestas in Mexico. Further, different communities receive this shock at different times throughout the year as determined by the feast day associated with each parish's patron saint. By exploiting the quasi-random variation in the timing of fiestas across municipalities and by exploiting an exogenous change in Mexico's electoral calendar, we obtain unbiased estimates of the effect of community activity on turnout. Moreover, because saint's day fiestas are exactly the types of prescriptions recommended by social capital scholars, our results speak the potential of social capital as a policy prescription for increased democratic performance.

We are aware of only one other study that employs a quasi-experimental approach to study social capital and turnout. Condon (2009) conducts an ongoing study of U.S. elementary schools randomly assigned into the FAST (Families and Schools Together) program. Parents of students in the FAST schools are encouraged to become more involved in their child's school. Condon estimates a negative effect of the program on turnout. Parents assigned to the FAST program were less likely to vote than the control parents. Unfortunately, the way in which subjects were recruited for the study led to significant pre-treatment differences between parents in the two comparison groups. Specifically, parents in the FAST program were typically poorer and less likely to vote before the study which could generate biased estimates. We admire this approach and exploit a separate quasi-experiment to address the effects of social capital in a different political setting.

68

Saint's Day Fiestas in Mexico

The generation of social capital is not a formal process. Rather, citizens become connected to one another by coming together, engaging in casual conversation, eating, drinking, and having fun. In their review of communities which succeed in generating social capital, Putnam and Feldstein (2003) identify dinner parties, picnics, music, local art, and dancing as important sources of community engagement and trust. In one specific example, the authors argue that a multicultural festival helped the Dudley Street Neighborhood Initiative to build social capital in a previously deteriorating Boston neighborhood: "At countless community meetings, at the multicultural festival, through hard side-by-side labor, they [the initiative] helped people connect and reconnect." According to Putnam and Feldstein, social capital arises from casual community interaction, and it can arise quickly over the course of days or weeks. By this account, saint's day fiestas in Mexico generate social capital in abundance. These are exactly the interventions that scholars would prescribe to increase the social capital and subsequent democratic performance of a community.

Saint's day fiestas offer a unique opportunity to test the effects of community activity and the subsequent social capital for the following reasons. Roman Catholic churches are predominant throughout Mexico, and each church or parish has a patron saint. Each patron saint has a particular feast date, typically the day of the year that the saint died. Around the feast date of a particular church's patron saint, the members of the church community hold a large festival, celebrating their saint and their community. In most Mexican communities, these saint's day fiestas are the biggest social event of the year, comparable to if not bigger than the celebrations coinciding with Easter and Christmas.

Our subsequent analysis makes three assumptions about saint's day fiestas in Mexico. First, the time of year at which each parish celebrates its fiesta is exogenous to other features of the communities. For example, a community that celebrates its fiesta in January is on average no

different from one that celebrates its fiesta in July.[2] Second, these fiestas temporarily increase the social capital of the community. For several weeks leading up to and following the fiesta, the community experiences an increase in social connectedness, sense of belonging, trust in the community, and discussion of important issues in the community. Third, if these fiestas do in fact influence voter turnout, we can attribute this effect to the boost of community activity generated by the fiestas. The following sections provide empirical support for these assumptions.

Fiesta Timing is Quasi-Random

To initially assess whether fiesta dates are quasi-random, we surveyed 14 Catholic priests and officials in Mexico whose e-mail address were listed in online church directories.[3] These individuals are not representative of all Catholic communities in Mexico, but they provide general, qualitative support for our assumptions about when fiestas occur and the extent to which they increase social capital.

The patron saint of each church is typically chosen for historical or idiosyncratic reasons. Moreover, the particular fiesta date for each saint is arbitrary, typically the day of the year that the saint died centuries ago. In our survey, we asked respondents how their particular parish chose its patron saint. No respondent indicated that the time of year for the fiesta was considered in this decision. Rather, patron saints resulted from idiosyncratic events or the preferences of one

---

[2] Later in the paper, we present an estimate which is not sensitive to this assumption. By exploiting an exogenous change in the election date, we test for the effects of increased social capital within municipalities.

[3] E-mail addresses were obtained from online directories of the 68 dioceses and 18 archdioceses in Mexico. Links to each diocese website are located at http://www.cem.org.mx/diocesis/.

particular priest or bishop. For example, one respondent from a church called "Our Lady of Refuge" provided the following account: "The people of God were consulted with the approval of the bishop. Here in Tamaulipas, there is great devotion to Our Lady of Refuge because we were officially put under the patronage of Our Lady of Refuge by the Spanish royalty during colonial times."[4] In this case, as in many other cases, Spanish colonizers chose the patron saint of the community for arbitrary reasons. These stories indicate that the fiesta date of a particular parish is exogenous to the characteristics of the community. Later in the paper, we provide further empirical evidence that the fiesta dates are exogenous. Demographic variables are uncorrelated with the fiesta dates, and communities with fiestas occurring near an election are no different from others in terms of their economic and demographic characteristics.

Fiestas Provide a Temporary Shock to Social Capital

To discern whether saint's day fiestas increase the short-term connectedness of their communities, we asked a series of survey questions regarding the nature of these fiestas. The priests and officials indicated that their fiestas last anywhere from 1 to 10 days, require 8 to 30+ days of preparation, and involve 400 to 5000 attendants. When asked about the types of activities at the fiestas, respondents listed numerous social and religious activities including eating, dancing, theatrical performances, wheelbarrow races, egg tosses, lotteries, singing, musical performances, mass, communion, confession, and religious processions. We expected that priests and religious officials would focus on the religious aspects of the event. However, more than half of the activities mentioned were social and secular. Several respondents specifically mentioned that coexistence of neighbors is a primary component of the fiestas.

---

[4] All responses have been translated from Spanish.

When we asked more specifically whether fiesta attendants discussed important political issues, respondents indicated that community members discuss municipal administration, public safety, unemployment, and the performance of political leaders at the fiestas. Finally, we asked whether the fiesta helps to build trust within the community. All respondents confirmed that this was the case. Community trust and discussion of important issues are critical elements of capital, and fiestas are successful in fostering these phenomena. Saint's day fiestas bring members of a community together. By preparing food, singing, dancing, and discussing important issues, the citizens are temporarily raising the connectedness of the entire community. As one clergy-member remarked, "Saint's day fiestas are a means to increase communion between the faithful."

Lastra, Sherzer, and Sherzer (2009) conducted an in depth ethnographic study of two saint's day fiestas in Central Mexico. In their descriptive account, the authors note the significance of these events for community connectedness: "The event reinforces peoples' sense of community. . . . It is striking that all the events of the patron saint fiesta are group activities that require collaboration. . . . Every year the collaboration necessary for the complex organization of the fiesta reaffirms its [the community's] social and ritual structure. . . . One aspect of sociability that is prevalent during fiestas is the cultural theme of accompanying, that is, or being with, sharing the moment with, friends, *compadres*, and the saint" (pp. 116-117). This account affirms our claim that saint's day fiestas temporarily increase a community's capital.

The account of Lastra, Sherzer, and Sherzer and our survey responses suggest that saint's day fiestas provide quasi-random shocks to social capital. Every community receives this positive shock at some point throughout the year, but the particular time of year is essentially random. We focus our study on Mexican municipalities with one Catholic church. On average, municipalities which celebrate their fiesta around an election should be no different than those that do not, except for the timing of their fiesta date. Both sets of municipalities have the same general level of social

activity, but they receive these shocks to connectedness and trust at different times. We exploit this quasi-random variation to estimate the effects of community activity and the subsequent boost of social capital on voter turnout.

Some scholars assume that social capital is a long-term characteristic of a community that cannot vary over short periods of time, but according to Putnam himself, this view is a misconception. "While some early work was understood to imply that stocks of social capital were immutable except on a time-scale of centuries, we now are beginning to explore ways in which individual behavior and collective choice can have important effects on social capital over even relatively short periods" (2002, p. xxii). According to Putnam, social capital can be increased in the short term, and saint's day fiestas are precisely the types of events prescribed to do so. In fact, if social capital is indeed "immutable except on a time-scale of centuries," then the concept has no policy relevance and can only be useful for historical summary. Our subsequent empirical analysis assesses the consequences of these short term shocks to social capital.

The Effect of Saint's Day Fiestas on Voter Turnout

We have collected census and electoral data for all municipalities across Mexico in which there is only one Catholic church and in which we could confidently identify the patron saint and corresponding feast date of the church.[5] We focus specifically on municipalities with one church because we want to ensure that the saint's day fiesta and corresponding social capital is affecting a

---

[5] Census data were downloaded from the web site of the Instituto Nacional de Estadística y Geografía, the Mexican government agency which administers the census. Electoral data were downloaded from the web site of the Instituto Federal Electoral, the Mexican government agency charged with administering elections and certifying the results.

large proportion of the community that we observe.[6] In total, we examine 325 municipalities across

7 national elections for a total of 2255 observations.[7] Summary statistics for all municipalities are

presented in the Appendix. These communities are predominantly rural, agricultural, low-income,

and Roman Catholic.

Our surveys indicate that saint's day fiestas increase social capital for several weeks before

and after the actual fiesta date. As a result, we code a "treatment" variable, *Fiesta*, which takes a

value of 1 if the fiesta date is within 2 weeks of the election date and 0 if the fiesta date is further

from the election. We will say that a municipality is receiving a shock of community participation if

its fiesta date lies anywhere within the four week window surrounding the election in that particular

year. In the appendix, we show that our subsequent results are not sensitive to the choice of this

arbitrary window size. Our subsequent results would be nearly identical for any cutoff between 3

and 18 days.

Municipalities with fiesta dates after the election are included in our fiesta "treatment",

because fiesta preparations begin weeks before the actual fiesta date. The shock to social capital

begins several weeks before the fiesta date and continues for several weeks afterward. Moreover,

individuals may smoothly allocate their free time (Becker 1965), so an upcoming fiesta will consume

time and affect citizens before it begins. However, our subsequent results are unchanged if we only

include municipalities with a fiesta date before the election in our coding of the *Fiesta* variable. For

---

[6] We identified single parish municipalities by collecting online diocese directories and

determining which municipalities are served by only a single church.

[7] The total number of observations is less than 325*7 = 2275 because turnout data is missing

in 1 case, and 19 cases were dropped because the reported turnout was greater than the voting age

population. Subsequent results are robust to the inclusion of these municipalities.

any of the 7 election years in our data set, there are 52 to 54 municipalities with fiestas occurring within two weeks of the election date.

With these data in hand, we quantitatively test our assumption that fiesta dates are quasi-random and then evaluate whether community activity is a complement or a substitute for political participation.

If fiesta timing is quasi-random, there should be few differences in observable pre-treatment characteristics between municipalities which have fiestas close to and far from the election date. The Appendix presents placebo regressions which test for differences between these two sets of municipalities along numerous demographic, economic, and political variables. There are few meaningful differences, and for the variables where we do see some small differences (earnings, partisan support, and government employment) these differences may result from the downstream effects of the fiesta date and the subsequent shock to voter turnout. Importantly, our subsequent results are robust to the inclusion of any of these covariates as controls in our analysis. Consistent with our prior argument that fiesta dates are quasi-random, this analysis suggests that the municipalities in our comparison groups are truly comparable to one another.

We can also assess the comparability of municipalities in our comparison groups by estimating a propensity score. We estimate a Logit model, regressing the *Fiesta* variable on 60 demographic variables from the census. A municipality's propensity score is its predicted value, representing the *a priori* predicted probability that the municipality would be in the fiesta treatment given its demographic characteristics. The Appendix presents the distribution of propensity scores for both sets of municipalities. The distribution of propensity scores is similar for both groups.

Having established that fiesta timing is quasi-random, we now turn to evaluating the effect of saint's day fiestas on turnout. If community activity and social capital influence voter turnout, we expect turnout to vary with fiesta dates. If this exogenous shock of community participation

increases voter turnout, we should see higher average turnout levels for municipalities with fiestas closer to the election date. Conversely, we should see the opposite trend if this shock to social capital decreases turnout.

We take several approaches to estimating the effect of fiestas on turnout. We begin by presenting a nonparametric approach: kernel regression. We calculate residual turnout for each observation removing variation associated with different election years and the mean turnout levels in each state. Figure 3.1 shows the predicted level of residual turnout relative to the number of weeks that a municipality's fiesta occurs before or after the election. Turnout is significantly lower for municipalities holding a fiesta close to the election date. Several weeks before the election we begin to see the negative effect, and it continues for municipalities holding fiestas several weeks after the election.

Next, we employ two parametric approaches to estimate the effect of fiesta timing on voter turnout. Both results are presented in Table 3.1. First, we estimate the effect of the *Fiesta* treatment by ordinary least squares (Pooled OLS). To improve the precision of the estimates, we include year fixed-effects and state turnout in the model. According to this estimate, a fiesta occurring within two weeks of the election decreases voter turnout by 3.5 percentage points. This estimate, which primarily exploits variation across different municipalities, is statistically ($p < .01$) and substantively significant. In the appendix (Tables A3.3 and A3.4) we show that a simple difference-in-means yields the same result, and the result is robust the exclusion or inclusion of different years or control variables.

Figure 3.1. Kernel Regression: Turnout across Fiesta Dates

Residual turnout removes year and state level variation in turnout. For each municipality-year, we regress turnout on year fixed effects and state level turnout, and then compute the residuals. This graph uses an epanechnikov kernel with bandwidth 3.5, but the result is not sensitive to the choice of kernel or bandwidth. Dotted lines indicate standard errors.

Table 3.1. The Effect of Saint's Day Fiestas on Voter Turnout

	(1) Pooled OLS	(2) Fixed Effects
Fiesta	-3.452	-2.525
	(1.038)**	(1.113)*
Year Fixed Effects	X	X
Municipality Fixed Effects		X
State Turnout	79.726	48.185
	(5.738)**	(6.203)**
Observations	2255	2255
R-squared	.40	.62
SER	11.01	9.37

*Municipality-clustered standard errors in parentheses; * significant at 5%; ** significant at 1%*

The dependent variable is voter turnout, coded as percentage points from 0 to 100. Each observation is a municipality-year between 1991 and 2009. "Fiesta" is a dummy variable indicating whether the municipality's saint's day fiesta occurred within two weeks of the election.

As a second test of the effect of saint's day fiestas, we leverage a shift in the federal election date. For the 1991 and 1994 elections, elections took place in mid-August, but the subsequent five elections took place in early July. Therefore, we can test for changes in voter turnout for individual municipalities which fell into or out of the *Fiesta* treatment as a result of the change in election date. Our second estimate in Table 3.1 includes municipality fixed effects, removing any variation in turnout across different municipalities. As the election date moves within (outside) two weeks of a municipality's fiesta date, turnout decreases (increased) by 2.5 percentage points ($p < .05$).[8] Even if fiesta dates are endogenous to turnout for some unknown reason, this fixed effects estimate would still provide an unbiased estimate because we have removed any variation across different municipalities.

We reject the hypothesis that social capital, along with the type of community activity thought to improve democracy, increases turnout. Rather, a fiesta occurring within two weeks of a federal election will reduce voter turnout in the municipality by 2.5 to 3.5 percentage points. This result is inconsistent with previous correlational observations, suggesting that those analyses are biased upward due to reverse causation and omitted variable bias. While social individuals and groups are more likely to vote, the random assignment of community activity does not increase but rather decreases voter turnout.

---

[8] We conduct a Hausman test (Hausman 1978) comparing our fixed effects estimates to those from a random effects model, and cannot reject the null that the estimands are equal. As a result, we have no evidence that our fixed effects estimate is statistically different from our cross-sectional estimates. Second, we cannot reject the null hypothesis that the independent effect of each municipality is uncorrelated with the treatment variable. This provides further support for our claim that fiesta dates are quasi-random.

Persistence of Turnout across Subsequent Elections

If saint's day fiestas decrease turnout in one election, we expect that such demobilization will continue to have an effect in subsequent elections. Empirical research suggests that voting is habitual (Plutzer 2002; Gerber, Green, and Schachar 2003; Meredith 2009; Davenport et al. 2010), so the decision to abstain from voting in one election will decrease the probability of voting in future elections. To explore this possibility we take a closer look at the switchers, those municipalities which fell into or out of the fiesta treatment over time.

Our cases fall into three categories in regard to their receipt of the fiesa treatment. There are approximately 50 municipalities with August fiesta dates which only experienced the social capital shock in 1991 and 1994. There are approximately 50 municipalities with late June or early July feast dates which only experienced the social capital shock in the five elections following 1994. And there are approximately 220 municipalities with other feast dates which never had a fiesta within two weeks of the election.[9]

Figure 3.2 shows the residual turnout rates for all three of these groups across each election year. We can see that the June/July municipalities voted at the same rate as the "control" municipalities in 1991 and 1994 when they had not yet received the treatment. Then, after the election date moved toward their fiesta date, their turnout rates dropped and remained lower through all subsequent elections. The August municipalities initially began with lower turnout rates because they fiestas coincided with the elections in 1991 and 1994. Following the change in election date, turnout remained low through subsequent years even though these municipalities no longer

---

[9] These numbers are approximate, because the election date does move slightly from year to year. However, the only big change occurred between 1994 and 1997, so we simplify our analysis here to designate three groups of municipalities.

received the social capital shock around the election. The observed trend is consistent with the hypothesis that negative shocks to voter turnout will persist over time.

Looking more closely at the municipalities treated from after 1994, we see further evidence that the fiesta treatment is persistent. Residual turnout is initially similar between the two groups when neither group has a fiesta near the election, but once the election date switches, residual turnout progressively declines with each election. The longer a municipality has had a fiesta near the election, the greater the negative effect. By 2006 when these municipalities have their fourth election around the time of a fiesta, their turnout is 4.5 percentage points lower than their counterparts. The repeated existence of a fiesta around the time of multiple elections has a particularly strong demobilizing effect. These results do not suggest that the social capital generated by a fiesta persists for several years; we saw earlier that the fiestas only increase community participation for several weeks. Rather, the act of voting itself is persistent, so an individual's failure to vote in one election decreases her chances of voting in the next election.

The persistence of the fiesta effect explains why our fixed effects estimate is slightly smaller (closer to zero) than our cross-sectional results. The fixed effects analysis focuses on individual municipalities in which the treatment changed. When the election date changed, the municipalities with August elections were no longer treated. However, the negative effects of their previous treatments persisted, causing turnout to remain low. Given this phenomenon, our fixed effects estimate is likely biased toward zero, suggesting that we should rely more heavily on our cross-sectional results. Alternatively, if we think that the data in Figure 3.2 results from imbalances between the August municipalities and the rest of the sample, then we should rely more heavily on our fixed effects estimate. Both estimates provide similar results, so the interpretation of these patterns is inconsequential for our main empirical result.

Figure 3.2. A Closer Look at the Switchers
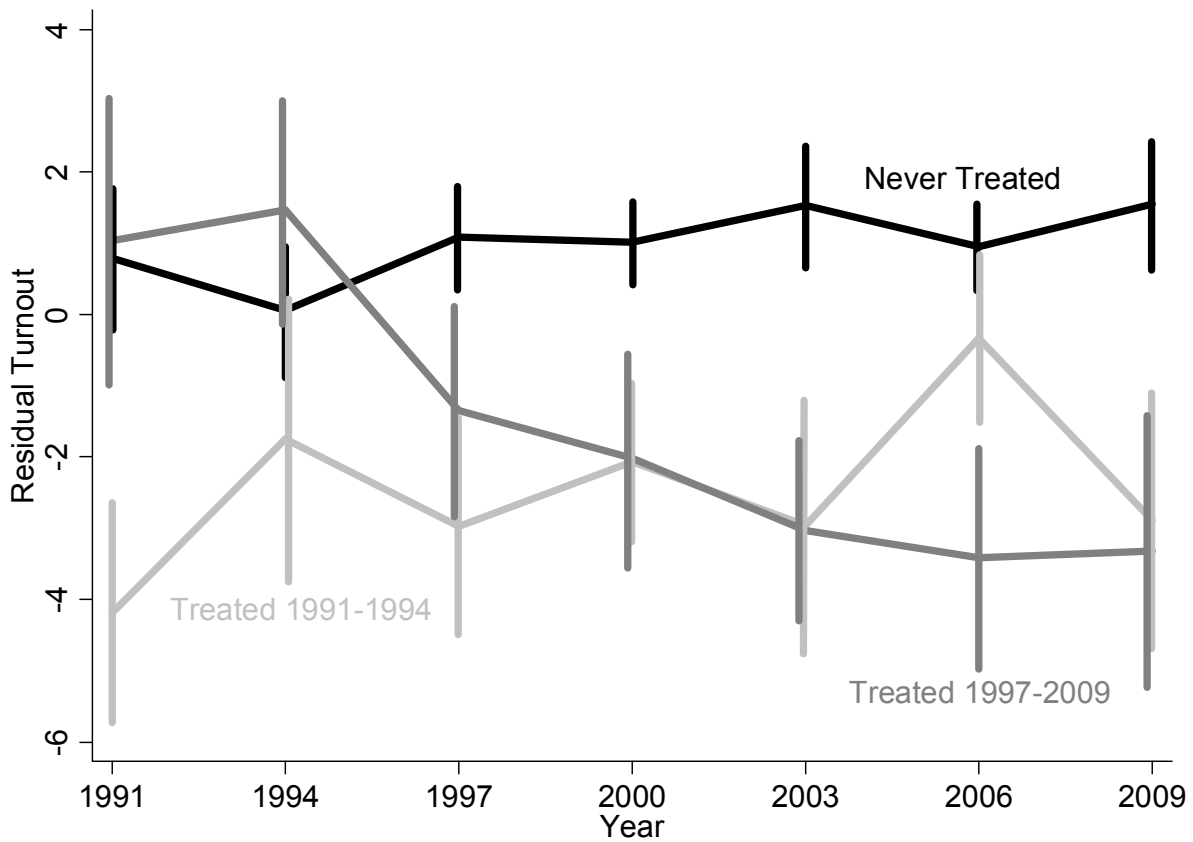
Again, residual turnout removes variation associated with each year and state. The graph presents the average level of residual turnout for three different subsets of Municipalities. Those with fiesta dates in August were "treated" in the 1991 and 1994 elections. Those with fiesta dates in late June/early July were treated in all elections from 1997 to 2009. All remaining municipalities were never treated in any election. The vertical bars represent standard errors.

Variation of the Effect across Municipalities

If saint's day fiestas decrease voter turnout, we expect the effect to vary across different type of municipalities. We modify our pooled OLS model to include interaction terms which assess the conditions under which the fiestas will have a stronger or weaker effect on voter turnout. Table 3.2 shows the results of three regressions which include various interaction terms. In each case, the variables have been re-coded to range from 0 to 1. Therefore, we can interpret the coefficient on an interaction term as the change in the effect of fiestas as we move from municipalities with the lowest level of the explanatory variable to those with the highest level.

First, we expect the effect to be larger in municipalities with a higher percentage of Catholics. In these municipalities, the fiesta will involve a higher proportion of residents and the shocks to social capital will be more intense. Column 1 shows that the effect is 8.3 percentage points greater (more negative) as we move from the municipalities with the lowest proportion of Catholics to those with the highest proportion of Catholics.

Second, the effect should be greater in smaller municipalities for several reasons. Smaller communities with fewer social alternatives will attract a higher proportion of residents to the fiesta. Also, we have restricted our analysis to municipalities with only one church. Larger municipalities may have other churches that we are unaware of. If this is the case, our effect will be diluted in these larger municipalities because a smaller proportion of residents will attend the fiesta. Column 2 indicates that the effect of fiesta timing is strongest for smaller municipalities. Saint's day fiestas actually decrease turnout by 21 percent for the smallest municipalities in our data set.[10] In a

---

[10] Normally, it is tricky to directly interpret the coefficients in interactive models. However, since we have coded the fiesta treatment and log population to range from 0 to 1, we can interpret the coefficient on *Fiesta* in column 2 as the effect of the fiesta treatment for the smallest communities in our data set.

Table 3.2. Variance of the Fiesta Effect across Municipalities

	(1)	(2)	(3)
Fiesta	3.354	-21.183	-3.179
	(4.324)	(4.082)**	(1.041)**
Fiesta * Catholic	-8.288		
	(5.219)		
Fiesta * Log Population		27.887	
		(5.885)**	
Fiesta * Local Election			-0.537
			(3.090)
Catholic	7.074		
	(2.888)*		
Log Population		-15.925	
		(2.885)**	
Local Election			11.14
			(1.351)**
Year Fixed Effects	X	X	X
State Turnout	79.15	66.073	70.64
	(5.272)**	(5.088)**	(5.207)**
Observations	2255	2255	2255
R-squared	.40	.43	.43
SER	11.84	11.57	11.58

Municipality-clustered standard errors in parentheses; * significant at 5%; ** significant at 1%

The dependent variable is voter turnout, coded as percentage points from 0 to 100. Each observation is a municipality-election between 1991 and 2009. "Fiesta" is a dummy variable indicating whether the municipality's saint's day fiesta occurred within two weeks of the election. "Local Election" is a dummy variable indicating whether the federal election coincided with a state or local election in that particular municipality. The variables "Catholic" and "Log Pop" are coded to range from 0 to 1. Therefore, the interaction terms can be interpreted as the change in the treatment effect as we move from the smallest (least Catholic) to the largest (most Catholic) municipalities in the data set.

community of 400 to 500 residents, 1 in 5 individuals are demobilized by a fiesta occurring near the election date.

Lastly, we test whether the effect is different if a local election happens to coincide with the federal election. Turnout is typically higher in local elections compared to federal elections, and citizens are more likely to care about the local races. In column 3, we do find that turnout is significantly higher when a local election coincides with the federal election, but the demobilizing effect of saint's day fiestas is unchanged. Even in elections where citizens are highly interested, the demobilizing effect of community participation remains.

Replicating the Findings in an Urban Setting

Our previous analysis has focused solely on Mexican municipalities with just one Catholic Church. Thus, our data set consists of primarily poor, rural, agricultural communities. The questions remain whether our results will generalize to other democratic communities. In order to ensure internal validity, we have limited our study to the subset of regions for which we can make valid inferences. To test for external validity, we test the effect of saint's day fiestas on turnout in an urban center, Monterrey. This analysis, presented in the Appendix, lacks the precision of our previous analysis. Nonetheless, we estimate a similar effect. Saint's day fiestas decrease turnout in Monterey by approximately 3 percentage points, just as they do in rural municipalities. Therefore, the same mechanisms that lead community participation to decrease turnout in rural municipalities are present in urban settings as well.

Testing Alternative Explanations

We have presented evidence that the occurrence of a saint's day fiesta near a federal election significantly decreases voter turnout in a community, and we have argued that this decrease is

causally attributable to the community activity and subsequent social capital generated by fiestas. In this section, we consider alternatives to our causal interpretation. Challenges to our findings will likely come in two forms. First, there may be unobserved differences between municipalities in our comparison groups that confound our results. Second, the observed effect may driven by some extraneous factor that also changes during the fiestas.

We have already attempted to rule out the first challenge regarding the comparability of municipalities with fiesta dates near and far from the election date. The placebo regressions and distribution of propensity scores demonstrate that there are few observable differences between these groups. Moreover, we obtain a similar result employing municipality fixed effects which is not sensitive to differences between municipalities. All of these results are consistent with our claim that fiesta dates are quasi-random. However, more subtle alternative explanations are still possible.

Catholic parishes choose their patron saints for many reasons. Perhaps certain characteristics are correlated with the type of patron saint that a church will select. Since some saints are quite common, one may worry that the results are driven by the churches of one or a few particular saints. There are 7 different fiesta dates that are shared by more than 15 churches in our data set. Our results are robust to the exclusion of any of these sets of churches. Another way to address this concern is to measure the degree of similarity within churches sharing a fiesta date. We find that the variance of census covariates among municipalities sharing a fiesta date is the same as the variance of covariates across all municipalities. Therefore, all evidence is consistent with our argument that the fiesta dates are quasi-random.

Another possibility is that certain types of municipalities will choose patron saints whose feast date occurs at a convenient time. Perhaps the family members of college students or migrant workers will prefer to have their fiestas during the summer or around Christmas so that their relatives can return home during the celebration. We find that municipalities with summer or

Christmas fiesta dates are demographically similar to all other municipalities. However, as previously discussed, even if there were some unobserved variable correlated with fiesta dates, our fixed effects estimate would still provide an unbiased estimate because it removes differences across municipalities.

We assume that saint's day fiestas influence social capital, and the way in which the fiestas affect voter turnout is through both community activity and the subsequent social capital generated by the fiestas. Since we cannot obtain a precise measure of community participation, we estimate the effect of fiestas on turnout. If fiestas influence turnout through some mechanism other than community participation, we could draw a misleading conclusion about the relationship between social capital and turnout. While we cannot prove the validity of this assumption, we can raise alternative possibilities and assess their plausibility.

As we have discussed, Saints Day Fiestas are large events spanning several days and requiring multiple weeks of preparation. Is there something other than social capital that changes during the fiesta that might explain our result? One possibility is that intense celebration and alcohol consumption exhaust citizens. On one level, exhaustion is not inherently at odds with social capital. We want to test the effect of citizens coming together and connecting. If they happen to consume alcohol or lose sleep when they convene, that would be one byproduct of social engagement. However, our results cannot be attributed solely to hangovers or sleeping in, because we look at a four week window. Tiredness might influence turnout if the fiesta occurs one or two days before the election. However, we find that fiestas occurring two weeks before or after the election date decrease turnout in the same way as those occurring the day before the election. The fact that the effect spans across a long time period both before *and* after the election rules out the possibility that our effect could be entirely driven by exhaustion, hangovers, or other factors which would not span this same time period.

Another possibility is that political candidates or government officials would visit the festivals in order to influence the community's political participation. Since turnout decreases during the fiestas, this type of political activity would have to decrease turnout in order to work against our conclusions. Perhaps citizens are disillusioned by political campaigning at the festivals or government officials actually attempt to decrease turnout in communities that oppose them. Both possibilities are unlikely, particularly for the small, rural municipalities in our data set. Additionally, our survey respondents indicated that there is never any official political or government activity at their fiestas. We can also rule out the possibility that voter registration plays a role because the deadline to register is several months before the election date.

A related explanation is that saint's day fiestas interfere with the campaign activity that would typically mobilize voters. Again, this story may not be at odds with our claim if community activity and social capital in general tend to interfere with campaign activity. This campaign effect would be part of the downstream effect of community participation. Nonetheless, altered campaign activity cannot entirely explain our results, because of the duration of our effect. We would not expect a fiesta *after* the election to interfere with campaign activity in the same way that a fiesta before the election might. Again, the long time period across which we see an effect rules out the possibility that our results are driven by a short-term factor such as altered campaign activity before the election. If anything, the occurrence of the fiesta should make it easier for campaigns to reach citizens and mobilize citizens, but we do not observe this phenomenon.

Since saint's day fiestas have an inherently religious purpose, the fiestas may alter the community's level of personal religiosity which in turn influences turnout. Ten of 12 survey respondents indicated that church attendance increases during the fiestas, so this hypothesis is plausible. However, religious participation is thought to be both a source and consequence of social capital (Putnam 2000; Putnam and Campbell 2010), so this possibility is not at odds with our

conclusions. In fact, we take this increased religiosity as supporting evidence that fiestas temporarily increase social connectedness and trust within a community. Moreover, evidence suggests that religious activity may causally increase, not decrease, voter turnout (Gerber, Gruber, and Hungerman 2008), so we suspect that our results cannot be entirely attributed to religiosity.

Discussion and Conclusion

Social scientists have long considered vibrant civic political associations a basic requisite of democracy. Lipset (1959) claims, "In a large complex society, the body of the citizenry is unable to affect the policies of the state. If citizens do not belong to politically relevant groups, if they are atomized, the controllers of the central power apparatus will completely dominate the society." More recently, social capital scholars have argued that non-political civic associations promote participation and effective governance (Putnam 1993, 1995, 2000; Tavits 2006). This study focuses on the hypothesis regarding social capital and political participation.

We exploit a natural experiment to test the effect of participation in local community festivals on voter turnout. Saint's day fiestas bring together individuals in a community, allowing them to connect with one another and discuss important issues. Contrary to previous theories and observational findings, this exogenous shock to community activity around the time of a federal election actually decreases voter turnout. To be clear, we do not conclude that social capital on its own depresses political participation. Rather, the types of community activities which increase social capital can have adverse consequences. As these activities increase trust and connectedness, they also distract citizens from the political process and decrease participation. This finding is not obvious. In fact, when we described our design to other researchers, many predicted that turnout would increase and none expected that it would decrease.

While the results presented in this paper undermine one of the most important social capital hypotheses, our findings are not a general indictment of the role of civic engagement in democratic governance and collective behavior. The mechanisms connecting citizen behavior and quality of governance are manifold and complex. The direct effect of community activity on turnout is only one component. There are other ways that community engagement could improve the quality of democracy. For example, governments might be more responsive to a socially-connected citizenry. Also, our analysis takes advantage short-term shocks to social capital, but we cannot say whether the development of long-lasting social capital would have the same effect. Perhaps the temporary shock of a community festival decreases turnout but the long-term social capital necessary to hold regular community festivals is actually beneficial for participation. Nonetheless, we provide the best available test of the hypothesis that social capital will increase turnout and representation, and we find the opposite. Moreover, any policy prescription which aims to increase social capital will likely generate this short-term variety of connectedness, so a test of the consequences of short-term social capital is warranted.

How can we reconcile our findings with previous observations? In virtually every case, measures of social capital are correlated with political participation. However, these correlations may be driven by confounding variables or reverse causation, in which case they tell us nothing about the causal effect of community activity.

Several previous studies are consistent with our finding that community participation decreases turnout. As previously mentioned, Condon (2009) finds that parents assigned to the FAST (Families and Schools Together) program are less likely to vote. Additionally, Stoker and Jennings (1995) find that young couples experience a decrease in political participation around the time of their weddings. Marriage ceremonies bring families and friends together like no other event, and despite the elevated level of social activity, turnout decreases. We present a similar result but

90

hope to overcome the methodological limitations of previous studies by exploiting the quasi-random timing of saint's day fiestas.

Community festivals do not always detract from voter turnout. Addonizio, Green, and Glaser (2007) find that festivals held on the election date near the polling location can increase voter turnout by several percentage points. Why do festivals increase turnout in their context and decrease it in ours? The key difference is that their festivals were specifically designed to attract voters to the polls. In a sense, their festivals lower the cost of voting, which increases turnout as expected. In this case, fiestas do not happen on the exact election date (except in a few cases). Instead, saint's day fiestas bring individuals together at a time and location removed from voting, raising social capital without directly altering the cost of voting. Their result suggests that planned festivals can raise turnout by bringing people to the polls, but their finding says nothing about the direct effects of social capital.

How can a positive shock to social capital decrease voter turnout? We have proposed three ways in which heightened social connectedness might detract from political participation. All three possibilities are plausible, and we suspect that all of them are at work in this context. First, time is an essential resource for political participation (Verba, Schlozman, and Brady 1995) which is consumed by social capital. Rational individuals will smoothly allocate their free time (Becker 1965) such that community activity will decrease political participation even when the community event occurs weeks before or after an election. As citizens become more involved with the community, they have less time to learn about the election, form an informed opinion, and visit the polling place.

Second, social interactions present conflicting views to potential voters which might create uncertainty (Mutz 2002). Our surveys indicate that citizens talk about contentious issues at the fiestas and express discontent with political leaders and the political system. Moreover, Joel and Dina Sherzer, two authors of *Adoring the Saints: Fiestas in Central Mexico*, told us that political

discontent is pervasive throughout the fiestas. "If you look at the kinds of figures, giant puppets, etc. which parade about during fiestas, they are sardonic, politically biting, and humorously critical of government leaders."[11] The exposure to conflicting views may cause members of the community to lose confidence in their own political opinions or become disillusioned with politics altogether.

Lastly, citizens often vote because of the sense of civic duty or the fulfillment they derive from the act itself (Riker and Ordeshook 1968). Social capital may serve as a substitute for that type of satisfaction. When citizens contribute to their community at the saint's day fiesta, they may no longer feel the need to vote, because they have already achieved the fulfillment that they would otherwise obtain through voting. Lastra, Sherzer, and Sherzer (2009) argue that this sense of duty explains the high levels of participation in saint's day fiestas. "These networks of relationships are effective because of ethical principles that govern the behavior of the inhabitants of the communities. They feel a sense of duty . . . to carry out one's promise" (p. 116).

We are among many political scientists concerned with low voter turnout and inequality in the political process (e.g. Verba, Schlozman, and Brady 1995). Socioeconomic status is highly correlated with voter turnout, which may bias public policies in favor of the few. Many scholars and activists hope that increased community participation and social capital will open the doors of political representation for underrepresented communities like those in our analysis. If social connectedness does increase turnout, then we can improve representation by building community centers, opening parks, and throwing community-wide festivals. However, we find no evidence that such proposals can solve the collective action problem of turnout. To the extent that social capital is correlated with turnout, it is likely a byproduct, not a cause, of a healthy democracy. Contrary to previous thinking, community activity is a substitute, rather than a complement, for political participation.

---

[11] This quote was taken from e-mail correspondence with Joel and Dina Sherzer.

Study #4

Pivotality and Turnout: Evidence from a Field Experiment in the Aftermath of a Tied Elections

Co-authored with Ryan D. Enos

Abstract

Many democratic citizens habitually abstain from the political process, and the reasons for this abstention are of great interest to scholars, campaigns, activists, and policymakers. Most social scientists and political pundits assume that greater electoral competition and the increased chance of pivotality will motivate citizens to turn out to the polls. First, analyzing observational and survey evidence, we find little support for this claim. Then exploiting the rare opportunity of a tied election for major political office, we conduct a large-scale field experiment. Informing citizens that an upcoming election will be close has little mobilizing effect. To the extent that we do detect an effect of electoral competition on turnout, it is concentrated among a small set of frequent voters. Our evidence suggests that increased electoral competition is not a solution to low turnout, and the predominant models of turnout which focus on pivotality are of little practical use.

Electoral competition and political participation are essential components of a healthy democracy (Verba, Schlozman, and Brady 1995), and most scholars argue that these two factors are complementary. In this paper, we exploit an extremely rare opportunity to test the important claim that electoral competition increases citizen participation and we find little support for this notion. Using multiple data sources and an unprecedented experiment, we subject the individual calculus of voting to the most thorough test yet, allowing us to rigorously test the impact of electoral competition on individual behavior in the real world. .

Many American citizens abstain from the political process. Since reliable data has become available, at least 30% of eligible citizens have failed to vote in any given presidential election (McDonald and Popkin 2000). Moreover, voting is a habitual act (Verba and Nie 1972; Miller and Shanks 1996; Green and Shachar 2000; Plutzer 2002; Gerber, Green, and Shachar 2003; Meredith 2009), and habitual voters are unrepresentative of the wider population (Verba and Nie 1972; Verba et al. 1995; Wolfinger and Rosenstone 1980), so low turnout has significant partisan and policy consequences (Anzia 2011; Citrin et al. 2003; Fowler 2013; Hajnal and Trounstine 2005; Hansford and Gomez 2010; Hill 2010; Knack and White 1998; Martinez and Gill 2005).

Dominant theories of voter participation predict that close electoral competition will increase participation (Riker and Ordeshook 1968), and previous scholars have even offered competition as a cure for unequal participation (Blais, Young, and Lapp 2000). In this paper, we thoroughly review the available evidence and find no convincing causal evidence for a relationship between the closeness of an election and voter turnout. Then we conduct a field experiment in the aftermath of a tied election to test the effect of priming and informing voters about an upcoming close election. Again, we find no consistent evidence that the closeness of elections spurs higher turnout or greater equality in participation. If anything, the closeness of elections may exacerbate the gaps in participation between groups by primarily mobilizing high-frequency voters. If policy

makers hope to increase participation among under-represented groups or the general population, they should focus on the costs of voting, civic duty, and the non-political benefits of voting because electoral competition appears to have little effect. Moreover, the concept of pivotality may hold little practical use to scholars who hope to model and understand the calculus of voting.

The closeness of elections may theoretically lead to increased turnout through two different mechanisms: individual voter psychology or the heightened activity of campaigns and media associated with close elections (Blais 2000, Cox and Munger, 1989). Here we focus on the former. The effects of media and campaigns on voter turnout are thoroughly studied. However, for a variety of reasons that we will explore in this paper, we still do not understand the effect of electoral closeness on voter psychology. Do close elections increase the individual incentive to participate? Isolating the effect of closeness on voter psychology is paramount for both modeling the calculus of voting and also determining how public policies can increase voter turnout. If turnout is strongly influenced by media and campaign functions, these activities can, of course, be increased in the absence of close elections. However, if close elections increase turnout by convincing the voter that her vote is more valuable, then pivotality plays an important role in the calculus of voting and structural reforms that encourage closer elections may be a fruitful policy solution to low participation. In this paper, we undertake a significant test of the psychological mechanism of the effect of close elections.

We proceed by examining the role of electoral competition in traditional models of voter participation. We examine and retest previous attempts to empirically establish a causal relationship between competition and turnout. Then, we provide our own experimental test of this relationship following a fortuitous tied election. Finally, we conclude by arguing for an altered focus in voter participation research.

Pivotality and Turnout

In their classic theory, Riker and Ordeshook (1968) attempt to solve the paradox of voter turnout (Downs 1957) by asserting that citizens receive utility from the act of voting itself. In their well-known model, an individual's decision to vote is influenced by four factors: her probability of casting a decisive vote (P), the value the she would derive from her preferred candidate winning over that candidate's closest competitor (B), her utility derived from the act of voting itself (D), and the cost of voting (C). Therefore, a citizen's returns from voting (R) can be written as follows:

(1) $$R = PB + D - C,$$

and a citizen will vote when

(2) $$PB + D > C.$$

This model provides the basis for much of the current understanding of the decision to vote. When researchers, pundits, and practitioners attempt to explain turnout, they often refer to P and B, assuming that they are as important if not more important than C and D. For example, after New York state ranked last in voter turnout in 2010, *The New York Times* asked political scientist Michael McDonald to diagnose the problem. "Mostly, I suspect that the uncompetitive elections is the main cause."[12] However P is usually infinitesimal (Gelman, King, and Boscardin 1998), so changes to P should have little effect on turnout. As Schwartz (1987) aptly points out, "Saying that closeness increases the probability of being pivotal . . . is like saying that tall men are more likely than short men to bump their heads on the moon."

Nevertheless, the widespread belief among political scientists is that electoral competition, and thus higher P, causally influences voter turnout. Classic texts have asserted this relationship: V.

---

[12] Sam Roberts. "New York State Ranks Last for Voter Turnout."*The New York Times*. November 16, 2010.

O. Key (1949) attributed low turnout in the South during the early 20th century to the lack of electoral competition in general elections and Anthony Downs (1957) famously stated that citizens have no rational reason to turn out in a one-sided election. Since then, numerous scholars have argued that electoral competition is the cure for low turnout (Kelley, Ayres, and Bowen 1967; Teixeira 1992; Wattenberg 2002; Franklin 2008) or the explanation for trends in turnout over time (Burnham 1969; Piven and Cloward 2000; Franklin 2004).

The theoretical strength and intuitive appeal of this model is demonstrated by the number of scholars who assume that electoral competition influences voter psychology and subsequently voter turnout. To provide a recent example of this widespread assumption, Arceneaux and Nickerson (2009) state that "When the race is close and many people care about the outcome, more people decide to vote relative to races in which general interest is low."

Of course, not all researchers assume a central importance of *P*.   For example, Gerber, Green, and Larimer (2008) note that "Because the probability of casting a decisive vote in an election is typically infinitesimal, the calculus of voting boils down to the relative weight of *C* and *D*." But the widespread assumption remains that the closeness of an election influences turnout. Andre Blais writes that election closeness increasing turnout "is the most firmly established result in the literature. I cannot see how this finding could be wrong" (2006, p. 119).

To illustrate the dominance of this assumption, we reviewed all articles published on voter turnout since 1980 which appear in five leading political science journals: *American Journal of Political Science*, *American Political Science Review*, *British Journal of Political Science*, *Journal of Politics,* and *Quarterly Journal of Political Science*.[13] An article was included if the title indicated the study of voter turnout. After reading the papers, a research assistant recorded any clear appeals to *P*. There were 70 papers published since 1980 in these five journals which addressed the causes of voter turnout. Of those,

---

[13] *QJPS* begin publication in 2006, so the count for that journal begins in that year.

more than half, 41, made a clear appeal to the importance of *P* or electoral competition. Political scientists continue to put electoral closeness at the center of models of turnout.

Pivotality and Equality

Inequality in political representation is a topic of significant scholarly attention (APSA Task Force 2004; Bartels 2008; Dahl 2006; Gilens 2012; Schlozman, Verba, and Brady 2012; Verba, Schlozman, and Brady 1995). While not the focus of this paper, inequality in representation has received significant attention in the literature on voter participation and our experiment allows for a rare opportunity to test some claims from previous scholars, so we consider the topic briefly here. In addition to claims that pivotality will increase average levels of turnout, scholars have also argued that close elections and increased pivotality will reduce disparities between voters and nonvoters, thus improving equality in political representation (e.g. Blais, Young, and Lapp 2000; Key 1949; Piven and Cloward 2000; Schattschneider 1960; Franklin 2004). Given the importance of this claim, we briefly address this question from both a theoretical and empirical standpoint.

Even if pivotality increases average levels of turnout, it will not necessarily lead to greater equality in turnout. For example, get-out-the-vote interventions which significantly increase average turnout also increase the disparities between voters and nonvoters by primarily mobilizing more of the types of individuals that were turning out anyway (Enos, Fowler, and Vavreck 2012). Moreover, the canonical theory of the calculus of voting does not imply that changes to *P* will affect all citizens equally. Specifically, *P* should only influence those citizens who have a large enough value of *B* and for those where $D - C$ is barely less than zero. For most under-represented citizens, *B* is low and $D - C$ is significantly less than zero, so *P* can have no impact on their decision to turn out. Conversely, the small subset of voters who can be influenced by *P* (high value of *B* and *D* almost equal to *C*) are likely to be demographically similar to the population of voters. Therefore, from a theoretical

standpoint, we would not expect increases in pivotality to improve equality in political representation; if anything, it could have the opposite effect. Later in this paper, we explicitly test whether our field experiment appeared to exacerbate or reduce inequalities in voter turnout. Consistent with theory, we find that informing voters about a high value of P primarily mobilizes a small subset of citizens who have voted frequently in the past, thereby exacerbating inequalities in participation.

Previous Evidence on Pivotality and Turnout

Despite the widespread belief that electoral competition increases voter turnout, the empirical evidence is weak. The existing literature asserts the importance of close elections and pivotality but almost all previous studies suffer from one or more shortcoming. Observational studies suffer from the confounding of close elections or individual perceptions of closeness with other, unmeasured variables. Experimental studies, whether in the laboratory or field, lack external validity. In short, despite the heavy scholarly attention to P, existing studies have yet to find a satisfactory way to measure the influence of close elections on voters. Using a multi-method approach and harnessing unique and original data, we overcome the limitations of previous studies and address the direct influence of closeness on behavior.

While many studies report higher turnout during close elections (Barzel and Silberberg 1973; Kim et al. 1975) other studies find no such correlation (Ferejohn and Fiorina 1975; Matsusaka 1993). More importantly, even if electoral competitiveness is correlated with turnout, these measures tell us little about the causal effect of closeness on turnout through individual voter psychology. Campaign activity (Patterson and Caldiera 1983; Schachar and Nalebuff 1999; Hill and McKee 2005, Gimpel, Kaufmann, and Pearson-Merkowitz 2007), elite mobilization (Cox and Munger 1989), campaign donations (Ansolabehere and Snyder 2000; Erikson and Palfrey 2000), and media coverage (Clarke

and Evans 1983; Jackson 1996) are significantly greater during competitive elections which may increase turnout even if individual citizens are unaware of or uninfluenced by the closeness of the election. Moreover, higher turnout may lead to increased electoral competitiveness instead of competitiveness driving turnout. In fact, Panagopoulos and Green (2008) demonstrate that experimental increases in turnout may lead to tighter election outcomes. Both of these confounding factors will lead the already weak correlations between competitiveness and turnout to overstate the true effect of P.

Some studies attempt to assess the independent psychological effect of electoral closeness over and above campaign activity by including closeness and campaign spending in the same model. The effect of closeness in these models is small (Seidle and Miller 1976; Tucker 1986; Cox and Munger 1989; Berch 1993), mixed (Matsusaka and Palda 1993), or zero (Jackson 1996), and this multivariate approach does not overcome the problems of endogeneity. First, this approach does not account for the possibility of reverse causation, as discussed above. Second, the addition of control variables is unlikely to account for all differences between close and uncompetitive elections.

In addition to these aggregate observational studies, researchers have attempted to assess the effects of close elections at the individual level using survey data. Respondents who think an upcoming election will be close are more likely to vote (Riker and Ordeshook 1968; Aldrich 1976; Blais et al. 2000), suggesting that a high perception of P may drive citizens to the polls. Again, however, these correlations likely overestimate the true causal effect of electoral competition on turnout. Omitted variables may influence both turnout and respondents' propensity to say an election will be close. For example, campaign activity increases citizens' perceptions of competition (Bowler and Donovan 2011). Moreover, the act of voting itself may lead voters to increase their perceptions of closeness to avoid cognitive dissonance. Since the act of voting is costly, voters might convince themselves that their vote is meaningful in order to justify the expense. For these

reasons, we would expect a positive correlation between citizens' perceptions and turnout even if electoral competition does not drive turnout.

Lab experiments provide an additional opportunity to test for the effects of close elections. Ansolabehere and Iyengar (1994) randomly presented prospective voters with different poll results before the 1992 general election. However, subjects receiving news of a competitive race were no more likely to express an intention to vote than those receiving news of a one-sided race. Later lab experiments report that closeness increases turnout (Levine and Palfrey 2007; Duffy and Tavits 2008), but these results arise from artificial, non-political settings where the financial costs and benefits of casting a "vote" are explicitly laid out for subjects. These results show that experimental subjects can make arithmetic calculations and respond to clear financial incentives, but they suffer from a lack of external validity when speaking to causes of voter turnout in real-world elections.[14]

Lastly, field experiments provide a final opportunity to test for the effects of close elections. Experimental treatments telling prospective voters that an upcoming election will be close are equally if not less effective than other treatments in mobilizing voters (Gerber and Green 2000; Bennion 2005; Dale and Strauss 2009). These results suggest that priming electoral competition is not particularly effective in generating turnout, but they should be interpreted cautiously. In one study (Bennion 2005), the "close election" treatment was paired against a "civic duty" treatment, and both treatments may have their own independent effect. In the other two studies, the experiments took place in non-competitive elections, so the "close election" treatments may have lacked credibility among subjects. In this respect, field experiments suffer from problems of external

---

[14] Recent lab experiments by Blais, et al (2011), which arguably are designed to better approximate voter decision made during actual elections, have shown that most subjects *do not* abstain when it is rational for them to do so.

validity similar to those of lab experiments; we do not know how voters would have responded to these interventions in the real-world situation of a close election.

The ideal opportunity to test for the effect of electoral closeness on voter psychology arises when an election is known to be close beforehand. In this circumstance, a field experiment could credibly inform citizens of the close election and test for the effects of this information. To our knowledge, the best opportunity in this regard is a special election held after an exact tie. Later in the paper, we exploit precisely this rare opportunity. First, we turn to a reassessment and extension of the existing observational data to see if any evidence can be found for an effect of closeness on turnout.

Observational Evidence: Real World Turnout is Not Responsive to Electoral Competition

U.S. presidential elections provide an opportunity to assess the relationship between pivotality and turnout because the Electoral College effectively weights some states more than others. In the same race, a voter in Ohio may have a significantly higher pivotal probability than a voter in Wyoming. Several studies exploit this variation to argue that electoral competitiveness increases political participation (Kim et al. 1975; Hill and McKee 2005; Wolak 2006; Lipsitz 2008). However, the observed correlations between closeness and turnout are typically small and, of course, might not result from a direct effect of closeness on voter psychology. Increased campaign activity or other omitted variables may explain the results of these studies. Illustrating the substantively small size of this effect, Gerber et al. (2009) find that turnout is only 2 percentage points higher in battleground states compared to spectator states. Of course, this small difference could easily be the result of campaign activity alone considering the intensity of campaigning in these states (Shaw 1999, Shaw 2006). Moreover, to assess the relative size of this 2 percent effect, Gerber

et al. (2009) point out that other election effects "dwarf the effect of being in a pivotal (battleground) state" (p. 4).

In Figure 4.1, we present our own simple analysis of turnout in Presidential elections from 1964 to 2008 by plotting turnout in each state against the winning candidate's margin of victory.[15] If electoral completion or considerations of pivotality significantly influence voter turnout, then we would expect strong negative correlations between turnout and the margin of victory in each state. However, in most election years, the observed relationship between these two variables is weak, and we actually observe a positive relationship in 5 of these 12 elections. In short, there is little meaningful relationship between pivotality and turnout in presidential elections.

Previous researchers have cut the data in different ways to argue that electoral competitiveness drives turnout. This simple graphic demonstrates that the effect of electoral competitiveness is limited. Consistent with previous findings, electoral competition accounts for a tiny percentage of the overall variation in turnout (Gerber et al. 2009). Moreover, even these weak relationships likely overstate the causal effect of electoral competition on turnout because of omitted variables and reverse causation. Higher turnout in competitive states may be driven by increased campaign activity, media coverage, or other factors unrelated to citizens' perceptions. One way to avoid these difficulties is to use survey evidence to directly measure perceptions of closeness. In the next section, we turn to the survey evidence, examine the shortcomings, and attempt to better estimate the effect of close elections.

---

[15]Data obtained from David Leip's U.S. Election Atlas http://www.uselectionatlas.org/.

Figure 4.1. Turnout by Margin of Victory in Presidential Elections (1964-2008)
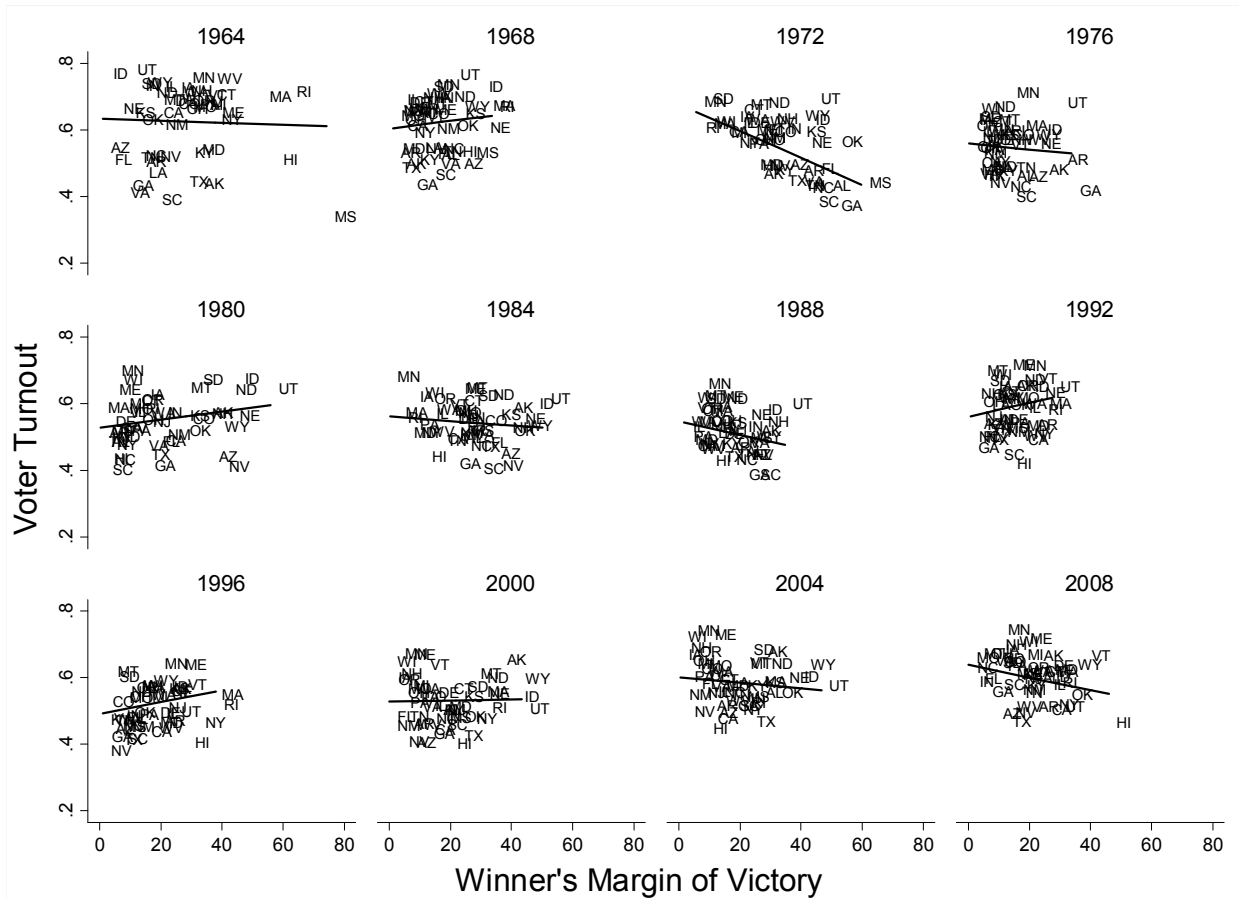
The figure indicates voter turnout in each state and each presidential election from 1964 to 2008, plotted against the winning candidate's margin of victory in the state. If pivotality exhibits a significant effect on voter psychology and voter turnout, we would expect significant, negative correlations between turnout and the margin of victory. However, in most years, the relationship between these two variables is weak, and we actually observe a positive correlation in 5 of these 12 elections.

Survey Evidence: Perceptions of Closeness are Not Meaningfully Correlated with Turnout

The first and most cited evidence that close elections influence turnout comes from survey data. Citizens who say that an upcoming election will be close are more likely to vote (Riker and Ordeshook 1968; Aldrich 1976; Blais, Young, and Lapp 2000). For every U.S. presidential election since 1952, the American National Election Study (ANES) has asked respondents to gauge whether the upcoming election will be close. Analyzing these surveys in 1952, 1956, and 1960, Riker and Ordeshook (1968) and Aldrich (1976) – adding 1972 to the analysis – show that the prediction of a close election is correlated with turnout, concluding that perceptions of closeness causally influence individuals' decisions to turn out. However, the evidence deserves further scrutiny. The relationships reported by Riker and Ordeshook (1968) and Aldrich (1976) were not subject to the rigors of multivariate regression analysis, raising the possibility that omitted variables are driving the results. Additionally, as discussed above, the analyses might be plagued by reverse causality.

In Table 4.1, we analyze all ANES surveys in presidential election years between 1952 and 2008. Like the previous studies, we find that the prediction of a close election is strongly correlated with turnout. Regressing turnout on a dummy variable, *Close* (1 = respondent thinks election will be close, 0 = not close), and election fixed effects, we find that turnout is 7 percentage points higher among those who predict a close election (Column 1). However, the effect shrinks to 4 percentage points when we control for each respondent's turnout in the previous presidential election (Column 2), and it shrinks further to 2 percentage points when we include additional demographic controls such as age, education, and political knowledge (Column 3).

Even these smaller estimates of the effect of P may not represent the true effect. Turnout itself could influence perceptions of P through the well-known process of cognitive dissonance (Festinger 1957). Voters may justify their decisions to turn out by convincing themselves that their vote matters. In this way, voters may turn out for reasons unrelated to P but increase their stated

Table 4.1. Perceived Closeness and Turnout in Presidential Elections, ANES 1952-2008

	DV = Turnout			Turnout, t-1
Close	.071	.039	.021	.067
	(.006)**	(.005)**	(.006)**	(.006)**
Turnout, t-1		.483	.393	
		(.006)**	(.006)**	
Election Fixed Effects	X	X	X	X
Demographic Controls			X	
N-Observations	27,725	27,725	25,429	27,725
N-Elections	15	15	15	15
R-squared	.005	.217	.272	.005
SER	.47	.42	.40	.45

*Standard errors in parentheses; ** significant at 1%*

Each OLS regression includes all ANES respondents in presidential election years from 1952 to 2008. Close is a dummy variable indicating whether the respondent perceives that the upcoming election will be close. In the first three models, the dependent variable is reported turnout in that presidential election, and in the last model, the dependent variable is reported turnout in the previous presidential election. Demographic controls include white racial identity, gender, age, squared age, income, education, church attendance, party identity, and political knowledge.

perceptions of closeness to avoid cognitive dissonance. Our analysis in Table 4.1 is consistent with this pattern of dissonance. In column 4, we conduct a placebo test, regressing each respondent's turnout in the previous presidential election on their perception of closeness in the current election. If dissonance is at work, we would expect that perceptions of closeness in the current election are related to higher levels of turnout in previous elections, since voters may justify their past behavior by adjusting their current perceptions. Consistent with this hypothesis, voters who think the upcoming election will be close are 7 percentage points more likely to have voted in the previous presidential election. Moreover, this placebo effect is the same size as the observed effect in column 1, where turnout in the current election is regressed on perceptions of closeness in the current election, suggesting that most, if not all, of the previously observed effects are driven by omitted variables, reverse causation, and cognitive dissonance. The observed patterns in Table 4.1 are unchanged if we restrict our analysis to only those years analyzed by Riker and Ordeshook (1968) or Aldrich (1976), indicating that the small effect of P on turnout is not simply a result of the expanded time-frame of our analysis.

Field Experiment during a Tied Election

Using aggregate observational data, we have demonstrated that the effect of closeness on voter turnout is likely small and we are left with little room for an effect of P on the individual calculus of voting. Re-examining the survey data, we have again shown that the effect of P is likely small and difficult to separate from explanations of reverse causality. We now turn to a very rare circumstance that allows us to directly test the effect of the psychological impact of close elections in an actual election. The 2010 November election for Massachusetts State House in the 6th Worcester District ended in a dead heat. After a series of recounts and a court case, Geraldo Alicea, the Democratic candidate, and Peter Durant, the Republican candidate, had each received exactly 6,587

votes - an exact tie. A special election was scheduled for May 10, 2011, and the race was likely to be close again. The same candidates who had just produced the tie would square off again with the same voters[16]. If closeness could ever directly motivate turnout in a large election, it would do so in the May special election. We took advantage of this situation to directly test P by manipulating voters' knowledge of closeness.

By experimentally manipulating the awareness of P in an actual election, we are subjecting the effect closeness to a test that avoids many of the pitfalls – such as reverse causation and omitted variables – that have plagued previous. Moreover, because P is credibly approaching its theoretical maximum value that is achieved only in an exact tie, we are allowing for the greatest chance of observing closeness impacting the decision to vote in an actual election of any significance. We also have the advantage of testing for P in an election where the B term is likely to be large because of the value of seats in the Massachusetts House of Representative: the legislature is among the most highly professionalized in the country (King 2000; Squire 2007), allowing legislators to exert great influence on state agencies with which voters interact (Woods and Baranowski 2006). Over 13 million dollars was spent on elections to the State House in 2010 alone.[17] And in this particular race, the state parties, independent groups, and the media took a strong interest in the election – indicating the high value attached to the seat. The large B term allows experimentally induced changes in the P value to affect the utility of voting and thus affect turnout.

To measure the overall effect of closeness on voter turnout, including both individual psychology and other effects, we would have to manipulate of the actual closeness of elections. Given the likely impossibility of such a design, we test the effect of closeness by manipulating voter

---

[16]Two independent candidates entered the race, but this did not change the fact that the special election would also be a tight contest between Alicea and Durant.

[17] www.followthemoney.org

knowledge of closeness in an election that is credibly close. This design allows us to separate the direct effect of electoral competition from the effects of campaigns, media, and other omitted variables. In this way, while this study does not actually measure the overall effect of close elections, it isolates the effects of close elections on individual voter psychology as hypothesized in the Riker and Odeshook model.

Despite the interesting circumstances of the election, the average citizen in the district was relatively uninformed about the race. Conducting a brief phone survey of registered voters in the three days leading up to the election, we found that only 64% knew that a special election was coming up on May 10th, only 52% could name both candidates, and only 41% knew that the previous election had ended in a tie. These figures likely overestimate the true level of knowledge because the surveys only involved registered voters who answered the phone and agreed to answer questions about an upcoming election. Walking around the district on Election Day, we saw little campaign material aside from the polling places and the two candidates' headquarters. Subsequent interviews with the candidates and campaign staffs revealed that the campaigns had focused their efforts on likely voters, targeting based on previous voter turnout, party registration, ethnicity, and expressed support.[18] One of the campaigns even admitted to performing no voter contact in several towns in the district. Furthermore, neither candidate directly mentioned the tied election in their campaign materials (author interviews, July and August 2011).

In the days prior to the election, we placed phone calls to registered voters to remind them about the special election and, for some, about the exact tie in the previous election and that their vote had a relatively high chance of being pivotal. As it turned out, the special election was also an

---

[18] The Alicea campaign concentrated disproportionally on Latino voters, which may have had some residual effect in targeting low-propensity voters, but this was not the explicit strategy of the campaign.

extremely tight race. Durant defeated Alicea by only 56 votes, so our "pivotal" treatment was credible in informing citizens that the race would be close. At first glance, it appears that citizens in the district failed to respond to the unique circumstances of the extremely tight election. Only 20% of the district's residents turned out in the special election, compared to 33% for the November election. Moreover, 16% of the special election voters supported an independent candidate who had little chance to win. These observations show little support for the claim that voters are motivated by electoral competition, but our experimental design allows us to explicitly test the causal effect of perceived pivotality on the individual calculus of voting.

Experimental Design

Given the rare circumstances around the election and the fact that most citizens were unaware of the tie, the special election in the 6th Worcester District provided the ideal opportunity to test for the effects of electoral competition on voter turnout. Some individuals in the district received a simple reminder that a special election was coming up, while others received the same reminder plus information about the tie and the unusually high chance that their vote could be pivotal. If the "pivotal" treatment mobilized more voters than the "reminder" treatment, then we would conclude that knowledge of electoral closeness mobilized voters. We applied the treatment through phone calls during the three days leading up to the election. The phone study allowed us to first gauge each respondent's level of knowledge and then provide an experimental treatment. Phone experiments have been shown to substantially mobilize voters if conducted in a certain fashion (Nickerson 2006, 2007). Our calls were conducted in a fashion that included the elements of Nickerson (2006) and Nickerson (2007) that have been demonstrated to increase turnout: we recruited student volunteers to make the calls and we provided a script but instructed the callers to

convey the appropriate information in a casual, conversational style without reading directly from the page.

We obtained the list of registered voters in the district and their phone numbers from Catalist, a for-profit political data vendor. Catalist provided us with 19,327 registered voters in the district for whom they could connect with a phone number. There were 9,318 unique phone numbers in the list. To check the quality of numbers, we used Call Fire, a robo-call service, to identify and remove eighty-four invalid phone numbers from our list. Then, we removed 369 phone numbers that had more than four registered voters associated with it to make it easier for our callers to identify the recipient of the call. This strategy of removing large households has been previously employed to improve both the administration and the interpretation of field experiments (Nickerson 2008; Huber, Gerber, and Washington 2010; Nickerson and Rogers 2010). For the remaining 8,865 phone numbers, we randomly assigned them into one of four conditions. Phone numbers were stratified according to their pre-treatment characteristics, and randomization was conducted within each stratum. We provide a detailed account of the randomization procedure in the Appendix.

Two thousand, nine hundred fifty-five phone numbers, one-third of the population, were randomly assigned into the "reminder" condition. In this condition, a caller would introduce herself and identify the recipient of the call. If a registered voter did not answer the phone, then the caller would try to ask for one of the voters associated with that phone number. In a casual manner, the caller would ask whether the recipient knew about the special election coming up. If the recipient claimed to know, then the caller would ask for the day that the election will take place. These two questions allowed us to identify the voters' prior knowledge about the election, which was used in our subsequent analysis. Then, the caller would simply remind the recipient that the election would be on May 10th to fill a seat in the State House. If the recipient asked for additional information, callers where permitted to provide the names and parties of the two candidates. However, for any

additional questions, the recipients were referred to the website of the Massachusetts Secretary of the Commonwealth. The complete script given to callers is available in the Appendix.

Another 2,955 phone numbers were randomly assigned to the "pivotal" condition. This treatment was identical to the "reminder" condition with one exception. At the end of the call, the caller would also inform the recipient about the reason for the special election, by saying: "The reason that there is a special election is that the last election ended in an exact tie. Had one more or one less person voted in the last election, your candidate would have won. The special election on Tuesday is likely to be close again, so there is a high chance that your vote could make a difference." The response of voters to this script, which our callers recorded, indicated a successful manipulation. For example, voters often responded by assuring the caller that they had voted despite personal hardship or relaying an anecdote about a friend of relative that had not voted and proclaiming that this had made the difference in the election.

A separate 296 phone numbers, just one-thirtieth of the population, were randomly assigned to a survey condition. This condition allowed us to better gauge the extent of political knowledge among our population. As in the "reminder" and "pivotal" conditions, we asked recipients whether they knew about the upcoming election and whether they knew the date. Then, we also asked whether they could name the candidate or whether they knew the reason for the special election. As mentioned previously, only 41% of respondents knew that the previous election ended in a tie, suggesting that the majority of our sample was unaware of the unusually high chance that their vote could make a difference. Lastly, the remaining 2,659 phone numbers were assigned to a "no contact" condition, where no phone call would be attempted. After numbers were assigned to conditions, they were randomly sorted before given to callers. Therefore, the timing of the call was also random and each caller had approximately the same proportion of reminder, pivotal, and survey calls. Prior to randomization, we stratified the subjects according to previous voter turnout and

other characteristics. A complete description of our randomization procedures can be found in the Appendix.

Our experiment allows us to assess the causal effect of being informed about electoral competition on turnout in the special election. Lists of special election voters were obtained from the 5 different town clerks in the district and matched to the list of treated individuals and treated phone numbers. 97% of our subjects were successfully matched to the towns' records of registered voters. Importantly, rates of attrition are the same for voters in both treatment groups. The difference in turnout between the pivotal group and the reminder group represents the causal effect of interest, because the conditions were randomly assigned, and the only difference between the two conditions is the information we provided about electoral closeness. If closeness has a substantial causal effect on the individual calculus of voting, then we expect a large difference. However, if the effect is negligible, as suggested by our observational and survey estimates, then we expect little difference between the pivotal and reminder groups. Our experiment also allows us to test the effects of our manipulation across voters with different voting propensities, thereby testing whether being informed of electoral closeness had the greatest effect on non-regular voters and thereby diminished inequalities in turnout or the greatest effect on regular voters and thereby increased inequalities in turnout.

Estimation strategy

Randomization and stratification, in addition to the advantage of avoiding imbalance between treatment groups, allows more precise estimates, because we can estimate the effect of the

pivotal treatment versus the reminder treatment with strata fixed effects. Looking at all individuals i, in all strata j, we employ OLS[19] to estimate the following equation:

(3) $$Y_{ij} = \beta * Pivotal_{ij} + \gamma_j + \varepsilon_{ij}.$$

Y_{ij} is a dummy variable indicating whether the individual turned out in the May 10th special election. $Pivotal_{ij}$ is a dummy variable, taking a value of 1 if that individual is in the "pivotal" condition and a value of 0 if the individual is in the "reminder" condition. Individuals in the survey or no contact conditions are omitted from the subsequent analysis. γ_j represents a specific effect for each stratum. Some strata are likely to turn out at high rates, while others will not. This procedure removes that source of variation and estimates the effect of the pivotal treatment within strata. Where applicable, standard errors are clustered by phone number, since that is the level at which individuals and households are treated.

Results

As described earlier, we randomly assigned 5,910 phone numbers to receive either the "reminder" or the "pivotal" treatment during the three day period preceding the special election. Due to time and labor limitations, we were only able to call 5,157 (87%) of those numbers. Because the order of call attempts was randomly assigned, similar numbers of attempts were made to phone numbers in the reminder and pivotal conditions, 2,599 and 2,558 respectively. Of those attempts,

---

[19] Since the dependent variable, turnout, is dichotomous, we can also estimate the treatment effect using logit or probit. In all cases, marginal effects from logit or probit are nearly identical to our OLS coefficients. Moreover, OLS is superior to other estimators because linearity is guaranteed because all independent variables are also dichotomous (this does not imply that treatment effects are homogeneous) and OLS overcomes the incidental parameters problem which would bias the results of other estimators.

1,021 (20%) resulted in complete treatments. This figure is comparable to response rates in previous phone experiments (Nickerson 2006). Importantly, the response rates were similar for the reminder and pivotal conditions, 20.3% and 19.3% respectively.

In Table 4.2, we estimate the effect of the pivotal treatment for 5 different subsets of experimental subjects. The Appendix provides the mean turnout rate for each relevant subset of experimental subjects, showing that simple differences-in-means yield the same results as our strata fixed-effects models. Row 1 shows our intent-to-treat estimate, the average effect of being assigned to the pivotal treatment compared to the reminder treatment. The estimate of .006 indicates that after removing strata fixed effects, subjects in the pivotal treatment group were 0.6 percentage points more likely to vote in the special election compared to those assigned to the reminder group. This effect is substantively tiny and statistically insignificant, suggesting that a pivotal treatment is not an effective method for increasing aggregate turnout. We go on to analyze smaller subsets of the data according to our theoretical predictions about who should be most affected by the treatment. Keep in mind that the precision of our estimates declines and the subsample becomes smaller, but the results are interesting nonetheless.

In Row 2, we isolate those subjects who answered the phone and received a treatment. According to this estimate, the pivotal treatment increased turnout among this subgroup by 1.2%, but this estimate is not statistically significant. In Row 3, we home in on just those individuals who were directly contacted and who were uninformed about the upcoming election. There are only 317 such individuals so this estimate is necessarily imprecise. Here, the estimated effect of the pivotal treatment is larger, 5.2%, but still statistically insignificant.

Compared to other results from get-out-the-vote experiments, our 5.2 percentage point estimate for uninformed, contacted individuals may, at first, appear large. However, it is important to recognize that the overall intent-to-treat effect is 0.6 percentage points, and we only obtain this

115

Table 4.2. Experimental Effect of Pivotal Treatment

	Estimate	S.E.	P-value	Obs
Intent-to-Treat	.006	.009	.491	11361
Contacted Individuals	.012	.029	.670	936
Contacted, Uninformed Individuals	.052	.043	.225	317
Contacted, Uninformed, voted in > 2 recent elections	.185	.088	.037	139
Contacted, Uninformed, voted in ≤ 2 recent elections	.018	.028	.525	178

The table presents the estimated effect of the pivotal treatment relative to the reminder treatment for five different subsets of experimental subjects. The estimates are OLS coefficients resulting from a regression of turnout on the pivotal treatment and strata fixed effects. Standard errors are clustered by phone number for the intention-to-treat estimate.

5.2 percentage point figure after narrowing our sample down to the least informed individuals that we contacted. If other get-out-the-vote experiments narrowed their samples in this way (imagine focusing only on those subjects who received the message and were clearly informed by the message), they would likely find much larger treatment effects. Most importantly, if electoral competition exhibits a strong effect on voter psychology as previous literature suggests, these treatment effects should be massive – much larger than the 5 percentage point estimate that we obtain.

We cannot know whether some of these positive coefficients have arisen by chance alone or through some real effect of our pivotal treatment. Our observed effects are not larger than we would expect by chance and they are smaller than we would expect if closeness weighs heavily on the calculus of voting. Even in this case of an extremely close election, voters did not turn out at high rates, and informing them about closeness had no statistically significant mobilizing effect. However, our estimate of the average effect on uninformed voters is not zero, so we can look more closely at the effect on subsets of voters to see if closeness has a mobilizing effect on habitual non-voters.

The Effect on Infrequent Voters

As previously discussed, electoral closeness has frequently been posited as a cure for unequal participation (Blais, Young, and Lapp 2000; Key 1949; Piven and Cloward 2000; Schattschneider 1960; Franklin 2004). Our experiment provides a rare opportunity to test this claim by examining the differential effects of the treatment on regular and non-regular voters. If close elections can help to equalize turnout, then we should see a stronger effect of the treatment on voters have voted infrequently in previous elections.

I rows 4 and 5 of Table 4.3, we estimate the effect of the pivotal treatment separately for those who voted in 3 or more of the 9 most recent elections and for those who voted in 2 or less.[20] We choose this arbitrary cutoff because it splits our sample roughly in half, but Figure 4.2 shows the effect across all levels of vote history. We find a large, statistically significant effect for regular voters (18.5%) but no effect for infrequent voters. A single interactive model yields the same result. The treatment effect is zero for those who have not previously voted, but the interaction between the pivotal treatment and previous turnout is large and statistically significant. The size of the effect for regular voters is large but not unprecedented (e.g. Michelson, Garcia Bedolla, and Green 2008; Green and Gerber 2008), especially for a small subgroup. More important than the exact size of the point estimate, which is, of course, subject to statistical uncertainty, is that the effect for regular voters is statistically and substantively larger than the effect for non-regular voters.

Figure 4.2 further demonstrates the variation of the pivotal treatment effect across prior levels of voter turnout. The figure presents separate kernel regressions of turnout in the special election across the probability of turnout in the 9 previous elections for uninformed individuals in the pivotal and reminder groups. There is no effect of the pivotal treatment for infrequent voters, but there is a sizable effect for frequent voters.

While considerations of pivotality may mobilize a small subset of individuals, it has virtually no mobilizing effect for the underrepresented citizens who vote infrequently. In fact, the size of the treatment effect on regular voters was large enough to open a substantively significant gap between the turnout of regular voters and those who vote infrequently. The effect of the additional turnout stimulated by the treatment condition reaches an average of over 18 percentage points among the most frequent voters. This finding suggests that, if anything, increased competition and pivotality

---

[20] The 9 recent elections are the 2010 general; 2009 special; 2008 general, primary, and special; 2007 special; and 2006 general, primary, and special elections.

Figure 4.2. Effect of Pivotal Treatment for Uninformed Subjects across Previous Turnout
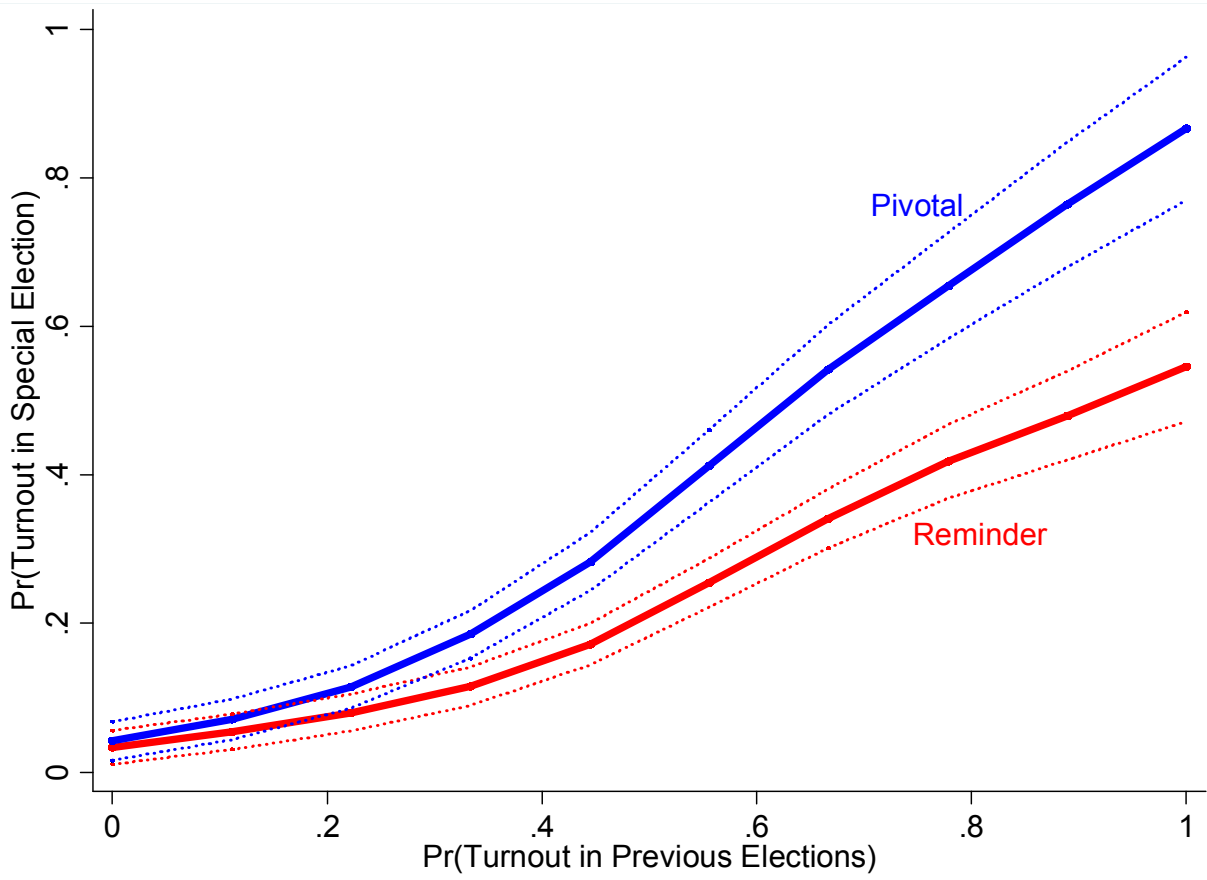
Solid lines represent kernel regressions of turnout in the special election across each respondent's probability of turnout in the previous nine elections. The sample includes only contacted subjects who were initially uninformed about the upcoming special election. Dotted lines represent standard errors.

may widen the turnout gap between the politically active and the politically under-represented. This result is consistent with our theoretical prediction that closeness should influence those with high B terms and D terms large enough (close enough to C) to make them almost indifferent between voting and abstention.

Placebo/Balance Tests

Although we find minimal effects of the pivotal treatment compared to the reminder condition, we might worry that a real effect has been masked or counteracted by unobserved differences between the two experimental groups. Random assignment guarantees that the two groups are asymptotically identical, but for our finite sample size, small differences may arise by chance. Our stratified randomization procedure ensured greater balance between treatment conditions for certain variables, but small differences could arise for other variables. Moreover, the subset of households or individuals who answered the phone and received a treatment could be slightly different by chance in the two conditions.

To demonstrate that these concerns do not plague experimental results, researchers often present the difference between the two experimental conditions for a number of pre-treatment variables. Since our analysis first requires the removal of strata fixed-effects, a simple difference-in-means alone would not assure us that our estimates are valid. Instead, we present a number of placebo tests, replicating our previous estimates but substituting our dependent variable for numerous pre-treatment variables. Regressing a pre-treatment variable on a dummy for the pivotal treatment and strata fixed-effects, we can assess the extent of imbalance between treatment conditions for that particular variable. The Appendix presents the results of 45 placebo tests. Each cell in the table presents the coefficient on the pivotal treatment and the corresponding standard

error for a particular sample and for a particular placebo outcome. As in Table 4.3, standard errors are clustered by phone number where appropriate.

There are few meaningful differences in pre-treatment characteristics between treatment conditions. Individuals in the two experimental groups were similar in terms of their turnout in the 2010 general election, the 2009 special election, the 2008 primary election, Hispanic racial identity, age, party registration, absentee voting in the special election, and whether we were able to match them to the town voter files. If anything, uninformed individuals in the pivotal group were, by chance, more likely to vote in previous elections, suggesting that our previous estimates may overestimate the effect of the pivotal treatment. Only 1 out of 45 coefficients is statistically significant at the 5% level. Contacted, uninformed individuals in the pivotal condition were less likely to be Hispanic than their counterparts in the reminder condition. Since only 317 individuals fall into this category, we expect some imbalances to arise by chance. However, this difference does not bias our experimental results. None of our results in Table 4.3 are significantly changed by the inclusion or exclusion of pre-treatment controls. These placebo tests confirm that subjects in the pivotal and reminder conditions are comparable to each other and only differed in the extent to which we informed them about the closeness of the upcoming election.

Discussion and Conclusion

In this paper, we assess the effect of pivotality on the calculus of voting. We examine pivotality in an aggregate context; we examine closeness using survey measures; and, finally, we identify and exploit a rare opportunity to credibly test for the effect of closeness in the aftermath of a tied election for a major political office. This last test exploits such a rare event that this test is

unlikely to be replicable.[21] Through all these tests, we find little evidence that the closeness of elections and considerations of pivotality motivate voters to turn out.

Of course, we also leave many unanswered questions. As discussed previously, since close elections affect turnout through mechanisms other than individual voter psychology, we cannot measure the overall effect of close elections nor can we measure the long-term effects of close elections. However, we do isolate the short-term effects of close elections on individual voter psychology. The long-term effects of close elections remain to be studied, although the fact that Ohio's turnout in presidential elections is not dramatically greater than Wyoming's suggests that these effects cannot be huge. Also, we cannot say whether our results would hold in a more salient election, but the inability of competition and pivotality to dramatically mobilize citizens in a large state legislative election casts doubt on its potential in other electoral settings as well.

Our findings are consistent with literature that questions the usefulness of the classic models of voting. Some individuals like voting and others do not. However, turnout decisions are rarely influenced by an individual's perception that her vote could influence political outcomes. To the extent that these considerations do matter, the effect appears to be concentrated among a small subset of regular voters. Citizens who are not inclined to vote will not be mobilized in competitive elections, even in a rare case when their vote could make a difference.

While our findings challenge the usefulness of traditional models of turnout in which P factors into the calculus of voting, they do not reject the canonical model or distinguish between other models which exclude P. Other models, often forgotten in recent scholarship, may prove useful in understanding turnout. For example, the minimax-regret model (Ferejohn and Fiorina

---

[21]Muligan and Hunter (2003) analyzed all contested state legislative elections between 1968 and 1989. In 40,036 state elections, two ended in a tie – one of which resulted in another election. In the 16,577 contested federal elections between 1898 and 1992, one ended in a tie.

1974) posits that citizens may vote to avoid the dreadful outcome of a tied election in which she failed to vote, and Schwartz (1987) offers a model in which turnout per se leads to direct benefits for an individual or her community. Further progress in understanding and modeling the calculus of voting will likely arise from careful testing of these alternative models and from the identification of forces that factor into D and C.

Close elections are certainly different from blowouts or uncontested elections, and future researchers will continue to identify such differences. For example, Fraga and Hersh (2011) argue that voters are less sensitive to changes in the cost of voting when elections are close and Kam and Utych (2011) argue that voters are more cognitively engaged in close elections. Our results suggest that these findings and other such effects of close elections are minimally motivated by individual voter psychology. Rather, they are likely driven by the heightened campaign activity, elite discourse, or media coverage during close elections.

Voting is possibly the single most important act of democracy, and competitive elections are fundamental to a functioning democracy (Dahl 1970). Even though citizens often fail to vote or pay close attention to politics, democratic theory maintains that they will mobilize when it matters (Dahl 1961). One such event when a responsive citizenry should mobilize is a close election. However, our results suggest that voters fail to consider the chances that their vote will be pivotal and therefore fail to participate when the stakes are high. These findings cast doubt on the ability of an apathetic electorate to hold their government accountable. The available evidence suggests that considerations of pivotality and electoral competition do not factor into the calculus of voting in a significant manner.

Study #5

Increasing Inequality: The Effect of GOTV Mobilization on the Composition of the Electorate

Co-authored with Ryan D. Enos and Lynn Vavreck

Abstract

Numerous get-out-the-vote (GOTV) interventions are successful in raising voter turnout. However, these increases may not be evenly distributed across the electorate and may actually increase the differences between voters and non-voters. This phenomenon is particularly notable given the many GOTV strategies that explicitly aim to reduce inequalities in representation. By analyzing individual level-data, we reassess previous GOTV experiments to determine which interventions mobilize under-represented versus well-represented citizens. We develop a generalized and exportable test which indicates whether a particular intervention reduces or exacerbates disparities in political participation and apply it to 26 previous experimental interventions. Despite raising mean levels of voter turnout, more than two-thirds of the interventions in our sample widened disparities in participation. On average, voter mobilization strategies tend to increase the participation gap, thereby exacerbating representational inequality. We conclude by discussing substantive implications for political representation and methodological implications for experimenters.

Scholars are increasingly interested in inequalities in political participation and their consequences for political outcomes (APSA Task Force 2004; Bartels 2008, 2009; Dahl 2006; Gilens 2005, 2012; Schlozman, Verba, and Brady 2012). At the same time, political scientists increasingly use large-scale field experiments to test methods for increasing political participation. The findings of these experiments have been adopted by political campaigns and are an increasingly important feature of electoral politics (Green and Gerber 2008; Gerber et al. 2011; Issenberg 2010, 2012). Often, a goal of these get-out-the-vote (GOTV) strategies is to reduce the participation gap - the extent to which the electorate differs from the voting-eligible population. While many experiments do successfully increase average levels of voter participation, they may not affect all citizens equally. Instead, an experiment may affect regular voters more than infrequent voters. In this paper, we explicitly test whether GOTV treatments tend to reduce or exacerbate the gap in political participation. We find that GOTV interventions, on average, tend to magnify the participation gap. This finding is widely important for the study and equalization of political representation, political campaigns, and voting behavior.

In a typical GOTV experiment, a researcher will randomly assign individuals, households, or geographic regions to receive a particular treatment. Then, the researcher will typically collect voting data from public records and estimate the average effect of the treatment on voter turnout by calculating a simple difference-in-means. To our knowledge, there are four primary motivations for conducting GOTV experiments. First, scholars and civic groups may want to increase political participation as an end in and of itself. Second, field experiments can improve the efficiency of political campaigns. Third, these experiments might inform the study of voting behavior and human behavior in general. Fourth, the participation gap may hold critical political and policy outcomes, and GOTV experiments can identify methods for reducing this gap. For all but the first motivation, the traditional GOTV analysis of calculating average treatment effects is usually insufficient. In

addition to knowing *how many* voters were mobilized by a particular treatment, we also need to know *which types* of voters were mobilized. We offer a generalized method for assessing this type of heterogeneity in experimental treatment effects which can lend insights for all three of the preceding motivations, although we focus mainly on the last.

Because voters are systematically unrepresentative of the rest of the population, some individuals and groups may be significantly under represented relative to others (Verba, Schlozman, and Brady 1995; Lijphart 1997). By studying the determinants of voter turnout, scholars may be able to propose policies to reduce differences in political participation. We develop a single statistical procedure that can be applied to any GOTV experiment to explicitly test whether the intervention reduced or exacerbated disparities in voter turnout. The test can be easily implemented for any previous experiment, often without the collection of any additional data. We apply this method to 26 experimental interventions from 12 published papers to assess the overall implications of the get-out-the-vote paradigm for equality and representation. Eighteen of the 26 interventions, more than two-thirds, exacerbated the participation gap, and this effect is statistically significant in 9 cases. Alternatively, only 2 interventions significantly reduced the participation gap. A pooled analysis of all interventions reveals a large and statistically significant exacerbating effect of GOTV on participation differences.

By analyzing many experiments at once, we build upon the pioneering work of Arceneaux and Nickerson (2009) who re-analyze 11 previous door-to-door canvassing experiments to determine which subset of citizens should be targeted to achieve the most efficient allocation of campaign resources. They argue that campaigns should target higher-propensity voters in low-salience elections and lower-propensity voters in high-salience elections. In this paper, we expand the number of interventions to 26, examine many different types of mobilization strategies, and move beyond the focus of campaigns and electioneering to assess the effects of voter mobilization

on inequality and political representation. Modern GOTV strategies, which are being actively used by interest groups and campaigns, can dramatically change the composition of the electorate.

Get-Out-the-Vote and the Political Consequences of the Participation Gap

"The existence of political equality is the fundamental premise of democracy" (Dahl 2006). Dahl's assertion presents a difficult challenge for democratic societies because political equality is undermined by disparities in voter participation. As Verba, Schlozman, and Brady (1995) put it: "since democracy implies not only government responsiveness to citizen interests but also equal consideration of the interests of each citizen, democratic participation must also be equal….No democratic nation – certainly not the United States – lives up to the ideal of participatory equality." As Citrin, Schickler, and Sides (2003) explain, "there are meaningful differences between voters and non-voters," and "tiny changes in the distribution of the popular vote (and thus small differences in turnout) can make an enormous difference to the nation's politics" (p. 88). Scholars have shown that increased turnout is associated with more equitable municipal spending (Hajnal and Trounstine 2005), altered election results (Hansford and Gomez 2010), less interest group capture (Anzia 2011), increased working class representation (Fowler 2013), and general legislative responsiveness (Griffin and Newman 2005).

Recognizing the importance of participation to a healthy democracy and electoral outcomes, political scientists have long studied the determinants of political participation (e.g. Wolfinger and Rosenstone 1980; Verba, Schlozman, and Brady 1995; Putnam 2000). The introduction of field experiments gives political scientists another tool for understanding how campaigns and policy-makers can decrease inequalities in turnout, and scholars note the potential for field experiments to increase participation among underrepresented groups. For example, Avery (1989) cites door-to-door canvassing as a method to increase participation among regular non-voters; Michelson, Garcia

127

Bedolla, and Green (2008) point to a "multi-year effort to increase voting rates among infrequent voters" that relied entirely on GOTV experiments; and Green and Michelson (2009) argue that GOTV campaigns can alter the "age, socioeconomic, and racial/ethnic disparities between voters and nonvoters" (see also Garcia Bedolla and Michelson 2012).

What has not been acknowledged by practitioners of voter mobilization is that GOTV methods may actually exacerbate the differences between voters and the voting-eligible population. Furthermore, no previous study has explicitly assessed the effects of voter mobilization on the participation gap and representational inequality. Given the prevalence of GOTV field experiments and the fact that political campaigns are now adopting the methods that these experiments have shown to be effective (Green and Gerber 2008; Gerber et al. 2011; Issenberg 2010, 2012), an assessment is warranted.

How Can GOTV Treatments Exacerbate the Participation Gap?

For both theoretical and empirical reasons, our claim that get-out-the-vote treatments may exacerbate the participation gap may be counterintuitive. First, many scholars assume that increased turnout is unambiguously good for democracy and equality (Key 1949; Schattschneider 1960; Lijphart 1997). For example, Schattschneider writes "Unquestionably, the addition of forty million voters (or any major fraction of them) would make a tremendous difference [in the quality of representation]" (1960, p. 101). Schattschneider's observation is obviously true in the extreme because with universal turnout there would be no difference between the population of voters and the voting-eligible population. However, when not everyone votes regularly, especially in systems with widespread non-participation like the United States, a marginal increase in voter turnout may actually exacerbate differences if the increase is concentrated among high-propensity voters, the types of citizens who are voting at high rates anyway.

Figure 5.1. Three Hypothetical Experiments with the Same Average Effect
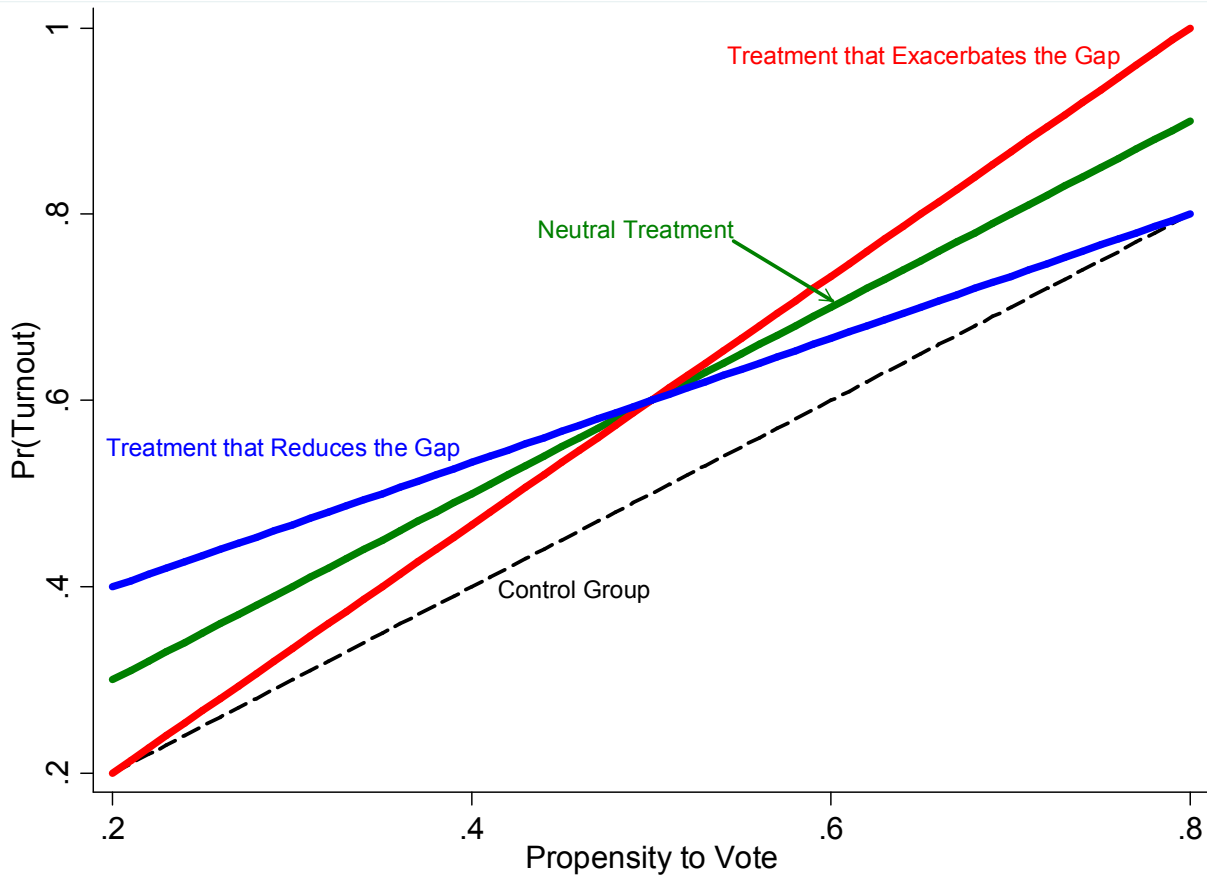
The figure presents three hypothetical experimental treatments. Each one significantly boosts the average level of turnout, but they have starkly different implications for the participation gap. The blue line reduces the participation gap by primarily mobilizing low-propensity citizens. The green line mobilizes citizens equally at all propensity levels and is therefore neutral in regard to the participation gap. The red line actually exacerbates the participation gap by primarily mobilizing high-propensity citizens, thereby making the electorate less representative of the voting-eligible population.

To illustrate this possibility, Figure 5.1 shows three hypothetical experimental treatments. For all three treatments and a control group, the lines plot voter turnout across varying underlying propensities to vote. The underlying propensity to vote can be thought of as the probability that the voter would have voted absent a treatment. All three treatments have the same average effect. For an individual with an average propensity, the treatments all increase the probability of voting by 10 percentage points. Despite this large, positive average effect, the treatments have starkly different implications for equality of participation. One treatment effect (blue) is concentrated among low-propensity voters and therefore reduces the participation gap. Another treatment (red) is concentrated among high-propensity voters, thereby exacerbating differences. The other treatment effect (green) is homogeneous and is therefore neutral in regards to these disparities.

Our subsequent findings may be counterintuitive because at high-propensity levels we expect to see a ceiling effect. If an individual is sure to vote in the absence of a treatment, then no treatment can exhibit a positive effect. For this reason, we might expect most treatments to naturally favor equality. After all, low-propensity voters have more room to increase their turnout probabilities because they are voting at lower rates. These factors make our subsequent results all the more surprising. Despite the possibility of a ceiling effect and despite the large numbers of low-propensity citizens that can be mobilized, we still find that most GOTV interventions *exacerbate* the already stark disparities in voter turnout.

Previous Studies of "Who is Mobilized?"

The varying treatment effects across different underlying propensities to vote that are illustrated in Figure 5.1 can be thought of as interactions between the propensity to vote and the treatment. Some previous GOTV studies have examined interactive effects in attempts to test mechanisms or to look for stronger or weaker effects across subgroups. For example, Nickerson

and Rogers (2010) show that discussing a "voting plan" tends to mobilize individuals who live alone but has little effect for others; Alvarez, Hopkins, and Sinclair (2010) find that partisan campaign contacts are most effective in mobilizing new registrants; Gerber and Green (2000b) find that non-partisan leaflets are most effective among non-partisans who have recently voted; and Green and Gerber (2008) find that canvassing typically boosts turnout among those who voted in the previous election.

Despite these examples, this type of analysis is rare. Experimental researchers do not typically explore all possible interactive effects of their treatment – and rightfully so because atheoretical testing of multiple interactive effects will generate false positive findings (Pocock et al. 2002; Gabler et al. 2009). While researchers could reduce this problem by testing for interactions with split samples (Green and Kern 2012), the atheoretical testing of many hypotheses is not recommended. This methodological issue poses a problem for those interested in the participation gap and the related effects on representation and policy outcomes. As previously discussed, the typical GOTV study simply tells us how many people were mobilized by a particular treatment but not which types of individuals were mobilized. To learn who is mobilized, an interactive test is called for, but this increases the danger of producing false positives. When experimenters do present interactive experimental results, readers cannot know whether they arose from a theoretically motivated hypothesis test or a post-hoc search for statistically significant coefficients.

A Novel Test for the Effect of a Treatment on the Participation Gap

We develop a single test which explicitly assesses whether an experimental treatment exacerbates or reduces disparities in participation. We propose a procedure which reduces an individual's pre-treatment characteristics onto a single dimension, her propensity to vote. Having

estimated this propensity, we test whether the experimental treatment effect varies across this single variable.

Many demographic factors predict an individual's propensity to vote. In our analysis we reduce all of the factors onto one scale - a propensity score - that indicates the *a priori* probability that a person with those characteristics will vote in the absence of any GOTV intervention. Of course, not all relevant demographic variables are available through public records used in GOTV experiments. Voter files typically indicate an individual's age, gender, geographic location, and previous turnout history. Variables like race and party registration are only available in certain states, and personal information such as income, education, or church attendance is never available. Even the limited variables that are available serve as a proxy for the extent to which a particular individual may be represented in the political process. For the purposes of this paper, we do not care *why* these demographic variables predict voter turnout. The fact that any variables can systematically predict the propensity to vote suggests that there are meaningful disparities in political participation, and our test assesses whether a particular treatment reduces or exacerbates those disparities. Later in the paper, we employ survey data to show that this propensity variable is strongly correlated with demographic factors such as income, education, church attendance, and marital status as well as political attitudes on taxes, minimum wage, federal spending, affirmative action, and other important issues.

If a specific experimental treatment mobilizes low-propensity citizens more than high-propensity citizens, then we will say that the treatment has reduced the participation gap. In other words, the demographic gap between voters and the greater population has been reduced. However, if a treatment mobilizes high-propensity citizens more than low-propensity citizens, we will say that the intervention has exacerbated the participation gap. In this scenario, the voting population has become even more unrepresentative of the general population.

Our estimation procedure involves three specific steps. First, we estimate a propensity score for every individual in the sample by regressing voter turnout on every available demographic variable for each individual in the control group.[22] The specification for estimation should be as flexible as possible given the amount of available data. We restrict this part of the analysis to subjects assigned to the control group because we want to estimate the propensity of each individual in the absence of any treatment. Because individuals have been randomly assigned, we know that the propensity scores of the control group are representative of the treatment group as well. Then, for every individual in the sample, we calculate their predicted probability of voting. For each individual, this score indicates their *a priori* probability of voting in the absence of a treatment. This score represents our single propensity variable that we employ to assess the effect of each treatment on differences in participation. For the second step, we rescale the propensity variable such that the mean equals zero and the standard deviation equals one. This is done by subtracting the mean from each individual score and then dividing by the standard deviation. This step allows us to reasonably compare treatments across different types of elections and populations, and also improves the interpretation of our subsequent estimates. Lastly, we estimate the following interactive model by OLS to test whether the treatment effect increases or decreases as the propensity score increases:

(1) $Turnout_i = \alpha + \beta * Treatment_i + \gamma * Propensity_i + \delta * Treatment_i * Propensity_i + \varepsilon.$

The two coefficients of interest are β and δ. The *treatment* coefficient, β, represents the treatment effect for an individual with an average propensity score (this is not necessarily the same as the

---

[22] We use OLS, a linear probability model, for this step but the results are virtually unchanged if we use Logit or Probit instead. Advantages of the linear probability model in the setting as compared to other models include computational speed, fewer assumption, and the avoidance of the "incidental parameters problem" (Angrist and Pischke 2009).

average treatment effect). The interactive or multiplicative coefficient, δ, represents the extent to which the treatment effect varies with the propensity score. For example, a δ of .01 indicates that the treatment effect increases by 1 percentage point, on average, for every standard deviation increase in propensity. If δ is greater than zero, then the treatment has exacerbated the participation gap, while a δ less than zero indicates that the treatment has reduced the gap. If the interactive effect is non-linear, then we should not interpret the model literally by imputing the treatment effect at various levels of propensity. Nonetheless, even relaxing any assumption of linearity, the sign of δ still indicates whether the treatment is on average greater or smaller for high-propensity citizens.

Two tricky methodological issues could theoretically pose a problem for our analysis but do not. First, the propensity variable is estimated from one regression and incorporated into a second regression. As a result, the second regression may yield incorrect standard errors by failing to incorporate the uncertainty from the first regression. We can assess this additional uncertainty with a non-parametric bootstrap, and in each case in our data the bootstrapped standard errors are virtually identical to those where we ignore the initial uncertainty of the propensity variable. Second, we might worry that our initial regression "over-fits" the data by fitting random noise rather than an underlying relationship. If this were the case, our subsequent estimates could be biased. Specifically, if we over-fit the data in the first regression, that would lead to a downward bias in the interaction term in the second regression because our propensity variable would not predict turnout out-of-sample as well as it does within-sample. We can address this issue by randomly partitioning the control group into two groups, estimating the propensity variable with one group, and then

running the second regression using only the second group. This procedure also produces results nearly identical to the original setup.[23]

Having estimated this parametric model, we also explore the non-linearity of the interactive effect through kernel regressions. By plotting the probability of turnout as a non-parametric function of the propensity score for both the treated and control observations, we obtain a more detailed picture of how the treatment effect varies across different propensity scores.

This novel analysis provides a single, theoretically-based method for evaluating the effect of a treatment on participation differences. In the following sections, we apply this procedure to 26 different GOTV interventions and then pool all of the data to assess the overall effect of these programs on the participation gap.

A Detailed Example: The "Neighbors" Treatment

To clarify our procedure, we will discuss one experiment; Gerber, Green, and Larimer's (2008) Neighbors treatment; in greater detail. Following this discussion, we will provide a broader analysis of the others. In this experiment, the authors randomly assigned registered voters to receive postcards in the run up to a low-salience 2006 primary election in Michigan. The sample included registered voters in Michigan who voted in the 2004 general election. Subjects in the control condition received no contact at all, while subjects in the treatment condition received a memorable postcard, indicating the previous turnout behavior of individuals in their neighborhood. All at once, this powerful treatment informed citizens about the upcoming election, indicated that a researcher was watching them, and threatened to share their turnout behavior with others in the neighborhood.

_____

[23] Future scholars who implement this test should be sensitive to these issues, particularly if the sample size is small.

The Neighbors treatment had a massive average treatment effect, raising turnout by 8 percentage points. However, despite a strong average effect, the treatment may not be equally effective for all types of citizens. For the purposes of our study, we want to know whether this treatment tends to reduce or exacerbate the participation gap by mobilizing low or high-propensity citizens.

To estimate our Propensity variable for each subject in the experiment, we utilize all the pre-treatment variables available in the data provided by the authors. We know whether each subject in voted the primary and general elections of 2000, 2002, and 2004; the gender of each subject; their age, and their household size (calculated as the number of voters registered at the same address). Restricting our analysis to the control group, we divided subjects' ages into 13 categories and then regressed the dependent variable (turnout in a 2006 primary) on turnout in each of the 6 previous elections, gender, age-category fixed effects, and household size fixed effects. We could have chosen a more flexible model with interactive effects, but we must be careful to avoid over-fitting. Having estimated the model by OLS, we predict for each individual in both the control and treatment groups their *a priori* probability of voting. These predicted probabilities have a mean of .30 and standard deviation of .13, which we then rescale so that the mean is 0 and the standard deviation is 1. The r-squared of .21 indicates that our model explains a significant proportion but not all of the variation in turnout.

Having estimated the propensity variable, we test whether the treatment effect varies across different propensities. As described earlier, we regress turnout on the treatment, propensity, and the interaction of the two. The coefficient on the treatment variable is .080 with a standard error of .003, indicating that the treatment effect was 8 percentage points for the hypothetical individual with

an average propensity score.[24] This is remarkably close to the average treatment effect reported by

Gerber, Green, and Larimer, but this need not be the case. The coefficient on the interaction term

is of greatest interest because it indicates how the treatment effect changes as propensity changes.

Here, we estimate a coefficient of .016 with a standard error of .002 ($p < .01$), indicating that on

average the treatment effect increases by 1.6 percentage points for every standard deviation increase

in propensity. According to our statistical test, the "neighbors" treatment exacerbated the

participation gap despite significantly raising the average level of turnout.

One way to assess the substantive size of this interactive effect is to compare the treatment

effects for two hypothetical groups whose propensity scores are respectively two standard deviations

above and below the mean. According to our model, the conditional average treatment effect of the

"neighbors" intervention was 11.2 percentage points (.080 + 2*.016 = .112) for the high-propensity

group, while the effect was only 4.8 percentage points for the low-propensity group. As such, the

neighbors treatment effect is more than 2 times greater for the highest propensity individuals in the

sample compared to the lowest propensity individuals. This difference is substantively large and

statistically significant, so we conclude that this experimental intervention significantly exacerbates

the participation gap.

Having conducted our formal statistical test, we now take a closer look at this interactive

effect with a graphical, nonparametric approach. In Figure 5.2, we employ kernel regressions to plot

the probability of turnout across individuals' *a priori* probability of voting. We plot these

relationships separately for both the control condition and the treatment condition. We display

these curves in the top panel of Figure 5.2. For low-propensity voters, there is little difference

---

[24] All standard errors are Huber-White heteroskedasticity-robust standard errors. However,

the standard errors are virtually unchanged if we cluster by household, where applicable, or if we

conduct a non-parametric bootstrap.

Figure 5.2. Variation in "Neighbors" Treatment Effect
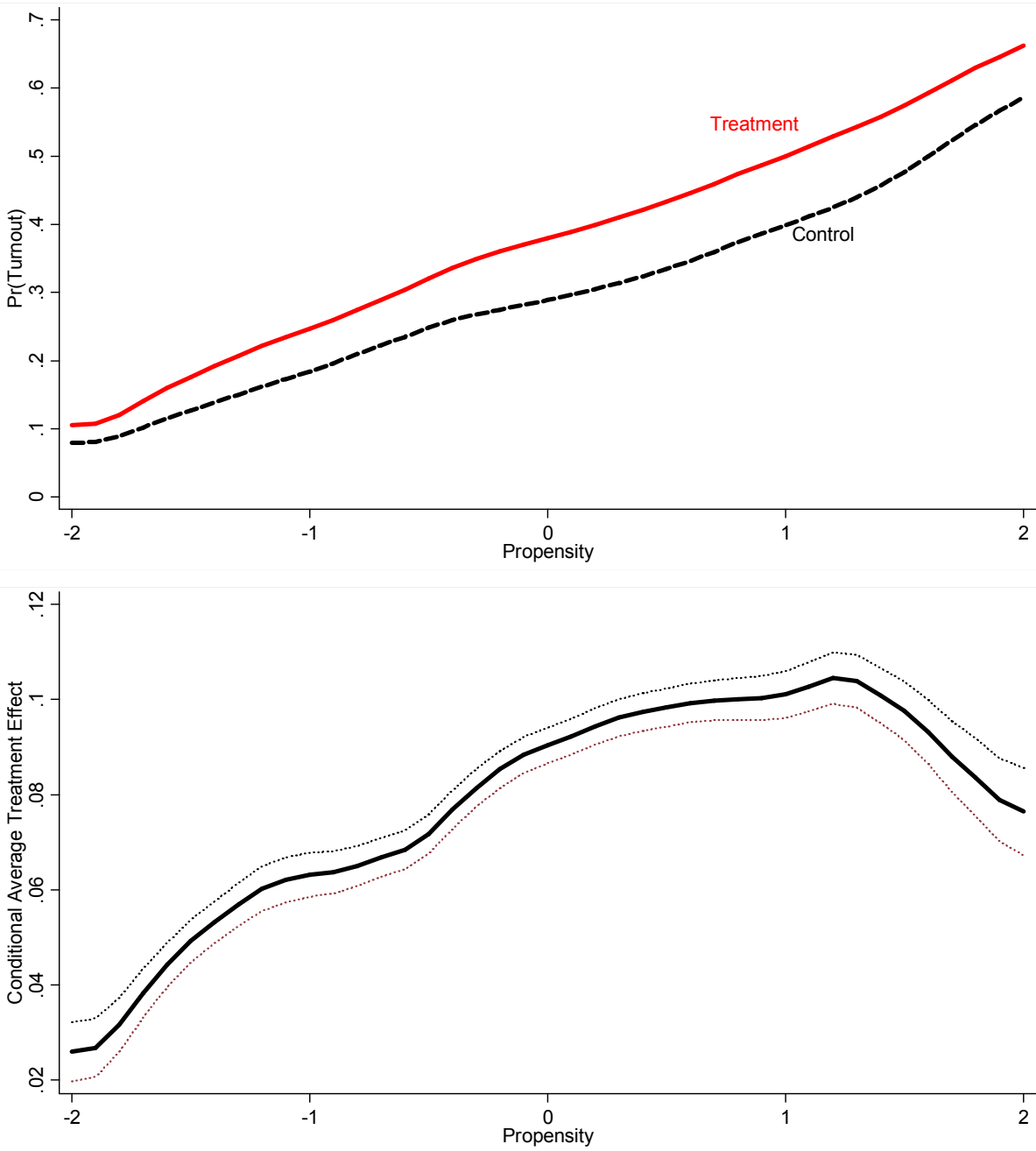
The figure shows variation in the "neighbors" treatment across different propensity levels. The top panel presents kernel regressions of voter turnout on propensity separately for the treatment and control group. The bottom panel presents the difference between these two kernel regressions, presenting the conditional average treatment effect at each propensity level. The dotted lines indicate standard errors. We see that the "neighbors" treatment is much larger for high-propensity citizens relative to low-propensity citizens, thereby exacerbating the participation gap.

between the curves, indicating a small conditional average treatment effect. However, as propensities increase, the gap between the groups increases as well, indicating that the treatment effect is becoming stronger for high-propensity citizens. In the bottom panel of Figure 5.2, we plot the difference between these two kernel regressions, showing how the treatment effect varies across different propensities. In this panel, the higher the curve, the greater the treatment effect, so a positively sloping curve indicates that the treatment effect increases as propensity increases. The curve in the bottom panel turns downward at the highest propensities, demonstrating the expected ceiling effect. However, prior to this downward turn, the curve is positive and increasing, indicating that the treatment has an increasingly strong effect as the underlying propensity to vote increases – thereby exacerbating existing differences in voter turnout.

Later, we apply these same procedures to 26 different experimental treatments. This allows us to statistically assess the effects of GOTV treatments on the participation gap and also to visualize the way in which these effects vary across different experiments. The Neighbors treatment is not an anomaly: numerous experimental interventions exacerbate the participation gap while few interventions significantly reduce it. First however, we establish that the propensity score measure we have described is associated with politically meaningful variation in the electorate.

Does the propensity score capture meaningful differences between voters?

We may not care whether GOTV interventions mobilize high or low-propensity citizens if our propensity variable does not capture meaningful characteristics of the electorate. Political scientists have extensively studied the correlates of turnout (e.g. Wolfinger and Rosenstone 1980; Verba, Schlozman, and Brady 1995; Putnam 2000), and this literature suggests that high-propensity citizens will be systematically different from low-propensity citizens across a number of politically-relevant variables. Even though we only have data on vote history and a few demographics, we

argue that our propensity variable is a proxy for many characteristics that we care about such as socio-economic status and issue positions.

We test the political relevance of our propensity variable using survey data. We employ data from the 2008 versions of the Cooperative Congressional Election Study[25] (CCES) and the Cooperative Campaign Analysis Project[26] (CCAP). For each survey respondent, we generate our propensity variable, using turnout in the 2008 general election as the dependent variable and only using information that would be available in public records as the independent variables: age, race, gender, household size, state, party registration, and vote history in previous elections. To keep in line with typical GOTV samples, we only include registered voters in the samples. Also, because survey respondents were matched to statewide voter files, we can use validated turnout data instead of reported behavior in the analysis.

Having generated our propensity variable using only those variables available in public records, we test whether this variable captures other meaningful features of the citizenry. Table 5.1 reports the results of 33 separate regressions. In each case, we regress a demographic characteristic or political attitude on the propensity variable. The coefficient indicates the extent to which the characteristic changes, on average, for every standard deviation increase in propensity. With the exceptions of family income and party identification, all dependent variables are coded as dummies, so the coefficients can be interpreted as changes in probability. For example, a single standard deviation increase in propensity corresponds with a $6,000 increase in family income, a 6 percentage point increase in the probability of a college degree, and a 3 percentage point increase in the

---

[25] Ansolabehere, Stephen, 2011-4-13, "CCES Common Content, 2008", http://hdl.handle.net/1902.1/14003 V4.

[26] Jackman, Simon and Lynn Vavreck. 2010. Cooperative Campaign Analysis Project, Release 2.1. YouGov, Inc., Palo Alto, CA.

probability of approving of George W. Bush. To summarize, high-propensity citizens, as identified by our method, are wealthier, more educated, more likely to attend church, more likely to be employed, more likely to approve of Bush, more conservative, and more Republican. They are more supportive of abortion rights and less supportive of withdrawing troops from Iraq, domestic spending, affirmative action, minimum wage, gay marriage, federal housing assistance, and taxes on wealthy families.

Even though our propensity variable relies on a small number of sparse measures that are readily available from public records, it corresponds strongly with numerous demographic characteristics and issue positions which are highly relevant in American politics. As a result, if GOTV interventions tend to mobilize high-propensity citizens over those with low propensities, they will make the electorate wealthier, more educated, more religious, and more conservative on a number of important issues. However, high-propensity citizens are not more conservative on all issues. In fact, they are more liberal on abortion and they are no different on healthcare and immigration. The abortion result is consistent with previous findings that high socioeconomic-status citizens are more supportive of abortion rights (Bartels 2008). The propensity variable is also highly correlated with intensity of preferences. For example, high-propensity citizens are much more likely to have an extreme ideology or a strong party identification. As a result, GOTV interventions which mobilize high propensity citizens are likely to increase the polarization of the electorate in addition to changing its demographic composition. This analysis demonstrates that our propensity variable does capture meaningful political differences. Having demonstrated the wealth of information captured by our propensity variable, we apply our test to experimental data.

Table 5.1. Does the propensity variable capture meaningful political characteristics?

	CCES	CCAP
Family Income	6117(305)**	6022(399)**
College	.065(.003)**	.059(.003)**
Church Attendance	.052(.003)**	.040(.005)**
Married	.097(.004)**	.041(.006)**
Employed or Retired	.086(.004)**	.068(.006)**
Bush Approval	.027(.003)**	.026(.004)**
Allow Abortion	.013(.004)**	.029(.004)**
Withdraw Troops from Iraq	-.033(.004)**	-.031(.006)**
5 Pt. Ideology (0 = Very Lib., 1 = Very Con.)	.013(.002)**	.010(.003)**
Extreme Ideology	.037(.003)**	.015(.005)**
7 Pt. Party ID (0 = Strong Dem., 1 = Strong Rep.)	.009(.003)**	.007(.004)*
Strong Party ID	.091(.004)**	.062(.005)**
Cut Domestic Spending	.014(.004)**	
Support Affirmative Action	-.018(.004)**	
Raise Minimum Wage	-.038(.003)**	
Ban Gay Marriage	.034(.004)**	
Support Federal Housing Assistance	-.020(.004)**	
Support Civil Unions		.006(.006)
Support Universal Healthcare		-.007(.006)
Raise Taxes on Wealthy Families		-.037(.005)**
Citizenship for Illegal Immigrants		.000(.006)
Sample Size	25,481	16,518

*Robust standard errors in parentheses; ** significant at 1%, * significant at 5%*

This analysis employs survey data from the 2008 CCES and 2008 CCAP studies to assess whether our propensity variable is meaningfully capturing politically relevant variables. The samples include verified registered voters who were successfully matched to statewide voter records. For both survey samples, we generate our propensity variable using validated turnout in the 2008 presidential election as our dependent variable, and using only variables available on voter files (age, race, gender, household size, state, and validated turnout in previous elections) as our independent variables. Then, we test whether this propensity variable is correlated with political characteristics and attitudes by regressing those variables on our propensity variable. Each cell in the table represents a separate regression of a dependent variable on the propensity variable. The first cell entry of 6,117 indicates that a single standard deviation increase in propensity corresponds with an increase in family income of $6,117. Aside from income, ideology, and party ID, all other variables are coded as dummy variables. The propensity variable is positively related to income, education, church attendance, marriage, employment, Bush approval, conservative ideology, Republican party identification, having an extreme ideology, having a strong party identification, and a desire to ban gay marriage. Propensity is negatively correlated with a desire to withdraw troops from Iraq, support for affirmative action, a desire to raise minimum wage, support for federal housing assistance, and support for raising taxes on wealthy families. This analysis shows that when GOTV treatments increase the prevalence of high-propensity citizens in the electorate, they are meaningfully changing the demographic and attitudinal composition of the electorate.

Applying the Test to 26 Different GOTV Treatments

We obtained our sample by identifying all GOTV field experiments published since 2000 in ten leading journals[27] where the data was available online. We augmented the sample by directly requesting data from authors. In principle we can apply our test to any GOTV effort. However, when the average effect of a get-out-the-vote effort is zero, we should not expect to find any interactive effect unless the treatment demobilizes some subset of individuals. For this reason, we restrict our analysis to available experiments with positive and statistically significant average treatment effects.[28] Even within a particular study, we only analyze those experimental treatments which exhibit a statistically significant effect on the average level of voter turnout. Table 5.2 presents a summary of the published studies utilized in this paper. For each study, the table reports the delivery method of the experimental treatment, the electoral context, and the set of covariates available for the calculation of our propensity variables.

The regression results for each treatment are presented in Table 5.3. For each experiment, the table presents the coefficient on the *Treatment* variable, indicating the effect of the treatment for the average citizen in the sample. More importantly, the table presents the coefficient on *Treatment*Propensity*, indicating the extent to which the treatment effect changes as propensity increases. Because the *Propensity* variable is recoded so that the mean equals 0 and the standard deviation equals 1, we can interpret the interactive coefficient as the extent to which the treatment

---

[27] *American Journal of Political Science*, *American Political Science Review*, *American Politics Research*, *British Journal of Political Science*, *Electoral Studies*, *Journal of Politics*, *Political Analysis*, *Political Behavior*, *Political Research Quarterly*, and *Quarterly Journal of Political Science*.

[28] Confirming this intuition, we checked for interactive effects in the experiments reported in Nickerson (2007a), Vavreck (2007), Green and Vavreck (2008), and Gerber, Karlan, and Bergan (2010) all of which had zero or close to zero average effect. We found no interactive effects.

Table 5.2. Summary of Studies and Available Data

Study	Method	Context	Age	Race	Gender	Household Size	Geography	Vote History	Party Registration	Registration Year	Survey Responses
											Available Covariates
Gerber and Green 2000 (GG00)	Multiple	1998 General - New Haven, CT	X				X	X	X		
Gerber, Green, Nickerson 2003 (GGN03)	Door	2001 Local - Multiple Cities	X	X	X			X	X	X	
Nickerson 2006 (N06)	Phone	2000 and 2001 - Multiple Cities	X		X		X	X	X	X	
Nickerson, Friedrichs, King 2006 (NFK06)	Multiple	2002 General - Michigan	X		X		X	X		X	
Nickerson 2007 (N07)	Phone	2002 General - Multiple Cities	X			X	X	X	X		
Gerber, Green, Larimer 2008 (GGL08)	Mail	2006 Primary - Michigan	X		X	X		X			
Middleton and Green 2008 (MG08)	Door	2004 General - Multiple States	X		X		X	X			
Nickerson 2008 (N08)	Door	2002 Primary - Multiple Cities	X		X		X	X	X		
Dale and Strauss 2009 (DS09)	Text	2006 General - Multiple States	X	X	X		X	X		X	
Gerber, Green, Larimer 2010 (GGL10)	Mail	2007 Local - Michigan						X			
Gerber, Huber, Washington 2010 (GHW10)	Mail	2008 Primary - Connecticut	X	X	X	X	X	X		X	X
Nickerson and Rogers 2010 (NR10)	Phone	2008 Primary - Pennsylvania	X	X	X	X	X	X		X	X

The table summarizes the available data employed for our study. Our data is drawn from 12 previously published studies, numerous different methods of delivery, and numerous types of electoral settings. To construct our propensity variable, we employ any pre-treatment variables that can help us to predict voter turnout: age, race, gender, household size, geography, vote history, party registration, registration year, and in one case survey responses.

144

Table 5.3. Summary of Results

Intervention	Treatment	Treatment*Propensity	N-treated	N-control	Study
Mail	.011(.006)*	.001(.006)	7,776	11,596	
Door	.037(.009)**	-.009(.008)	2,877	11,596	GG00
Mail+Door	.031(.011)**	.003(.010)	1,853	11,596	
Phone+Mail+Door	.024(.013)*	.024(.011)*	1,207	11,596	
Bridgeport	.049(.020)*	.052(.025)*	895	911	
Detroit	.027(.009)**	-.020(.006)**	2,472	2,482	GGN03
Minneapolis	.027(.013)*	.027(.010)**	1,409	1,418	
St. Paul	.035(.016)*	-.015(.011)	1,104	1,104	
Stonybrook	.071(.031)*	-.014(.031)	680	279	N06
Boulder	.036(.015)*	.003(.013)	1,796	1,616	
Phone	.033(.006)**	.001(.005)	7,148	13,540	NFK06
Phone+Door	.039(.005)**	-.005(.004)	14,342	13,540	
Professional	.013(.004)**	.001(.004)	30,114	29,518	N07
Prof.+Vol.	.016(.004)**	-.002(.004)	24,834	29,518	
Civic Duty	.018(.002)**	.002(.002)	38,218	191,243	
Hawthorne	.025(.002)**	.008(.002)**	38,204	191,243	GGL08
Self	.048(.003)**	.008(.002)**	38,218	191,243	
Neighbors	.080(.003)**	.016(.002)**	38,201	191,243	
MoveOn	.020(.004)**	-.015(.003)**	23,384	22,893	MG08
Minneapolis	.038(.013)*	.051(.015)**	876	1,748	N08
Text Message	.030(.010)**	-.009(.010)	4,643	4,688	DS09
Civic Duty	.017(.007)*	.017(.008)*	3,238	353,341	
Shame	.064(.006)**	.029(.006)**	6,325	353,341	GGL10
Pride	.041(.005)**	.005(.006)	6,307	353,341	
Party Reg.	.035(.008)**	.009(.010)	1,772	1,173	GHW10
Planning	.009(.003)**	.000(.003)	19,411	228,995	NR10
Pooled	**.033(.001)****	**.004(.001)****	**316,705**	**867,144**	

*Robust standard errors in parentheses; ** significant at 1%, * significant at 5%*
The table presents our regression results for 26 different experimental treatments. We only include those treatments which demonstrate a statistically significant effect on the mean level of voter turnout. The Treatment coefficient indicates the treatment effect for the average citizen in the sample. The Treatment*Propensity coefficient indicates the extent to which the treatment effect changes as Propensity increases. Because the propensity variable is recoded so that the mean equals zero and the standard deviation equals 1, we can interpret the coefficient as the extent to which the treatment effect increases as propensity increases by 1 standard deviation. In 9 cases, we see a statistically-significant, positive coefficient, indicating that the treatment exacerbated in the participation gap. In only 2 cases do we see statistically-significant evidence that the participation gap was reduced. The final row presents a pooled analysis, showing the overall effect of GOTV experiments on the participation gap.

Figure 5.3. Variation in 26 GOTV Treatment Effects
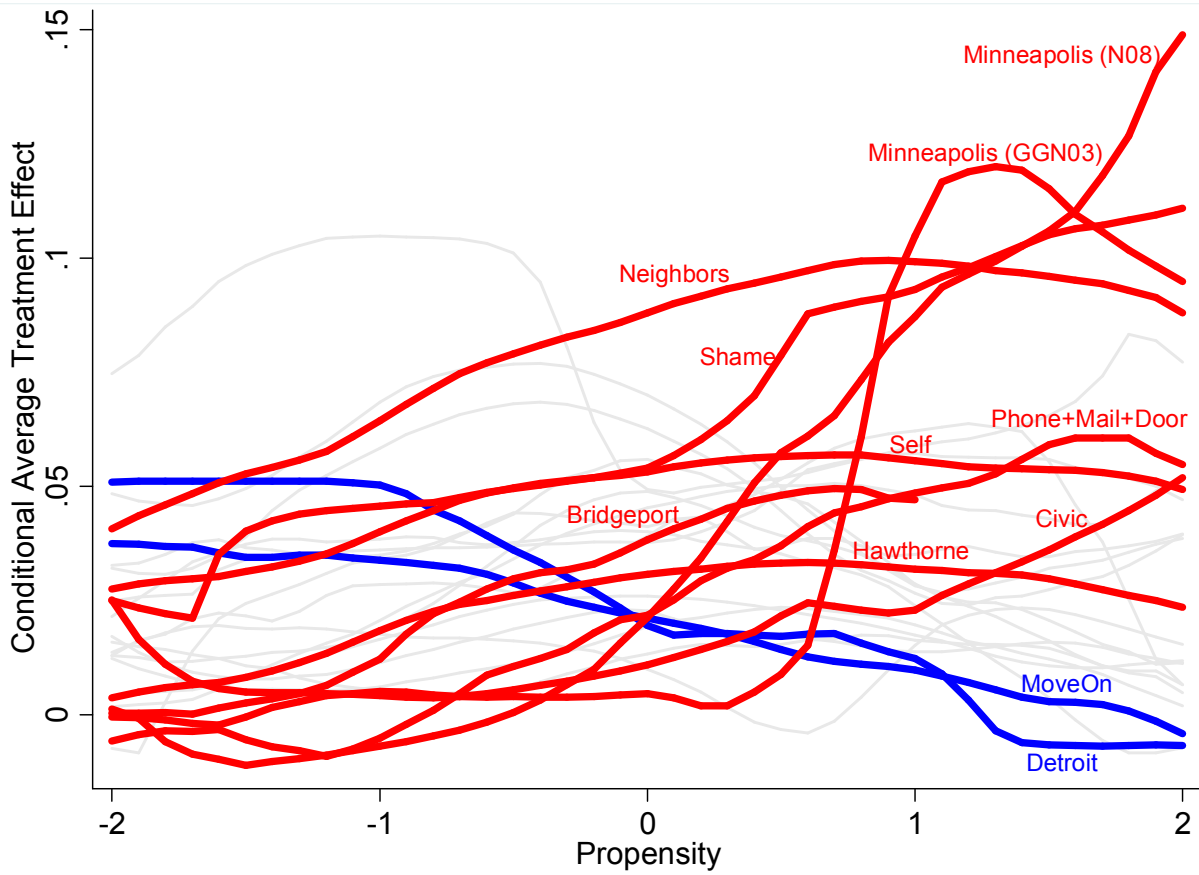
The figure presents the conditional average treatment effect for all 26 experimental interventions across different levels of propensity. These plots mimic the bottom panel of Figure 5.2. The red curves indicate treatments which significantly exacerbate the participation gap, the blue curves indicate treatments which significantly reduce the participation gap, and the gray curved indicate treatments which had no statistically significant effect.

effect increases for every standard deviation increase in propensity. Also, in Figure 5.3, we show

nonparametric analyses, similar to that in Figure 5.2, for each of these 26 interventions. Overall, we

see that GOTV interventions tend to exacerbate the participation gap. The interactive coefficient is

positive in 18 out of 26 cases (69 percent) and is positive and statistically significant for 9 of these

cases. More than two out of 3 GOTV treatments in our sample exacerbated the participation gap.

Alternatively, we find significant evidence for a reduction in the participation gap for only 2 out of

26 interventions (8 percent).

In the final row of Table 5.3 we conduct a pooled analysis by combining observations from

all experiments in a single regression. In total, over 300,000 individuals received one of these

GOTV treatments and almost 900,000 were assigned to a control group, receiving no treatment.

With these pooled data, we assess the overall effect of these GOTV treatments on the participation

gap. For 1.2 million individuals, we regress voter turnout on a treatment dummy variable, the

interaction of the treatment with each propensity score, study fixed effects, and the interaction of

each study with the propensity score. The inclusion of study dummy variables and study-propensity

interactions is necessary because the treatment variable is random within each study but may not be

random between studies.[1] Overall, these treatments exhibit a large positive effect on voter turnout

for the average individual, 3.3 percentage points. However, this treatment effect is much stronger

for high-propensity individuals. For every standard deviation increase in propensity, the treatment

---

[1] Suppose 50% of subjects are assigned to treatment in one experiment and only 10% of

subjects are assigned to treatment in another. The treated and control individuals would not be

comparable between studies, but they are comparable within studies. To prevent the different

treatment rates from influencing our estimates, we include study fixed effects along with study-

propensity interactions. This ensures that only the within-study variation (that which is randomly

assigned) contributes to our estimates.

effect increases by 0.4 percentage points. If we treat the relationship between these variables as truly linear, this suggests that GOTV treatments, on average, increase turnout by 4.1 percentage points for those whose propensity score is 2 standard deviations above the mean. However, these treatment effects are much weaker, only 2.5 percentage points, for low-propensity citizens. On average GOTV mobilization effects are more than 60 percent greater for the highest propensity individuals compared to the lowest. As a result, the typical treatment exacerbates the participation gap, despite the fact that GOTV interventions are often designed to reduce this gap. In the Supporting Information, we describe tests we conducted to ensure that our results are not driven by "deadweight" or other issues with the quality of data available on voter files.

What is the mechanism behind this effect?

Readers surprised by our results will naturally ask why these experiments tend to exacerbate the participation gap? Determining the mechanism is difficult to answer in any setting (Bullock, Green, and Ha 2010) but deserves attention nonetheless. Here, we provide several hypotheses and provide evidence that high-propensity individuals are easier to contact. First, the correlation of voting propensity with education and political knowledge might result in high-propensity citizens better understanding the treatment. Also, low-propensity citizens may have higher costs involved with voting so a treatment that reminds voters of the benefits of voting may have to be more powerful to stimulate low-propensity voters than high-propensity voters. Moreover, psychological, social, and economic differences between high and low-propensity citizens may explain why some people are simply more likely to comply with any policy or experimental intervention. Previous research shows that higher SES individuals are more likely to respond to many interventions including electoral reform (Berinsky 2005), public health campaigns (Pickett, Luo, and Lauderdale 2005), medical screenings (Wee et al. 2012), public housing (Blundell, Fry, and Walker 1988), and

Medicaid (Aizer 2003) even when such interventions are specifically designed to benefit low SES individuals.

A related potential mechanism lies in citizens' differential probabilities of being successfully contacted. High-propensity individuals may be more likely to read their mail, answer their phone, or talk to a canvasser. With our data, we can explicitly test whether high-propensity citizens are easier to contact via phone or door-to-door canvassing. For many studies including mail studies, the researchers is unable to know who actually received the treatment, but for phone and canvassing studies, the researcher can record which individuals or households were contacted. Table 5.4 presents the results of these test for all available studies. In each case, we regress a dummy variable for household contact on our propensity variable for the treatment group. The table shows that high-propensity individuals are much easier to contact via door-to-door canvassing or phone calls. For example, in Gerber and Green's (2000) New Haven study, a standard deviation increase in propensity corresponds to an extra 5 percentage point chance of canvassing contact and an extra 12 percentage point chance of phone contact. These results suggest that differential contact rates may explain much of the variation in intention-to-treat effects between high and low-propensity voters. If subsequent interventions hope to improve the participation and representation of low propensity voters, they must first overcome the challenge or reaching them in the first place.

Further information regarding the mechanisms behind our overall findings can be garners by examining variation in our interactive effects across different interventions. In the Supporting Information, we explore several sources of this variation. Consistent with Arceneaux and Nickerson (2009), we find that the exacerbating effect of GOTV treatment is greatest in low-salience elections. However, even in high-salience elections, these treatments, on average, increase disparities in participation. We also find that the most effective treatments in terms of increasing overall participation also exhibit the greatest exacerbating effects. Finally, we discuss the peculiar

Table 5.4. Are high-propensity citizens easier to contact?

	Study	Propensity	Constant
Door-to-Door Canvassing	GG00	.054(.007)**	.313(.007)**
	GGN03 - Bridgeport	.079(.018)**	.180(.015)**
	GGN03 - Columbus	.038(.013)**	.113(.011)**
	GGN03 - Detroit	.012(.006)*	.157(.007)**
	GGN03 - Minneapolis	.044(.009)**	.103(.009)**
	GGN03 - Raleigh	.019(.010)	.359(.012)**
	GGN03 - St. Paul	.051(.009)**	.127(.011)**
	NFK06	.021(.011)	.259(.010)**
	N08 - Denver	-.003(.008)	.220(.008)**
	N08 - Minneapolis	.035(.012)**	.300(.013)**
Phone Calls	GG00	.122(.007)**	.342(.007)**
	N06 - Albany	-.005(.018)	.616(.017)**
	N06 - Boston	.038(.014)**	.554(.014)**
	N06 - Stonybrook	.020(.012)	.886(.012)**
	NFK06	.025(.003)**	.310(.003)**
	N07 - Professional	.031(.003)**	.386(.003)**
	N07 - Volunteer	.060(.003)**	.422(.003)**
	NR10	.030(.002)**	.248(.002)**

*Robust standard errors in parentheses; ** significant at 1%, * significant at 5%*

For each experiment where contact information is available, we test whether contact rates are higher for high-propensity citizens by regressing contact on propensity for those individuals in the treatment group. The table shows that high-propensity citizens are much easier to contact via both door-to-door canvassing and phone calls. For example, in Gerber and Green's (2000) New Haven study, a standard deviation increase in propensity corresponds to an extra 5 percentage point chance of canvassing contact and an extra 12 percentage point chance of phone contact. These results suggest that differential contact rates explain much of the variation in intention-to-treat effects between high and low-propensity voters.

phenomenon that our few experimental settings where the participation gap was reduced involved largely African-American samples.

In one sense, why GOTV interventions tend to exacerbate the participation gap is of secondary importance. Whether the mechanism involves knowledge, costs, psychology, contact, or something else, the positive and normative implications of our study are the same. However, knowing that low-propensity citizens are harder to contact provides one promising avenue for future researchers to design interventions which may mobilize them. After all, reaching the population of interest may be enough to largely remove the exacerbating phenomenon that we identify. We hope that the continued application of our test on future experiments will allow us to understand the reasons why and the conditions under which campaign interventions exacerbate or reduce the participation gap.

Conclusion

In analyzing 26 field experiments, we find that over two-thirds of GOTV experiments mobilized high-propensity voters to a greater degree than low-propensity voters – thereby exacerbating the participation gap. Moreover, this exacerbating effect is statistically significant in 9 cases. Our pooled analysis demonstrates an average exacerbating effect that is substantively large and statistically significant. Most published GOTV methods developed by political scientists, many of which have been subsequently adopted by political campaigns (Green and Gerber 2008; Gerber et al. 2011; Issenberg 2010, 2012), exacerbate the disparities between voters and the voting-eligible population. Because turnout is politically consequential and because our *propensity* variable captures meaningful political differences between individuals and groups, this exacerbating effect should be of concern to both political scientists and practitioners.

These results increase pose and clarify a challenge to scholars and practitioners interested in political participation. Current mobilization methods employed by researchers and political campaigns appear to be mainly effective in bringing high propensity citizens to the polls. The search for a reliable mobilization method for bringing under-represented, low-propensity citizens to the polls continues, largely without resolution, and these results underscore the need to focus on this research moving forward. The method we present for assessing differential treatment effects provides a metric for future experimenters interested in developing and evaluating new methods for mobilizing these under-represented citizens.

Even for those uninterested in the participation gap *per se*, our findings hold practical implications for campaigns and scholars of voter mobilization. Some individuals, particularly those with low propensities, are harder to mobilize than others. As a result, the combinatorial effects of multiple GOTV efforts are called into question. Practitioners may wrongly assume that the effects of one mobilization will combine additively with the effects of others. However, once a campaign has hit a ceiling among high-propensity voters, it may be unable to further increase participation using traditional methods.

The findings of this paper also raise an ethical concern for experimenters and practitioners because experimental interventions and mobilization efforts are often conducted with the assumption that raising average participation levels can only be good for democracy. However, the evidence in this paper – that voter mobilization tends to exacerbate existing inequalities in the electorate – necessitates a more nuanced perspective. Despite good intentions, current GOTV efforts are not the solution to persistent inequalities in the political process. On the contrary, these efforts may contribute to the problem by making the electorate more polarized and less representative of the greater population.

Appendix for Study #1

Coding of Bendigo Turnout in the 1899 Referendum

The 1899 electoral rolls from Bendigo provide the name, occupation, address, property value, and ownership status of every registered voter in the town. I digitized these records by manually transcribing the information and double checking for accuracy. There are three different records, one for each of Bendigo's three wards. I removed all women from the record since they were not legally allowed to vote in the statewide election. Likely females were identified by first name. Also, the record typically did not list an occupation for women, which helped to resolve the registrants' gender in ambiguous cases.

The record of citizens who voted in the 1899 Referendum provides the name, address, town, and occupation of all Victorian men who voted in the election. This record is available in the State Library of Victoria and is available for purchase on cd-rom from MacBeth Geneology. The record includes over 3000 hand-written pages.

After digitizing the records of Bendigo registrants, I manually searched for each male registrant in the record of voters. In the simplest cases, matches were made by name and town. For example, there is one person named Benjamin Aarons in the Bendigo electoral rolls. I did not find a Benjamin Aarons from Bendigo in the list of voters, so I coded this individual as having abstained. Similarly, there is one person names James Abberton. I did find one James Abberton from Bendigo on the list of voters, so he is coded as having voted. When available, I recorded middle initials and suffixes (i.e. Jr.), but Thomas J. Alderson would still be coded as having voted if I found a Thomas Alderson (with no middle name or initial) from Bendigo in the list of voters.

When multiple Bendigo registrants had the same name, the coding was more complicated. To simplify this problem, I removed the 7 last names from the analysis where more than 20 Bendigo registrants shared the name (Brown, Jones, Miller, Roberts, Smith, Thomas, Williams). If I attempt

to code these names as well, the results are unchanged. Suppose that two Bendigo registrants share the same name, Arthur Armstrong. I would search for an Arthur Armstrong from Bendigo in the list of voters. If I found two, then I coded both as having voted. If I found none, then both were coded as having not voted. If I found 1, I would compare middle initials, suffixes, occupations, and street names to determine which one voted and which one did not.

In some rare cases, the same voter appears to be listed in the Bendigo rolls twice. For example, someone might reside in two residences in different wards, appearing on the roll in two separate wards. In these cases, I deleted one of the records. As a rule, I deleted the observation where the address did not match with the list of voters, assuming that this record constituted a secondary residence.

Importantly, I prevented myself from seeing the data on property ownership or property value when making these decisions, so I could not have influenced the results by adjusting my coding decisions to fit the data. This data was hidden and merged in after all coding was complete. As a result, any mistakes in coding will likely diminish rather than augment any differences in turnout across property ownership and property values.

More Details Regarding Standard Errors in the Analysis of State Assembly Elections

Statistical inferences are tricky with difference-in-difference designs, particularly when the number of treated units is small (Bertrand et al. 2004). In this section, I focus on the 9 percentage point estimate of the effect of compulsory voting on Labor vote share (Model 3 in Table 1.2 of the original paper), and I discuss various approaches to estimating the standard errors. Table A1.1 presents the standard error (point estimate = .092) estimated in 6 different ways.

The first row presents the traditional OLS standard errors. This approach could over-estimate the precision of the estimate under the likely possibilities of serial correlation and heteroskedasticity. The second row corrects for heteroskedasticity using the Huber-White approach. The third row allows for state-specific serial correlation and heteroskedasticity, estimating state-clustered standard errors. Because this is the typical approach used for panel designs, I report these standard errors throughout the paper. However, this approach can be misleading if the number of clusters is too small.

The fourth row presents the result of a non-parametric bootstrap. With this approach, I randomly sample 85 observations from my data set with replacement (85 is the number of elections in the original analysis) and estimate the effect of compulsory voting with that new sample. I repeat this procedure 10,000 times and estimate the standard error as the standard deviation of these bootstrap estimates. The block bootstrap, shown in row 5, is identical to the non-parametric bootstrap with one key exception. Instead of sampling observations independently, I sample states as blocks. This approach accounts for the possibility that we do not have 85 independent observations but 6 independent states.

Lastly, I design a random permutation test specifically for the problem at hand. For each state, I randomly choose a year between 1910 and 1950 (drawn from a uniform distribution) and assign that year as the adoption of compulsory voting in that state. Then I estimate the placebo

effect of compulsory voting under this hypothetical scenario. I repeat this procedure 100,000 times and report the standard error as the standard deviation of those permutation estimates. Figure A1.1 presents the results of those random permutations. We see that the estimates are normally distributed with mean zero and standard deviation of .035. The estimated effect from the real data is shown as a red line. Less than 0.5% of the random permutations resulted in an estimate larger than the actual estimate of .092, suggesting that our estimate is much larger than we would expect to see under the null hypothesis where compulsory voting has no effect.

The estimated effect of compulsory voting on Labor vote is statistically significant ($p < .05$) for all approaches to estimating the standard errors. Due to the difficulty in making statistical inferences with this type of data, I would not recommend focusing on any one particular approach. Instead, I would emphasize that every approach yields the same inference: this effect is much larger than we would expect by chance alone.

The figure shows the results of a random permutation test, showing the statistical significance of the difference-in-difference estimate of the effect of compulsory voting on Labor voteshare. For each state, I randomly choose a year between 1910 and 1950 (drawn from a uniform distribution) and assign that year as the adoption of compulsory voting in that particular state. Then I estimate the placebo effect of compulsory voting under this hypothetical scenario. I repeat this procedure 100,000 times. The figure presents the results of those random permutations. We see that the estimates are normally distributed with mean zero and standard deviation of .035. The estimated effect from the real data is shown as a red line. Less than 0.5% of the random permutations resulted in an estimate larger than the actual estimate of .092, suggesting that our estimate is much larger than we would expect to see under the null hypothesis where compulsory voting has no effect.

Table A1.1. Standard Error on the Estimated Effect of Compulsory Voting on Labor Vote

Traditional	(.028)**
Heteroskedasticity Robust	(.031)**
State-Clustered	(.033)*
Non-parametric Bootstrap	(.040)*
Block Bootstrap	(.046)*
Random Permutation	(.035)*

Figure A1.1. Random Permutation Results
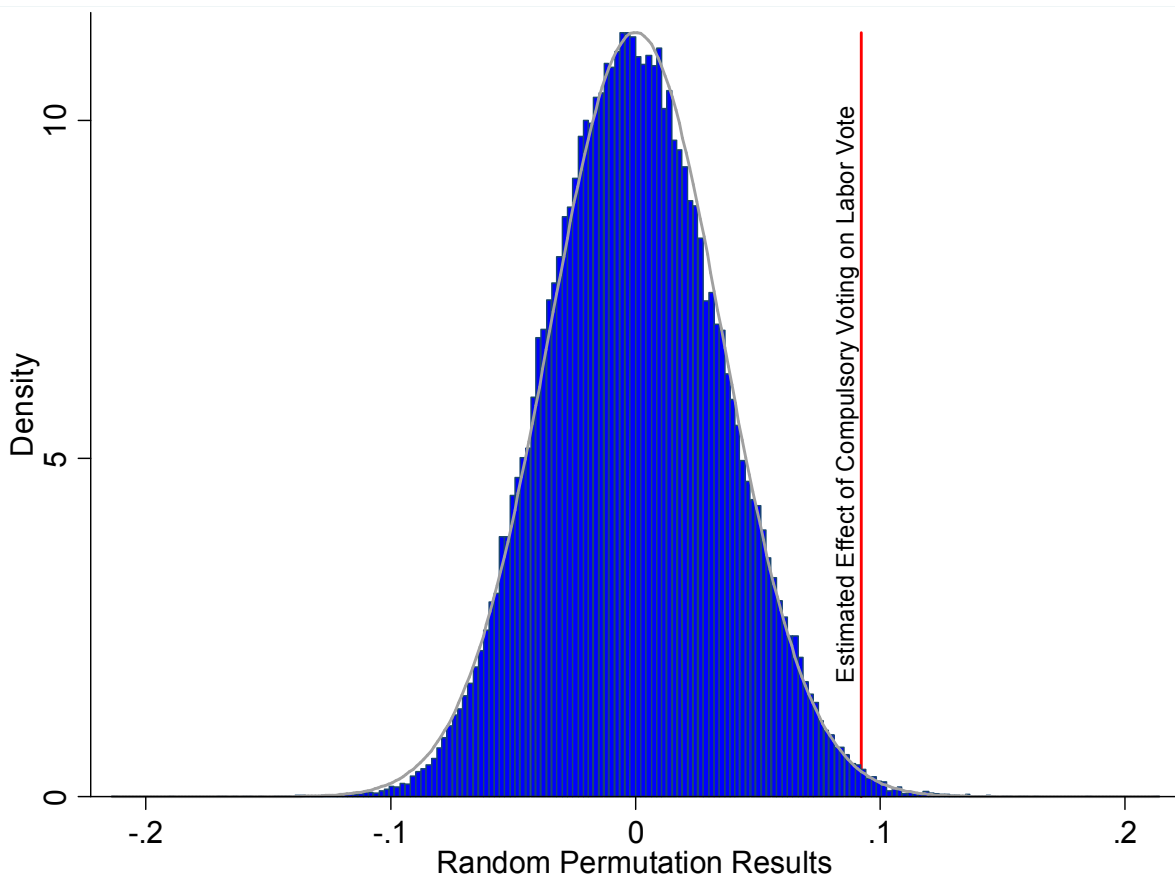
Robustness of State Assembly Results to Alternative Specifications

The table demonstrates the robustness of the results to various specifications. Rows 1 and 2 are copied from Column 3 and 4 in Table 1.2 of the original paper. This models estimate the effect of compulsory voting (coded from 0 to 1) on Labor Party voteshare (also coded from 0 to 1). Both estimates include state fixed effects and year fixed effects, and the latter allows for state-specific trends. The next 4 rows model time in different ways. Although these estimates are slightly smaller than in Row 1, the estimates are substantively and statistically significant for all model specifications. All models include state fixed effects.

Row 3 models the time trend linearly. Row 4 includes year and year squared, modeling the time trend as a second order polynomial. Row 5 also includes year cubed, modeling the time trend as a third order polynomial. Lastly, Row 6 includes different linear time trends before and after the adoption of compulsory voting. Compulsory voting may have changed a state's trend in Labor vote share in addition to its mean. This model mimics a regression discontinuity design where we estimate the expectation of Y at a threshold by local linear regression. This model estimates the expected jump in Labor vote share immediately after a state adopts compulsory voting. All standard errors are clustered at the state level.

Table A1.2. Robustness Checks

Specification	Estimate	Standard Error
Year Fixed Effects	.092	.033
State-Specific Time Trends	.093	.040
Linear Time Trend	.062	.028
Quadratic Time Trend	.061	.022
Cubic Time Trend	.057	.032
Different Time Trend after CV	.060	.023

More Details on the Synthetic Control Analysis

The paper employs the synthetic control method of Abadie, Diamond, and Hainmueller (2010) to estimate the effects of compulsory voting at the federal level on turnout pension spending. I estimate that compulsory voting increased voter turnout by 18.6 percentage points and pension spending by 0.41 percentage points of GDP. Below I provide more details on these estimates.

The synthetic control groups for each test are the weighted combinations of 20 other OECD countries that best match Australia in 1910 and 1920 (before compulsory voting) in terms of turnout, pension spending, and trends in these variables. Table A1.3 presents the weights estimated for each country for the two separate analyses. 10 comparison countries are excluded from the turnout analysis because turnout data is unavailable or missing, while all 20 comparison countries are included in the pension analysis. New Zealand, France, Canada, and the United Kingdom receive the greatest weights in the turnout analysis. Denmark and New Zealand receive the greatest weights in the pension analysis.

To address statistical significance of the findings, I conduct placebo tests in the other nations where there was no policy change between 1920 and 1930. For each nation, I construct a synthetic control unit, excluding Australia as a potential control country. Then, pretending that a policy change tool place between 1920 and 1930 in that country, I calculate the difference-in-difference estimate. The table presents these placebo estimates for each country. Again, only 10 comparison countries are analyzed for turnout while 20 countries are analyzed for pension spending. In both cases, only one country's absolute value is as great as the estimate in Australia, suggesting that the change in Australia's turnout and pension spending between 1920 and 1930 is statistically and substantively larger than we would expect to see by chance alone.

Table A1.3. Country Weights in Synthetic Control Analyses

Country	Turnout Analysis	Pension Analysis
Argentina		.002
Austria		.002
Belgium		.004
Brazil		.002
Canada	.100	.002
Denmark	.056	.741
Finland	.061	.002
France	.109	.009
Greece		.002
Italy		.002
Japan		.002
Mexico		.002
Netherlands	.055	.009
New Zealand	.427	.187
Norway	.031	.006
Portugal		.002
Spain		.003
Sweden	.015	.005
United Kingdom	.085	.014
United States	.060	.002

Table A1.4. Placebo Tests for the Effect of Compulsory Voting on Pension Spending

Country	Turnout Estimates	Pension Estimates
Australia	**.186**	**.414**
Argentina		-.069
Austria		.437
Belgium		.140
Brazil		-.069
Canada	-.251	.151
Denmark	.049	.090
Finland	-.155	-.069
France	-.018	.022
Greece		-.036
Italy		-.069
Japan		-.069
Mexico		-.069
Netherlands	.080	-.197
New Zealand	.074	.198
Norway	.013	-.063
Portugal		-.069
Spain		-.054
Sweden	.060	.192
United Kingdom	.095	-.052
United States	.074	-.069

Table A3.1. Sample Summary Statistics

| Variable	Mean	Median	Min	Max	St Dev	Observations
Population 2000	12198	9517	413	49462	9706	325
Population Density	101.7	51.1	0.9	1743.3	180.7	296
Pct Some High School	11.2	9.4	1.2	55.0	7.4	325
Pct Some College	4.3	3.5	0.4	28.4	3.2	325
Pct ≤ Min. Wage	16.6	13.4	1.5	58.0	10.7	325
Pct > 5x Min. Wage	2.5	1.8	0	14.3	2.2	325
Pct > 10x Min. Wage	0.9	0.6	0.0	5.6	0.9	325
Pct Agriculture	51.3	51.1	1.3	96.4	22.7	325
Pct Construction	10.3	9.4	0.2	34.1	6.1	325
Pct Finance	0.1	0.1	0	1.8	0.2	325
Pct Health	1.3	1.0	0	8.4	1.1	325
Pct Real Estate	0.1	0.0	0	0.8	0.1	325
Pct Transportation	2.4	1.8	0	17.9	2.4	325
Pct Commerce	10.3	10.0	0	40.7	5.9	325
Pct Education	4.2	3.3	0	21.3	3.3	325
Pct Government	3.1	2.8	0.3	14.8	1.9	325
Pct Manufacturing	15.6	12.0	0.7	57.0	12.5	325
Pct Mining	.7	.1	0	27.0	2.4	325
Pct > 70 Years Old	1.5	1.3	0.3	5.1	0.7	325
Pct < 18 Years Old	44.5	44.9	27.8	56.8	5.2	325
Pct Catholic	91.2	93.4	51.7	99.8	8.2	325
Pct PAN	29.7	31.1	2.5	65.2	14.5	325
Pct PRI	52.2	51.9	14.3	82.6	12.6	325
Pct PRD	18.1	15.7	0.6	75.1	13.5	325
Turnout 1991	44.0	44.3	11.4	98.3	14.5	322
Turnout 1994	64.0	65.1	19.5	90.9	12.4	323
Turnout 1997	51.4	50.8	12.7	94.8	12.6	325
Turnout 2000	61.5	61.8	27.0	96.9	10.2	325
Turnout 2003	45.8	43.9	1.2	97.7	14.8	322
Turnout 2006	65.0	64.7	33.8	96.7	9.9	320
Turnout 2009	57.0	54.5	12.3	99.4	16.0	318
Fiesta	0.16	0	0	1	0.37	325

Data on population density is missing for 29 municipalities. Turnout data is missing for a few municipalities in 1991, 1994, 2006, and 2009.

Table A3.2. Placebo Regressions to Test for Balance

Dependent Variable	Coefficient	Standard Error	P-value
Log Population	-.120	.097	.217
Log Pop. Density	-.007	.016	.664
Pct Some High School	-1.272	.683	.063
Pct Some College	-.409	.283	.150
Pct ≤ Minimum Wage	.765	1.150	.507
Pct > 5x Min. Wage	-.479	.182	.009
Pct > 10x Min. Wage	-.168	.080	.036
Pct Agriculture	3.164	2.394	.187
Pct Construction	.303	.679	.655
Pct Finance	-.018	.022	.409
Pct Health	-.138	.103	.181
Pct Real Estate	-.017	.012	.165
Pct Transportation	-.261	.197	.186
Pct Commerce	-.712	.624	.255
Pct Education	-.320	.324	.324
Pct Government	-.524	.164	.002
Pct Manufacturing	-1.236	1.337	.356
Pct Mining	-.150	.155	.333
Pct > 70 Years Old	-.003	.070	.968
Pct < 18 Years Old	.413	.500	.410
Pct Catholic	-.010	.009	.279
Pct PAN	-2.730	1.241	.029
Pct PRI	.058	1.142	.959
Pct PRD	2.672	1.296	.040

Each row represents a placebo regression where the dependent variable is regressed on the fiesta treatment, state turnout, and year fixed effects, closely mirroring the model in Column 1 of Table 3.1. The coefficient on the fiesta treatment is shown, along with its standard error (clustered by municipality). The table shows that the fiesta treatment is not meaningfully correlated with these demographic, economic, or political variables. To the extent that we do see statistically significant differences (earnings, government employment, and partisan support), these variables may not represent imbalance but actually downstream effects of fiestas and their negative effect voter turnout. Despite these possibilities of post-treatment bias, Table A3.3 shows that the empirical results are robust to the inclusion or exclusion of all of these covariates.

Figure A3.1. Distribution of Propensity Score for Both Comparison Groups
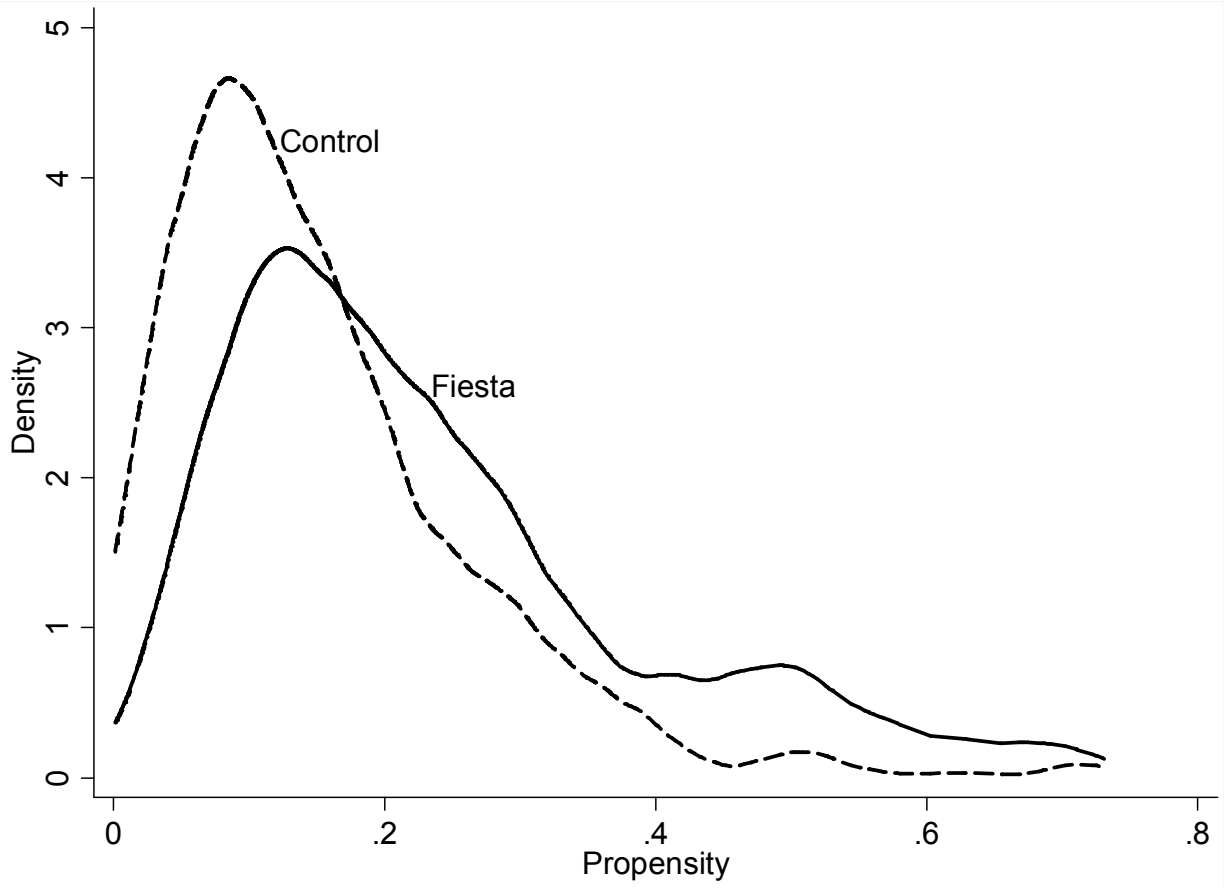
The distributions are estimated by a kernel density function. The propensity scores are estimated from a Logit model which regresses Fiesta on all census variables and year fixed effects. To avoid over-fitting, we using a jack-knife estimator which does not use observation "i" in determining the propensity score for that particular observation. Note that some of the variables included in the propensity model could actually be influenced by the treatment and the subsequent drop in turnout, so this graph may actually understate the comparability of the treated and untreated municipalities.

Table A3.3. Difference in Mean Voter Turnout by Year

Year	Fiesta w/in 2 Weeks?		Difference	Standard Error	P-Value
	No	Yes			
1991	44.76	40.22	-4.54	1.97	.024
1994	64.52	61.06	-3.47	1.94	.078
1997	51.92	48.53	-3.39	1.91	.080
2000	62.06	58.87	-3.19	1.75	.072
2003	46.58	41.93	-4.65	1.73	.008
2006	65.60	62.21	-3.40	1.68	.048
2009	57.62	53.55	-4.08	2.36	.088
Pooled	56.16	52.27	-3.89	1.15	.001

The table reports mean levels of voter turnout for our treated and untreated municipalities in each election. For each individual year, turnout in the treated municipalities is at least 3 percentage points lower, so our results are not sensitive to the inclusion or exclusion of particular years. Also, our results are not sensitive to modeling assumptions, because a simple difference in means yields the same estimate as the more complex regression models. For individual years, the standard errors allow for unequal variance between treatment groups (Huber-White standard errors), and for the pooled difference, the standard error allows for municipality specific correlation (municipality-clustered standard error).

Table A3.4. The Effect of Fiestas on Voter Turnout: Robustness Checks

Model	Effect of Fiesta Treatment	Standard Error	P-value
No Controls	-3.452	1.038	.001
Without Year Fixed-Effects	-3.466	1.034	.001
Without State Turnout	-3.817	1.152	.001
State Fixed Effects	-2.322	.951	.015
3 Population Variables	-3.625	1.045	.001
7 Education Variables	-3.320	1.024	.001
10 Wage Variables	-3.620	1.017	<.001
18 Economic Variables	-3.004	.969	.002
10 Age Variables	-3.637	1.067	.001
3 Partisan Variables	-3.180	1.012	.002
5 Miscellaneous Variables	-3.283	1.053	.002
All Controls	-3.122	.951	.001
Propensity Score Matching	-3.931	1.125	<.001
Nearest Neighbor Matching	-3.442	.777	<.001

All standard errors are clustered by municipality. The models above include year-fixed effects and state turnout, except rows 2 and 3 which exclude one at a time. Both matching estimators match on year exactly and include a linear bias adjustment. For both matching estimators, we report the standard errors proposed by Abadie et al. (2004). With propensity score matching, this procedure does not account for the uncertainty associated with the estimation of the propensity score. Many researchers bootstrap the standard errors in this case to account for such a problem. However, Abadie and Imbens (2008) show that the bootstrap does not yield valid standard errors for matching estimators. They argue that the closed-form standard errors are adequately conservative to account for uncertainty associated with the propensity score. For our pooled OLS and fixed effects estimates, we report municipality-clustered standard errors. In this case, they are nearly identical to standard errors calculated from a non-parametric bootstrap.

Figure A3.2. Robustness of the Fiesta Effect to Different Window Sizes
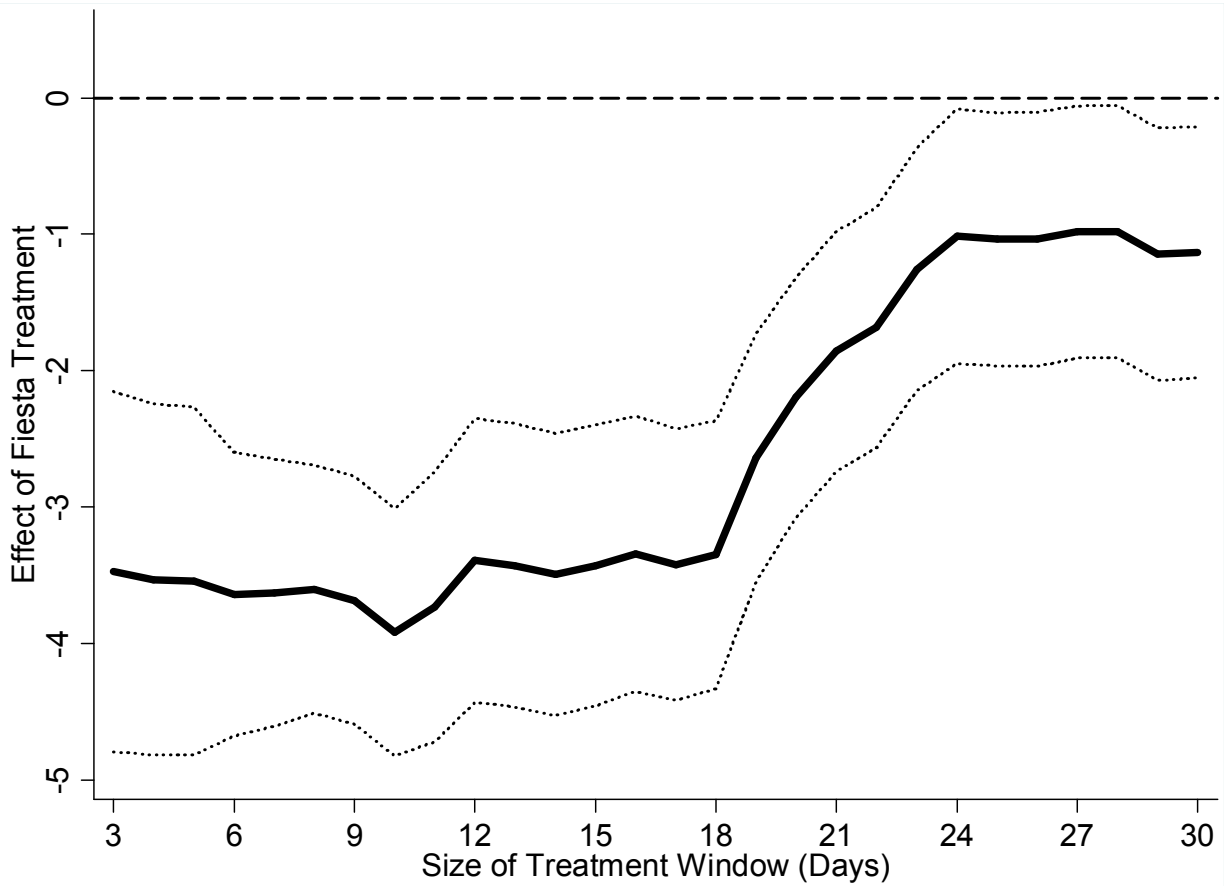
The figure presents the robustness of our main empirical result across different specifications of the fiesta treatment. Each point represents the coefficient on the fiesta treatment in a regression similar to Column 1 in Table 3.1. On the X axis, we vary the coding of the fiesta treatment. In the paper, we code the treatment to equal 1 if the fiesta occurs within 14 days of the election, but we can see that our result would have been nearly identical had we used any cutoff between 3 and 18 days. As expected, the treatment effect decays for very large windows, but our findings are not sensitive to an arbitrary choice about the size of the treatment window. The dotted lines indicate municipality-clustered standard errors.

Our previous analysis has focused solely on Mexican municipalities with just one Catholic Church. Thus, our data set consists of primarily poor, rural, agricultural communities. The questions remain whether our results will generalize to other democratic communities. In order to ensure internal validity, we have limited our study to the subset of regions for which we can make valid inferences. To test for external validity, we turn to the Mexican city of Monterrey. This urban center looks nothing like the previous communities in our data set. Monterrey has over one million residents, a well regarded health care system, four major universities, and a GDP per capita of more than 45,000 U.S. dollars.

For every church in Monterrey, the local archdiocese website lists the saint's day fiesta date and the neighborhoods served by the church. In many cases, we were able to match these listed neighborhoods to "secciones," the smallest geographic unit for which electoral results were reported in 2006 and 2009. As a result, we have compiled a dataset of 584 neighborhoods served by 93 different churches for which we know the fiesta date and voter turnout levels for 2006 and 2009. With only 93 churches, two election years, and no census variables our tests will be much less precise than in our previous analysis. However, we can test for the fiesta effect just as before to see if fiestas have the same type of effect in an urban area.

The table below presents the results of this analysis. We conduct OLS to test whether neighborhood turnout changes when the church's saint's day fiesta occurs within two weeks of the election date. Both models control for the year, and model 2 includes partisan controls. The partisan controls are the proportion of the vote earned in each neighborhood by the three major political parties - the National Action Party, the Institutional Revolutionary Party, and the Party of the Democratic Revolution. Since we do not have census data for each neighborhood, we rely on these partisan variables to serve as a proxy for the social structure and unobserved characteristics of

neighborhoods. These controls simply explain some of the variance in turnout and allow for a more efficient estimate.

As with the rural municipalities, we estimate a negative effect of the saint's day fiesta on turnout in Monterrey. Model 2 estimates that the occurrence of a saint's day fiesta within 2 weeks of an election reduces neighborhood turnout levels by 2.9 percentage points ($p < .01$). While the evidence is not as strong as before due to our lack of data, this analysis suggests that our findings in rural Mexico may apply to a much broader set of democratic communities. Saints Day Fiestas decrease voter turnout in Monterrey just as they do in rural communities, and the same mechanisms by which social capital decreases turnout in rural communities are present in urban centers.

Table A3.5. The Effect of Fiestas on Turnout in Monterrey

	(1)	(2)
Fiesta	-0.782	-2.931
	(1.610)	(0.738)**
Year Fixed Effects	X	X
Partisan Controls		X
Observations	1160	1160
R-squared	.07	.48
SER	8.99	6.74

*Church-clustered standard errors in parentheses; ** significant at 1%*

Appendix for Study #4

Randomization Procedures

In order to ensure greater balance between treatment groups and achieve greater precision with our subsequent estimates, we first stratified the population of phone numbers and then randomly assigned treatment conditions within each stratum. Based on previous literature, previous turnout would likely be a good predictor of turnout in the special election. Therefore, we divided phone numbers into three categories: (1) those where at least one person had voted in a low-salience special election in 2009, (2) those where at least one person had voted in the 2010 general election but no one voted in 2009, and (3) numbers where nobody voted in either election.

Next we divided phone numbers into three additional categories depending on the number of voters listed: (1) 1 registered voter at that number, (2) 2 registered voters, and (3) 3 or 4 voters. Remember that we omitted numbers with more than four voters.

Then, we decided that town of residence may be a good predictor of turnout in the special election. In this district, roughly one-third of the voters live in Charlton, one-third live in Southbridge, and one-third are distributed across three smaller towns, East Brookfield, Oxford, and Spencer. We divided the phone numbers into three categories accordingly. Post-election interviews with the candidates revealed that their voter contact strategies were heavily influenced by town-based geographic strategy (author interviews, July and August 2011), so stratifying on town was likely a good decision on our part.

Lastly, we stratified based on our data collected from robo-calling. To ensure that similar numbers of calls in each condition are answered, we separated the numbers where a live person had answered the phone from all others.

As a result of these categorizations, an individual phone number would fall into one of 54 unique categories (i.e. 1 registered voter, voted in 2010, lives in Southbridge, failed to answer robo-

call). Since we expected many calls to be unanswered, we wanted to ensure that each stratum had at least 100 phone numbers. Most of the 54 categories had more, but we combined some of the smaller categories to meet this criterion. In 9 cases, we combined categories that shared everything in common except their answers to the robo-call. In 1 additional case, we also combined categories from different towns. This resulted in 44 unique strata with at least 100 phone numbers. Within each strata, we randomly assigned one-third to the reminder condition, one-third to the pivotal condition, one-thirtieth to the survey condition, and the remainder to the no contact condition; always rounding to the nearest integer when necessary.

This procedure generates near-perfect balance across previous turnout, town, household size, and propensity to answer the phone. Moreover, random assignment guarantees balance over all observable and unobservable pre-treatment characteristics in expectation. Nonetheless, there could be slight imbalances by chance between conditions over other variables of interest. To avoid this, we generated 500 different randomization schemes according to the protocol above and then selected the one scheme that generated the best balance on other variable such as party registration, Hispanic ethnicity, and turnout in other elections.

Wording of Scripts Given to Callers

Reminder Treatement:

Hello, this is XXXX from Harvard University. We are calling registered voters to provide

information about an upcoming election in your town.

Am I speaking with XXXXXXXX?

Did you know that there is a special election coming up?

[IF YES] Do you know when it is?

We just want to remind you that there's a special election on Tuesday, May 10th to fill the seat of

your representative in the Massachusetts State House. For more information on the election you can

visit the website of the Secretary of the Commonwealth.

Goodbye

Pivotal Treatement:

Hello, this is XXXX from Harvard University. We are calling registered voters to provide

information about an upcoming election in your town.

Am I speaking with XXXXXXXX?

Did you know that there is a special election coming up?

[IF YES] Do you know when it is?

We just want to remind you that there's a special election on Tuesday, May 10th to fill the seat of

your representative in the Massachusetts State House. For more information on the election you can

visit the website of the Secretary of the Commonwealth.

The reason that there is a special election is that the last election ended in an exact tie. Had one

more or one less person voted in the last election, your candidate would have won. The special

election on Tuesday is likely to be close again, so there is a high chance that your vote could make a difference.

Goodbye

Table A4.1. **Mean Turnout among each Subset of Experimental Subjects**

	Reminder	Pivotal	Obs	P-value
Intention-to-Treat	.311	.318	5771	.467
Contacted Individuals	.456	.472	489	.615
Contacted, Uninformed Individuals	.180	.237	167	.209
Contacted, Uninformed, voted in > 2 recent elections	.364	.532	77	.047
Contacted, Uninformed, voted in ≤ 2 recent elections	.022	.033	90	.652

<p align="center">**Table A4.2. Placebo Tests**[2]</p>

	Dependent Variable								
	Voted '10	Voted '09	Voted '08	Hispanic	Age	Dem	Rep	Absentee	Unmatched
Intention-to-Treat	.000	.002	.001	-.008	.136	-.001	-.002	-.002	.001
	(.007)	(.004)	(.010)	(.007)	(.352)	(.010)	(.007)	(.003)	(.034)
Contacted Individuals	.012	-.001	-.037	-.017	-.974	-.048	.032	-.005	-.001
	(.023)	(.017)	(.030)	(.013)	(1.062)	(.031)	(.021)	(.024)	(.022)
Contacted, Uninformed Individuals	.052	.029	-.019	-.048	1.842	-.053	-.013	.000	.002
	(.042)	(.026)	(.051)	(.021)*	(2.009)	(.056)	(.037)	(.006)	(.015)
Contacted, Uninformed, voted in > 2 recent elections	.071	.054	.048	-.033	-1.054	-.138	.076	.019	.002
	(.054)	(.046)	(.087)	(.043)	(2.743)	(.094)	(.050)	(.017)	(.024)
Contacted, Uninformed, voted in ≤ 2 recent elections	.039	.009	-.005	-.020	4.819	.051	-.067	-.013	-.003
	(.048)	(.016)	(.045)	(.043)	(2.796)	(.073)	(.054)	(.012)	(.016)

Standard errors in parentheses; * p < .05

---

[2] These results arise from 45 separate regressions which follow the same specification as those in table 4.3. The only exception is that the dependent variable is a pre-treatment variable, so we expect (and find) these coefficients to be close to zero.

Appendix for Study #5

Can the findings be explained by data limitations?

We may worry that our finding is a result of poor data quality from the public voter files typically employed in field experiments. Public voter records are often inaccurate and out of date. For example, an individual may have passed away or moved and could not possibly vote in an upcoming election, but a researcher would have no way of knowing this. We refer to these individuals as "deadweight," because they shouldn't be on the voter file at all, but the researcher hopelessly tries to mobilize them. Deadweight could be particularly troubling for our study if these ineligible individuals tend to be classified as low-propensity. What if the treatment effect is actually homogeneous across the eligible population, but many individuals categorized as low-propensity are actually deadweight, leading us to falsely conclude that GOTV interventions exacerbate the participation gap?[3]

To address this concern, we perform a simple sensitivity analysis. We cannot entirely rule out concerns about deadweight, but we can determine how extensive the problem would have to be in order to drive our results. Pooling all experiments in our sample and including study fixed effects, we estimate an average treatment effect of 3.7 percentage points. Then, we break the sample into 20 subsamples according to each individual's propensity score.[4] As expected, the conditional average treatment effect is larger for the higher propensity subsamples. The largest treatment effect we

---

[3] From the perspective of the participation gap, the reason that an individual cannot be mobilized is highly relevant. If a person is deceased, then they truly should not be in the sample. However, if the tendency to move and be unreachable by a political campaign is correlated with turnout as well as demographics and policy preferences, then the systematic tendency of GOTV treatments to miss these individuals will increase the participation gap.

[4] Our general results are robust to different numbers of subsamples.

observe for a subsample is s 5.2 percentage points. If the treatment effect were truly homogeneous and our result were driven by deadweight in the voter file, then we could say that the true treatment effect for non-deadweight individuals in each sample would have to be at least 5.2 percentage points. Therefore, the minimum proportion of deadweight in our sample would have to be 30% (1 - .037/.052) in order for us to obtain the results that we do. Put another way, deadweight could only explain our results if deadweight individuals constitute at least 30% of these GOTV samples.

While voter records surely contain errors, this 30% figure is implausibly large.[5]   We demonstrate this by estimating the proportion of deadweight on a typical voter file. Campaigns and for-profit data vendors have a strong incentive to identify and remove deadweight from the file because targeting deadweight is costly. According to the data base of Catalist, a widely used political data services company based in Washington D.C., only 4% of the individuals on statewide voter files are classified as "deceased" or having a "bad address." Focusing specifically on those individuals who voted in the most recent election, that number shrinks to 2.5%. Also, large-scale mail based surveys in Florida and Los Angeles County suggest that less than 10% of individuals on voter files are deadweight (Ansolabehere et al. 2010). Our sensitivity analysis indicates that data quality cannot reasonably be argued to explain our results, because the actual amount of bad data is much less than it would have to be in order to pose a threat to our inferences.

---

[5] This is especially true in cases where the researchers selectively chose their sample to minimize deadweight. For example, Gerber, Green, and Larimer (2008) only include individuals who voted in 2004 in their sample, so the proportion of their sample that could have moved or passed away in the two year interim period is small.

Variation across Different Experimental Interventions

Having seen that GOTV interventions tend to exacerbate the participation gap, on average, we would like to know whether this effect varies across different contexts, settings, or mobilization methods. Our strategy allows us to assess this variation. By conducting our test across many experiments, we hope to identify the types of interventions that are most effective (or least ineffective) in reducing the participation gap. Of course, even though this is the largest analysis of different types of interventions yet undertaken, the sample size prevents us from making strong claims about variation across different interventions. Nevertheless, we hope that our test here will guide future researchers in applying this test to identify the types of treatments that might effectively reduce the participation gap.

First, we test for variation across electoral salience. Arceneaux and Nickerson (2009) argue that high-propensity voters will be easier to mobilize in low salience elections and low-propensity voters will be easier to mobilize in high salience elections. To test this hypothesis, we compare empirical results across elections with different levels of voter turnout. Figure A5.1 plots the interactive coefficient from our regressions against the constant coefficients from the same regressions. The constant term indicates the predicted probability of turnout for the average subject in the control group. The interactive term indicates the extent to which the treatment exacerbated the participation gap. So, by looking for a relationship between these coefficients, we can see if interventions are more likely to increase the participation gap in low or high salience elections

The hypothesis of Arceneaux and Nickerson is largely confirmed: as electoral salience increases, the exacerbating effect of GOTV interventions decreases. However, looking at the graph, we would only expect an intervention to decrease the participation gap in elections with turnout of 50% or greater. Elections with 50% turnout are rare in the U.S. outside of presidential races. This

analysis suggests that GOTV interventions *can* actually reduce the participation gap in very high-salience elections, but these interventions have the opposite effect in most settings.

Next, we test for variation across the strength of treatment. We quantify severity using our treatment coefficient, the effect of the treatment for the average subject in the sample. Figure A5.2 plots the interactive coefficient against the additive coefficient from our analyses. We see that as the severity of the experimental treatment increases, the exacerbating effect also increases. One possible explanation is that the most effective treatments tend to have a psychological or social pressure component to them, as opposed to the more traditional GOTV messages. This psychological component which makes these interventions so effective may have a particularly concentrated effect among high-propensity citizens. As a result, the most effective interventions have the unexpected consequence of exacerbating the participation gap.

Only two interventions in our analysis demonstrate statistically significant evidence that the participation gap was reduced. What might explain the difference in these two cases? One intriguing similarity between the two experiments with negative interaction effects is that they both targeted citizens in communities with large African American populations. One explicitly targeted African Americans (Middleton and Green 2008) and the other was set in the largely African American city of Detroit[6] (Gerber, Green, and Nickerson 2003).

We tested for the possibility that African Americans respond differently to GOTV experiments by examining field experiments for which both Blacks and non-Blacks are identified in the experimental population. Most public voter files do not identify the race of the voters. As such, there are only three studies in our sample that could be used for this purpose: Dale and Strauss (2009); the Gerber, Green, and Nickerson (2003) study in Raleigh; and Nickerson and Rogers (2010). In these studies, Black citizens do appear to respond differently to GOTV efforts than non-

---

[6] Detroit was 83% African American in 2010.

Black citizens. In two studies, we find that high-propensity Blacks are actually *demobilized* by the treatment and in the other they were mobilized much less than whites or low-propensity Blacks. If GOTV efforts are demobilizing likely African American voters, this undermines the purpose of many efforts, in addition to presenting serious ethical concerns. Of course, this is a preliminary analysis on only three existing data sources so we draw no strong conclusions about these mechanisms. Moreover, we would expect to see some negative interactions by chance alone, so we should not draw strong conclusions from the few cases where we see this.

Figure A5.1. Variation across Electoral Salience
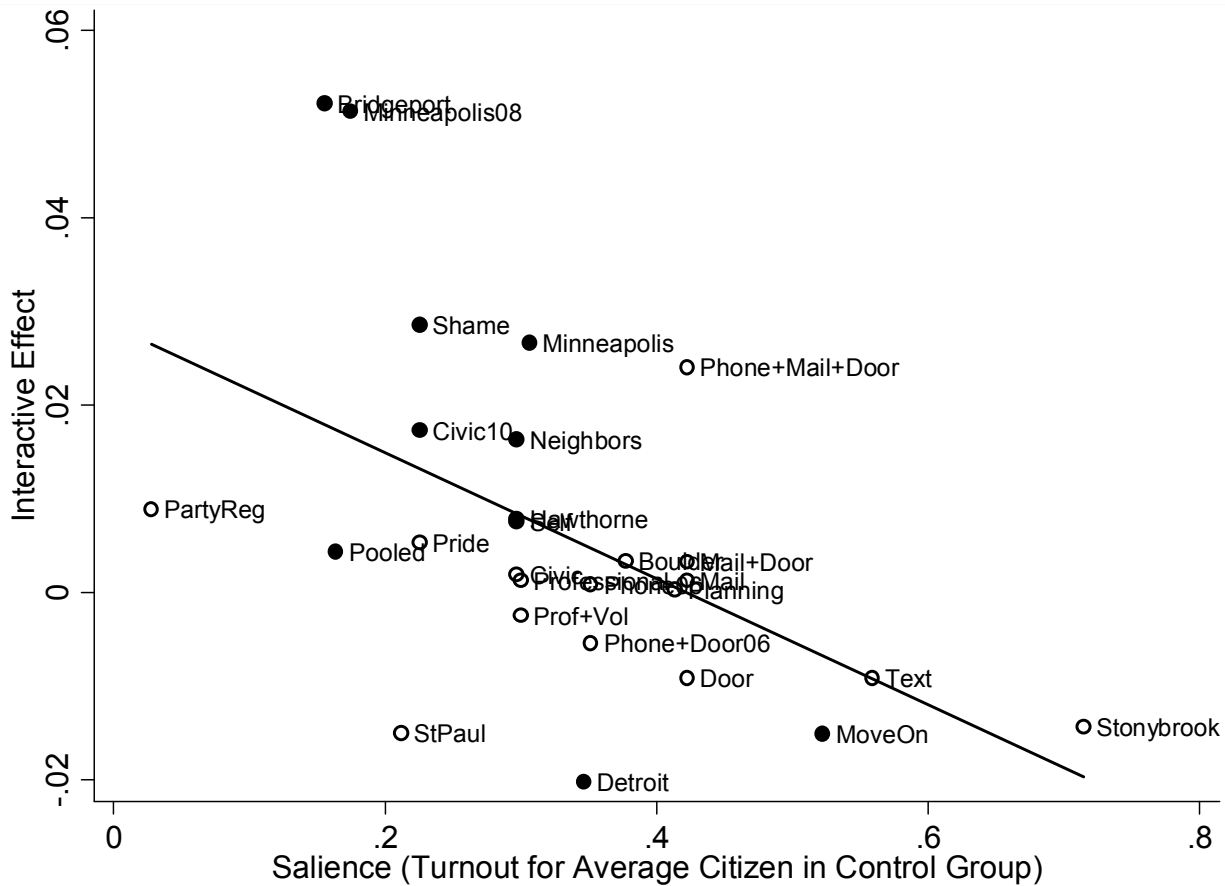
The figure assesses the salience hypothesis of Arceneaux and Nickerson (2009) that GOTV treatments will mobilize high-propensity citizens in low-salience elections and low-propensity citizens in high-salience elections. The y-axis is the multiplicative coefficient from these analyses, indicating the extent to which the treatment effect changes as propensity increases. The x-axis is the constant coefficient from our regression analyses, indicating the expected level of turnout for the average citizens in the control group. Solid circles indicate that the interactive coefficient is statistically significant (p < .05). The Arceneaux/Nickerson hypothesis is largely confirmed: as salience increases the effectiveness of GOTV treatments for low-propensity citizens increases relative to high-propensity citizens. However, significant variation remains in the data, suggesting that some treatments may be more or less effective in reducing the participation gap. Moreover, this analysis predicts that GOTV treatments will tent to exacerbate the participation gap in any setting where the average level of turnout is less than 50% -- this encompasses most elections in the U.S.

Figure A5.2. Variation across Strength of Treatment
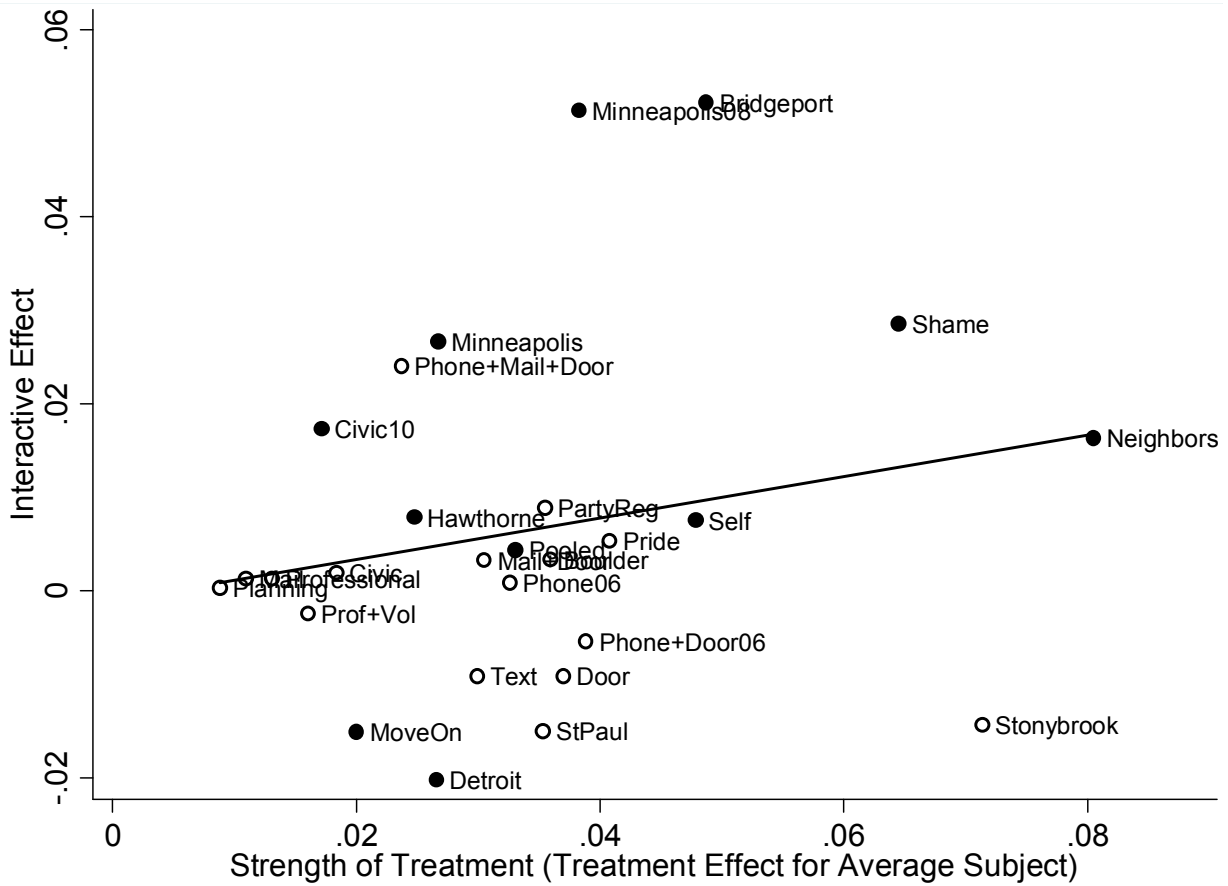
The figure presents the regression results from Table 5.3 graphically. The y-axis is the multiplicative coefficient, indicating the extent to which the treatment effect changes as the propensity variable increases. The x-axis is the additive coefficient, the treatment effect for the average citizen in the sample. Solid circles indicate that the interactive coefficient is statistically significant (p < .05). We see that the interactive effect tends to be larger for experiments with larger average effects – indicated by the upward sloping fitted line.

Bibliography

Abadie, Alberto, Alexis Diamond, and Jens Hainmueller. 2010. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association* 105(490):493-505.

Abadie, Alberto and Guido Imbens. 2008. On the Failure of the Bootstrap for Matching Estimators. *Econometrica* 76(6):1537-1557.

Addonizio, Elizabeth, Donald Green, and James Glaser. 2007. Putting the Party Back into Politics: An Experiment Testing Whether Election Day Festivals Increase Voter Turnout. *PS: Political Science and Politics* 40:721-727.

Aizer, Anna. 2003. Low Take-up in Medicaid: Does Outreach Matter and for Whom? *American Economic Review Papers and Proceedings* 3(2):238-241.

Aldrich, John. 1976. Some Problems in Testing Two Rational Models of Participation. *American Journal of Political Science 20*(4):713-733.

Aldrich, John. 1993. Rational Choice and Turnout. *American Journal of Political Science 37* (1): 246-278.

Alvarez, R. Michael, Asa Hopkins, and Betsy Sinclair. 2010. Mobilizing Pasadena Democrats: Measuring the Effects of Partisan Campaign Contacts. *Journal of Politics* 72(1):31-44.

American Political Science Association Task Force. 2004. American Democracy in an Age of Inequality. *Perspectives on Politics* 2(4):651-666.

Angrist, Joshua and Jorn-Steffen Pischke. 2009. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton: Princeton University Press.

Ansolabehere, Stephen, Eitan Hersh, Alan Gerber, and David Doherty. 2010. Voter Registration List Quality Pilot Studies. *Caltech/MIT Voting Technology Project* http://vote.caltech.edu/drupal/node/335.

Ansolabehere, Stephen and Shanto Iyengar. 1994. Of Horseshoes and Horse Races: Experimental

    Studies of the Impact of Poll Results on Electoral Behavior. *Political Communication*

    11(4):413-430.

Ansolabehere, Stephen, Jonathan Rodden, and James Snyder. 2006. Purple America. *Journal of*

    Economic Perspectives 20(2):97-118.

Ansolabehere, Stephen and James Snyder. 2000. Soft Money, Hard Money, Strong Parties.

    Columbia Law Review 100(3):598-619.

Ansolabehere, Stephen and James Snyder. 2002. The Incumbency Advantage in U.S. Elections: An

    Analysis of State and Federal Offices, 1942-2000. *Election Law Journal* 1(3):315-338.

Ansolabehere, Stephen and James Snyder. 2004. Using Term Limits to Estimate Incumbency

    Advantages when Officeholders Retire Strategically. *Legislative Studies Quarterly* 4:487-515.

Anzia, Sarah. 2011. Election Timing and the Electoral Influence of Interest Groups. *Journal of*

    Politics 73(2):412-427.

Anzia, Sarah. 2012. The Election Timing Effect: Evidence from a Policy Intervention in Texas.

    Quarterly Journal of Political Science 7(3):209-248.

Arceneaux, Kevin and David Nickerson. 2009. Who is Mobilized to Vote? A Re-Analysis of 11

    Field Experiments. *American Journal of Political Science* 53(1):1-16.

Avery, Michael. 1989. *The Demobilization of American Voters.* Greenwood Press.

Baldwin, Peter. 1990. *The Politics of Social Solidarity: Class Bases of the European Welfare State, 1875-1975*.

    Cambridge University Press.

Bartels, Larry. 2006. What's the Matter with What's the Matter with Kansas? *Quarterly Journal of*

    Political Science 1:201-226.

Bartels, Larry. 2008. *Unequal Democracy: the Political Economy of the New Gilded Age:* Princeton

    University Press.

Bartels, Larry. 2009. Economic Inequality and Political Representation. *The Unsustainable American State*, eds. Lawrence Jacobs and Desmond King. Oxford University Press.

Barzel, Yoram and Eugene Silberberg. 1973. Is the Act of Voting Rational? *Public Choice* 16(1):51-58.

Becker, Gary. 1965. A Theory of the Allocation of Time. *The Economic Journal* 75:493-517.

Bennion, Elizabeth. 2005. Caught in the Ground Wars: Mobilizing Voters During a Competitive Congressional Campaign. *Annals of the American Academy of Political and Social Research* 601:123-141.

Berch, Neil. 1993. Another Look at Closeness and Turnout: The Case of the 1979 and 1980 Canadian National Elections. *Political Research Quarterly* 46(2):421-432.

Berinsky, Adam. 2005. The Perverse Consequences of Electoral Reform in the United States. *American Politics Research* 33(4):471-491.

Bernhagen, Patrick and Michael Marsh. 2007. The Partisan Effects of Low Turnout: Analyzing Vote Absention as a Missing Data Problem. *Electoral Studies* 26(3):548-560.

Berry, Christopher and Jacob Gersen. 2011. Election Timing and Public Policy. *Quarterly Journal of Political Science* 6(2):103-135.

Blais, André. 2000. *To Vote Or Not To Vote? The Merits and Limits of Rational Choice*, Pittsburgh, University of Pittsburgh Press.

Blais, Andre. 2006. "What affects voter turnout?". *Annual Review of Political Science* 9: 111-25.

Blais, Andre, Robert Young, and Miriam Lapp. 2000. The Calculus of Voting: An Empirical Test. *European Journal of Political Research* 37(2):181-201.

Blais, André, Jean-Benoit Pilet, Karine van der Straeten, Jean-François Laslier, École Polytechnique, and Maxime Héroux-Legault. 2011. "To Vote or To Abstain? An Experimental Study of First Past the Post and PR Elections". Working paper.

Blundell, Richard, Vanessa Fry, and Ian Walker. 1988. Modelling the Take-up of Means-tested Benefits: The Case of Housing Benefits in the United Kingdom. *The Economic Journal* 98:58-74.

Booth, John and Patricia Bayer Richard. 1998. Civil Society, Political Capital, and Democratization in Central America. *Journal of Politics* 60:780-800.

Bowler, Shaun and Todd Donovan. Electoral Competition and the Voter. *Public Opinion Quarterly* 75(1):151-164.

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan. 2004. How Much Should We Trust Differences-In-Differences Estimates? *Quarterly Journal of Economics* 119(1):249-275.

Bruce, Stanley. 1925. *Prime Minister's Fighting Speech.* Published by Winn & Co.

Bullock, John, Donald Green, and Shang Ha. 2010. Yes, But What's the Mechanism? (Don't Expect an Easy Answer). *Journal of Personality and Social Psychology* 98(4):550-558.

Burnham, Walter Dean. 1965. The chaining shape of the American political universe. *American Political Science Review* 24(1):54-61.

Citrin, Jack, Eric Schickler, and John Sides. 2003. What if Everyone Voted? Simulating the Impact of Increased Turnout in Senate Elections. *American Journal of Political Science* 47(1):75-90.

Clarke, Peter and Susan Evans. 1983. *Covering Campaigns: Journalism in Congressional Elections.* Stanford University Press.

Condon, Meghan. 2009. The Effects of Social Capital on Voter Turnout: A Field Experiment in Two Southwestern Cities. Working Paper.

Cox, Gary and Michael Munger. 1989. Closeness, Expenditures, and Turnout in the 1982 U.S. House Elections. *American Political Science Review* 83(1):217-231.

Dahl, Robert. 1961. *Who Governs?* Yale University Press.

Dahl, Robert. 1970. *After the Revolution: Authority in a Good Society.* Yale University Press.

Dahl, Robert. 2006. *On Political Equality*. New Haven: Yale University Press.

Dale, Allison and Aaron Strauss. 2009. Don't Forget to Vote: Text Message Reminders as a Mobilization Tool. *American Journal of Political Science* 53(4):787-804.

Davenport, Tiffany, Alan Gerber, Donald Green, Christopher Larimer, Christopher Mann, and Costas Panagopoulos. 2010. The Enduring Effects of Social Pressure: Tracking Campaign Experiments over a Series of Elections. *Political Behavior* 3:423-430.

DeNardo, James. 1980. Turnout and the Vote: The Joke's on the Democrats. *American Political Science Review* 74(2):406-420.

Denny, Kevin and Orla Doyle. 2008. Political Interest, Cognitive Ability and Personality: Determinants of Voter Turnout in Britain. *British Journal of Political Science* 38:291-310.

De Tocqueville, Alexis. 1840. *Democracy in America*.

Downs, Anthony. 1957. *An Economic Theory of Democracy*. Harper and Row.

Duch, Raymond and Harvey Palmer. 2004. It's Not Whether You Win or Lose, but How You Play the Game: Self-Interest, Social Justice, and Mass Attitudes toward Market Transition. *American Political Science Review* 98:437-452.

Duffy, John and Margit Tavits. 2008. Beliefs and Voting Decisions: A Test of the Pivotal Voter Model. *American Journal of Political Science* 52(3):603-618.

Edlin, Aaron, Andrew Gelman, and Noah Kaplan. 2007. Voting as Rational Choice: Why and How People Vote to Improve the Well-Being of Others. *Rationality and Society* 19:293-314.

Enos, Ryan D., Anthony Fowler, and Lynn Vavreck. 2012. Increasing Inequality: The Effect of GOTV Mobilization on the Composition of the Electorate. MPSA Annual Meeting Paper.

Erikson, Robert. 1995. State Turnout and Presidential Voting: A Closer Look. *American Politics Quarterly* 23:387-396.

Erikson, Robert and Thomas Palfrey. 2000. Equilibria in Campaign Spending Games: Theory and Data. *American Political Science Review* 94(3):595-609.

Evans, Tim. 2006. Compulsory Voting in Australia. Australian Electoral Commission Report.

Ferejohn, John and Morris Fiorina. 1974. The Paradox of Not Voting: A Decision Theoretic Analysis. *American Political Science Review* 68(2):525-536.

Ferejohn, John and Morris Fiorina. 1975. Closeness Counts Only in Horseshoes and Dancing. *American Political Science Review* 69(3):920-925.

Festinger, Leon. 1957. *A Theory of Cognitive Dissonance.* Stanford University Press.

Fiorina, Morris. 1990. Information and Rationality in Elections. *Information and Democratic Processes.* University of Illinois Press.

Fisher, Stephen. 2007. (Change in) Turnout and (Change in) the Left Share of the Vote. *Electoral Studies* 26(3):598-611.

Fowler, Anthony. 2012. Regular Voters, Marginal Voters, and the Electoral Effects of Higher Turnout. APSA Annual Meeting Paper.

Fowler, Anthony. 2013. Electoral and Policy Consequences of Voter Turnout: Evidence from Compulsory Voting in Australia. *Quarterly Journal of Political Science*, forthcoming.

Fowler, James. 2005. Turnout in a Small World. *Social Logic of Politics.* Temple University Press.

Fowler, James. 2006. Altruism and Turnout. *Journal of Politics* 68:674-683.

Fraga, Bernard and Eitan Hersh. 2011. Voting Costs and Voter Turnout in Competitive Elections. *Quarterly Journal of Political Science* 5(4):339-356.

Franklin, Mark. 2004. *Voter Turnout and the Dynamics of Electoral Competition in Established Democracies since 1945.* Cambridge University Press.

Franklin, Mark 2008. Voter participation in the United States versus other democracies. *Voting in America.* Morgan E. Flechner, ed. Westport, CT: Praeger.

Gabler, Nicole, Naihua Duan, Diana Liao, Joann Elmore, Theodore Ganiats, and Richard Kravits. 2009. Dealing with Heterogeneity of Treatment Effects: Is the Literature Up to the Challenge? *Trials* 10:43.

Garcia Bedolla, Lisa and Melissa R. Michelson. 2012. *Mobilizing Inclusion: Transforming the Electorate through Get-out-the-vote Campaigns.* New Haven: Yale University Press.

Gelman, Andrew and Gary King. 1990. Estimating Incumbency Advantage without Bias. *American Journal of Political Science* 34(4):1142-64.

Gelman, Andrew, Gary King, and John Boscardin. 1998. Estimating the Probability of Events That Have Never Occurred: When Is Your Vote Decisive*? Journal of the American Statistical Association* 93(441):1-9.

Gelman, Andrew, David Park, Boris Shor, Joseph Bafumi, and Jeronimo Cortina. 2008. *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do.* Princeton University Press.

Gelman, Andrew, Boris Shor, Joseph Bafumi, and David Park. 2007. Rich State, Poor State, Red State, Blue State: What's the Matter with Connecticut? *Quarterly Journal of Political Science* 2(4):345-367.

Gerber, Alan S., James G. Gimpel, Donald P. Green, and Daron R. Shaw. 2011. "How Large and Long-lasting Are the Persuasive Effects of Televised Campaign Ads? Results from a Randomized Field Experiment". *American Political Science Review105* (1): pp. 135-150.

Gerber, Alan and Donald Green. 2000. The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment. *American Political Science Review* 94(3):653-663.

Gerber, Alan and Donald Green. 2000b. The Effect of a Nonpartisan Get-Out-The-Vote Drive: An Experimental Study of Leafletting. *Journal of Politics* 62(3):846-57.

Gerber, Alan, Donald Green, and Christopher Larimer. 2008. Social Pressure and Voter Turnout:

> Evidence from a Large-Scale Field Experiment. *American Political Science Review* 102(1):33-48.

Gerber, Alan, Donald Green, and Christopher Larimer. 2010. An Experiment Testing the Relative

> Effectiveness of Encouraging Voter Participation by Inducing Feelings of Pride or Shame.
>
> *Political Behavior* 32(3):409-422.

Gerber, Alan, Donald Green, 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.

Gerber, Alan, Donald Green, and Ron Shachar. 2003. Voting May Be Habit-Forming: Evidence

> from a Randomized Field Experiment. *American Journal of Political Science* 47:540-550.

Gerber, Alan, Jonathan Gruber, and Daniel Hungerman. 2008. Does Church Attendane Cause

> People to Vote? Using Blue Laws' Repeal to Estimate the Effect of Religiosity on Voter
>
> Turnout. NBER Working Paper No. 14303.

Gerber, Alan, Gregory Huber, David Doherty, Conor Dowling, Connor Raso, and Shang Ha. 2011.

> Personality Traits and Participation in the Political Process. *Journal of Politics* 73:692-706.

Gerber, Alan, Gregory Huber, Conor Dowling, David Doherty, and Nicole Schwartzberg. 2009.

> Using Battleground States as a Natural Experiment to Test Theories of Voting. APSA 2009
>
> Toronto Meeting Paper.

Gerber, Alan, Gregory Huber, and Ebonya Washington. 2010. Party Affiliation, Partisanship, and

> Political Beliefs: A Field Experiment. *American Political Science Review* 104(4):720-744.

Gerber, Alan S., Dean Karlan, and Daniel Bergan. 2009. "Does the Media Matter? A Field

> Experiment Measuring the Effect of Newspapers on Voting Behavior and Political
>
> Opinions". *American Economic Journal: Applied Economics* 1(2): pp. 35-52.

Gilens, Martin. 2005. Inequality and Democratic Responsiveness. *Public Opinion Quarterly* 69(5)778-796.

Gilens, Martin. 2012. *Affluence and Influence: Economic Inequality and Political Power in America.* Princeton University Press.

Gimpel , James G., Karen M. Kaufmann, and Shanna Pearson-Merkowitz. Battleground States versus Blackout States: The Behavioral Implications of Modern Presidential Campaigns. *Journal of Politics* 69(3): 786-797.

Gomez, Brad, Thomas Hansford, and George Krause. 2007. The Republicans Should Pray for Rain: Weather, Turnout, and Voting in U.S. Presidential Elections. *Journal of Politics* 69:649-663.

Gosnell, Harold. 1927. *Getting Out the Vote: An Experiment in the Stimulation of Voting.* Chicago: University of Chicago Press.

Green, Donald P. and Alan S. Gerber. 2008. *Get out the Vote: How to increase voter turnout.* Brookings Press.

Green, Donald and Holger Kern. 2012. Modeling Heterogeneous Treatment Effects in Survey Experiments with Bayesian Additive Regression Trees. *Public Opinion Quarterly*, forthcoming.

Green, Donald and Melissa R. Michelson. 2009. "ACORN experiments in minority voter mobilization" in *The People Shall Rule*, Robert Fisher (ed). Nashville: Vanderbilt University Press.

Green, Donald P. and Lynn Vavreck. 2008. Analysis of Cluster-Randomized Field Experiments: A Comparison of Alternative Estimation Approaches. *Political Analysis 16*: (2), pp. 138-152.

Griffin, Brian and Brian Newman. Are voters better represented. *The Journal of Politics* 67(4), pp. 1206-1227.

Grofman, Bernard, Guillermo Own, and Christian Collet. 1999. Rethinking the Partisan Effects of Higher Turnout: So What's the Question? *Public Choice* 99:357-376.

Hajnal, Zoltan and Jessica Trounstine. 2005. "Where turnout matters: the consequences of uneven turnout in city politics". *Journal of Politics* 67(2):515-536.

Hansford, Thomas and Brad Gomez. 2010. Estimating the Effects of Voter Turnout. *American Political Science Review* 104(2):268-288.

Hausman, Jerry. 1978. Specification Tests in Econometrics. *Econometrica* 46:1251-1271.

Herron, Michael. 1998. The Presidential Election of 1988: Low Voter Turnout and the Defeat of Michael Dukakis. Working Paper.

Highton, Benjamin and Raymond Wolfinger. 2001. The Political Implications of Higher Turnout. *British Journal of Political Science* 31:179-223.

Hill, David and Seth McKee. 2005. The Electoral College, Mobilization, and Turnout in the 2000 Presidential Election. *American Politics Research* 33(5):700-725.

Hill, Seth. 2010. *Changing Composition and Changing Allegiance in American Politics.* Doctoral Dissertation, University of California, Los Angeles.

Hirst, John. 2002. *Australia's Democracy: A Short History.* Allen and Unwin.

Hughes, William. 1922. Prime Minister's Policy Speech. *National Policy and Record.* Chatswood, North Sydney.

Issenberg, Sasha. 2010. "Nudge the Vote." *The New York Times Magazine*, October 29, 2010.

Issenberg, Sasha. 2012. *The Victory Lab: The Secret of Winning Campaigns.* New York: Crown.

Jackman, Simon. 1999. Non-compulsory Voting in Australia?: What Surveys Can (and Can't) Tell Us. *Electoral Studies* 18(1):29-48.

Jackson, Robert. 1996. The Mobilization of Congressional Electorates. *Legislative Studies Quarterly* 21(3):425-445.

Jacobson, Gary. 1997. *The Politics of Congressional Elections*. Longman Press.

Kam, Cindy and Stephen Utych. 2011. Close Elections and Cognitive Engagement. *Journal of Politics* 73(4):1251-1266.

Kelley, Stanley, Richard Ayres, and William Bowen. 1967. Registration and Voting: Putting First Things First. *American Political Science Review* 61(2):359-377.

Key, V. O. 1949. *Southern Politics in State and Nation*. Knopf, Inc.

Kim, Jae-On, John Petrocik, and Stephen Enokson. 1975. Voter Turnout among the American States: Systematic and Individual Components. *American Political Science Review* 69(1):107-123.

King, James D. 2000. Changes in Professionalism in U.S. State Legislators. *Legislative Studies Quarterly* 25(2):327-343.

Knack, Stephen. 1992. Civil Norms, Social Sanctions, and Voter Turnout. *Rationality and Society* 4:133-156.

Knack, Stephen and Philip Keefer. 1997. Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. *Quarterly Journal of Economics* 112:1251-1288.

Knack, Stephen and James White. 1998. Did States' Motor Voter Programs Help the Democrats? *American Politics Research* 26(3):344-365.

Krishna, Anirudh. 2007. How Does Social Capital Grow? A Seven-Year Study of Villages in India. *Journal of Politics* 69:941-956.

Lake, Ronald La Due and Robert Huckfeldt. 1998. Social Capital, Social Networks, and Political Participation. *Political Psychology* 19:567-584.

Lastra, Yolando, Dina Sherzer, and Joel Sherzer. 2009. *Adoring the Saints: Fiestas in Central Mexico*. University of Texas Press.

Leighley, Jan and Jonathan Nagler. 1992. Socioeconomic Class Bias in Turnout, 1964-1988: The Voters Remain the Same. *American Political Science Review* 86(3):725-736.

Letki, Natalia and Geoffrey Evans. 2005. Endogenizing Social Trust: Democratization in East-Central Europe. *British Journal of Political Science* 35:515-529.

Levine, David and Thomas Palfrey. 2007. The Paradox of Voter Participation? A Laboratory Study. *American Political Science Review* 101(1):143-158.

Lijphart, Arend. 1997. Unequal Participation: Democracy's Unresolved Dilemma. *American Political Science Review* 91(1):1-14.

Lipset, Seymour Martin. 1959. Some Social Requisites of Democracy: Economic Development and Political Legitimacy. *American Political Science Review* 53(1):69-105.

Lipsitz, Keena. 2008. The Consequences of Battleground and "Spectator" State Residency for Political Participation. *Political Behavior* 31(2):1-23.

Linder, Wolf. 1994. *Swiss Democracy: Possible Solutions to Conflict in Multicultural.* St. Martin's Press.

Lindert, Peter. 1994. The Rise of Social Spending, 1880-1930. *Explorations in Economic History* 31:1-37.

Mackerras, Malcolm and Ian McAllister. 1999. Compulsory Voting, Party Stability, and Electoral Advantage in Australia. *Electoral Studies* 18(2):217-233.

Martinez, Michael and Jeff Gill. 2005. The Effects of Turnout on Partisan Outcomes in U.S. Presidential Elections 1960-2000. *Journal of Politics* 67(4):1248-1274.

Martinez, Michael and David Hill. 2007. Was the Joke on the Democrats Again? Turnout and Partisan Choice in the 2004 U.S. Election. *American Review of Politics* 28:81-95.

Maslunka, Michele. 1983. *Bendigo: One Community's Response to Federation.* Honours Thesis, University of Melbourne.

Matsusaka, John. 1993. Election Closeness and Voter Turnout: Evidence from California Ballot Propositions. *Public Choice* 76(4):313-334.

Matsusaka, John and Filip Palda. 1993. The Downsian Voter Meets the Ecological Fallacy. *Public Choice* 77(4):855-878.

McAllister, Ian. 1986. Compulsory Voting, Turnout, and Party Advantage in Australia. *Australian Journal of Political Science* 21(1):89-93.

McAllister, Ian and Anthony Mughan. 1986. Differential Turnout and Party Advantage in British General Elections, 1964-1983. *Electoral Studies* 5:143-152.

McCarty, Nolan, Keith Poole, and Howard Rosenthal. 2006. *Polarized America. The Dance of Ideology and Unequal Riches.* Cambridge: MIT Press.

McDonald, Michael and Samuel Popkin. 2001. The Myth of the Vanishing Voter. *American Political Science Review* 95(4):963.974.

McFadden, Daniel. 1974. Conditional Logit Analysis of Qualitative Choice Behavior. *Frontiers in Econometrics*, ed. P. Zarembka. Academic Press.

Meredith, Marc. 2009. Persistence in Political Participation. *Quarterly Journal of Political Science* 4(3):186-208.

Michelson, Melissa R., Lisa García Bedolla and Donald P. Green. 2008. "Mobilizing the Registered: Tactics for Improving Voter Turnout." In Morgan E. Felchner, ed., *Voting in America*, (Praeger), pp. 198-211.

Middleton, Joel and Donald Green. 2008. Do Community-Based Mobilization Campaigns Work Even in Battleground States? Evaluating the Effectiveness of MoveOn's 2004 Outreach Campaign. *Quarterly Journal of Political Science* 3(1):63-82.

Muligan, Casey B. and Charles G. Hunter. 2003. "The Empirical Frequency of a Pivotal Vote". *Public Choice* 116(1/2): 31-54.

Mutz, Diana. 2002. The Consequences of Cross-Cutting Networks for Political Participation. *American Journal of Political Science* 46:838:855.

Nagel, Jack H. 1988. Voter Turnout in New Zealand General Elections, 1928-1988. *Political Science* 40:16-38.

Nagel, Jack and John McNulty. 1996. Partisan Effects of Voter Turnout in Senatorial and Gubernatorial Elections. *American Political Science Review* 90(4):780-793.

Nelson, Paul. 1995. *World Bank and Nongovernmental Organizations: The Limits of Apolitical Development.* St. Martin's Press.

Nickerson, David. 2006. Volunteer Phone Calls Can Increase Turnout. *American Politics Research* 34(3):271-292.

Nickerson, David. 2007a. Does Email Boost Turnout? *Quarterly Journal of Political Science* 2: 369-379.

Nickerson, David. 2007b. Quality Is Job One: Professional and Volunteer Voter Mobilization Calls. *American Journal of Political Science* 51(2):269-282.

Nickerson, David. 2008. Is Voting Contagious? Evidence form Two Field Experiments. *American Political Science Review* 102(1):49-57.

Nickerson, David, Ryan Friedrichs, and David King. 2006. Partisan Mobilization Campaigns in the Field: Results from a Statewide Turnout Experiment in Michigan. *Political Research Quarterly* 59(1):85-97.

Nickerson, David and Todd Rogers. 2010. Do You Have a Voting Plan? Impelmentation Intentions, Voter Turnout, and Organic Plan Making. *Psychological Science* 21(2):194-199.

Olson, Mancur. 1965. *The Logic of Collective Action.* Harvard University Press.

Panagopoulos, Costas and Donald Green. 2008. Field Experiments Testing the Impact of Radio Advertisements on Electoral Competition. *American Journal of Political Science* 52(1):156-168.

Patterson, Samuel and Gregory Caldiera. 1983. Getting Out the Vote: Participation in Gubernatorial Elections. *American Political Science Review* 77(3):675-689.

Paxton, Pamela. 2002. Social Capital and Democracy: An Interdependent Relationship. *American Sociological Review* 67:254-277.

Pettersen, Per Arnt and Larence Rose. 2007. The Dog that Didn't Bark: Would Increased Electoral Turnout Make a Difference? *Electoral Studies* 26(3):574-588.

Pickett, Kate, Ye Luo, and Diane Lauderdale. 2005. Widening Social Inequalities in Risk for Sudden Infant Death Syndrome. *American Journal of Public Health* 95(11):1976-1981.

Piven, Frances Fox and Richard Cloward. 1988. *Why Americans Don't Vote.* Pantheon Books.

Piven, Frances Fox and Richard A. Cloward. 2000. *Why Americans Still Don't Vote: and Why Politicians Want It That Way.* Boston: Beacon Press.

Plutzer, Eric. 2002. Becoming a Habitual Voter: Inertia, Resources, and Growth in Young Adulthood. *American Political Science Review* 96:41-56.

Pocock, Stuart, Susan Assmann, Laura Enos, and Linda Kasten. 2002. Subgroup Analysis, Covariate Adjustment, and Baseline Comparisons in Clinical Trial Reporting: Current Practice and Problems. *Statistics in Medicine* 23:2509-2525.

Powell, G. Bingham. 1980. Voting Turnout in Thirty Democracies: Partisan, Legal, and Socio-Economic Influences. *Electoral Participation: A Comparative Analysis*, ed. Richard Rose. Sage Press.

Putnam, Robert D. 1993. *Making Democracy Work: Civic Traditions in Modern Italy.* Princeton University Press.

Putnam, Robert D. 1995. Tuning In, Tuning Out; The Strange Disappearance of Social Capital in America. *PS: Political Science and Politics* 28:664-683.

Putnam, Robert D. 2000. *Bowling Alone.* Simon and Schuster.

Putnam, Robert D. 2002. Foreward. *The Role of Social Capital in Development: An Empirical Assessment.* Eds. Christian Grootaert and Thierry van Bastelaer. Cambridge University Press.

Putnam, Robert D. and David Campbell. 2010. *American Grace: How Religion Divides and Unites Us*. Simon & Schuster.

Putnam, Robert D and Lewis Feldstein. 2003. *Better Together: Restoring the American Community*. Simon & Schuster.

Radcliff, Benjamin. 1994. Turnout and the Democratic Vote. *American Politics Quarterly* 9:259-276.

Riker, William and Peter Ordeshook. 1968. A Theory of the Calculus of Voting. *American Political Science Review* 62(1):25-42.

Rotemberg, Julio. 2009. Attitude-Dependent Altruism, Turnout, and Voting. *Public Choice* 140:223-244.

Rubenson, Daniel, Andre Blais, Patrick Fournier, Elisabeth Digengil, and Neil Nevitte. 2007. Does Low Turnout Matter: Evidence from the 2000 Canadian Election. *Electoral Studies* 26(3):589-597.

Rupasingha, Anil, Stephan Goetz, and David Freshwater. 2006. The Production of Social Capital in US Counties. *Journal of Socio-Economics*. 46:83-101.

Schattschneider, E. E. 1960. *The Semisovereign People: A Realist's View of Democracy in America*. Holt, Rinehart, and Winston.

Scheufele, Dietram and Shavan Shah. 2000. Personality Strength and Social Capital: The Role of Dispositional and Informational Variables in the Production of Civic Participation. *Communication Research* 27:107-131.

Schlozman, Kay Lehman, Sidney Verba, and Henry E. Brady. 2012. *The Unheavenly Chorus: Unequal Political Voice and the Broken Promise of American Democracy*. Princeton University Press.

Schwartz, Thomas. 1987. Your Vote Counts on Account of the Way it is Counted: An Institutional Solution to the Paradox of Voting. *Public Choice* 54(2):101-121.

Seidle, Leslie and David Miller. 1976. Turnout, Rational Abstention, and Campaign Effort. *Public Choice 27*(1):121-126.

Shachar, Ron and Barry Nalebuff. 1999. Follow the Leader: Theory and Evidence on Political Participation. *American Economic Review* 89(3):525-547.

Shaw, Daron. 1999. The Effect of TV Ads and Candidate Appearances on Statewide Presidential Votes, 1988-96. *American Political Science Review* 93(2): 345-361.

Shaw, Daron. 2006. *The Race to 270*. Chicago: University of Chicago Press.

Skocpol, Theda. 1997. The Tocqueville Problem: Civic Engagement and American Democracy. *Social Science History* 21:455-479.

Sled, Sarah. 2008. *It's In the Mail: The Effect of Vote By Mail Balloting on Voter Turnout and Policy Outcomes in U.S. Elections.* Doctoral Dissertation, Massachusetts Institute of Technology.

Stein, Robert. 1998. Early Voting. *Public Opinion Quarterly* 62(1):57-69.

Stoker, Laura and M. Kent Jennings. 1995. Life-Cycle Transitions and Political Participation: The Case of Marriage. *American Political Science Review* 89:421-433.

Stolle, Dietlind. 1998. Bowling Together, Bowling Alone: The Development of Generalized Trust in Voluntary Associations. *Political Psychology* 19:497-525.

Squire, Peverill. 2007. Measuring State Legislative Professionalism: The Squire Index Revisited. *State Politics and Policy Quarterly* 7(2):211-227.

Tavits, Margit. 2006. Making Democracy Work More? Exploring the Linkage between Social Capital and Government Performance. *Political Research Quarterly* 59:211-225.

Teixeira, Ruy A.1992. *The Disappearing American Voter.* Washington, D.C.: Brookings.

Tingsten, Herbert. 1937. *Political Behavior: Studies in Election Statistics* P.S. King and Son.

Topf, Richard. 1995. Electoral Participation. *Citizens and the State*, ed. Hans-Dieter Klingemann and Dieter Fuchs. Oxford University Press.

Tucker, Harvey. 1986. Contextual Models of Participation in U.S. State Legislative Elections. *The Western Political Quarterly* 39(1):67-78.

van der Eijk and Marcel van Egmond. 2007. Political Effects of Low Turnout in National and European Elections. *Electoral Studies* 26(3):561-573.

Vavreck, Lynn. 2007. "The Dangers of Self-Reports of Political Behavior", Quarterly Journal of Political Science 2:287.

Verba, Sidney and Norman Nie. 1972. *Participation in America: Political Democracy and Social Inequality.* Harper and Row.

Verba, Sidney, Kay Schlozman, and Henry Brady. 1995. *Voice and Equality: Civic Voluntarism in American Politics.* Harvard University Press.

Wattenberg, Martin P. 2002. *Where Have All the Voters Gone?* Cambridge: Harvard University Press.

Wee, Liang En, Gerald Choon Huat Koh, Wei Xin Yeo, Run Ting Chin, Jolene Wong, and Braden Seow. 2012. Screening for Cardiovascular Disease Risk Factors in an Urban Low-income Setting at Baseline and Post Intervention: A Prospective Intervention Study. *European Journal of Preventative Cardiology*, forthcoming.

Wolak, Jennifer. 2006. The Consequences of Presidential Battleground Strategies for Citizen Engagement. *Political Research Quarterly* 59(3):353-361.

Wolfinger, Raymond and Steven Rosenstone. 1980. *Who Votes?* Yale University Press.

Woods, Neal D. and Michael Baranowski. 2006. Legislative Professionalism and Influence on State Agencies: The Effects of Resources and Careerism. *Legislative Studies Quarterly* 31(4): 585-609.

Wuffle, A and Christian Collet. 1997. Why Democrats Shouldn't Vote (with acknowledgements to R. Erikson). *Journal of Theoretical Politics* 9(1):137-140.