Essays on Legislative Behavior: Evidence from State Legislatures

by

Andrew Feher

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Political Science)
in The University of Michigan
2015

Doctoral Committee:

Assistant Professor Rocío Titiunik, Chair
Professor Elisabeth Gerber
Professor Kenneth W. Kollman
Professor Charles R. Shipan

For my mother

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support of a number of people. My dissertation committee—Liz Gerber, Ken Kollman, Chuck Shipan and Rocío Titiunik—collectively provided incalculable amounts of their time, and it was an absolute luxury to learn from each of these professors. Let me begin with my advisor, Rocío. I first met her as an overwhelmed first-year graduate student who couldn't distinguish between a casual and a causal inference. It wouldn't be an exaggeration to say she has permanently overturned the furniture in my mind. When my confidence or enthusiasm flagged, which it invariably did, her indefatigable energy gave me the jolt I needed to continue. One of the few windfalls I experienced in graduate school was having an office two doors away from hers. Some of my fondest memories in graduate school center on conversations during office hours where I simply tried to keep up as her speed boat of a brain began accelerating. Over the past four years, she has patiently mentored me, profoundly influencing how I think about and evaluate political science research. She invested more time and effort into not just this dissertation, but into my development as a researcher, than any adviser should.

I also benefited greatly from the other members of my committee. Ken, more than anyone, shaped my understanding of American Politics. I had the good fortune of teaching for him twice, and along the way acquired a mountain of insight into successful pedagogy. He also organized a weekly dissertation group that served as a much-needed commitment device during my final two years in graduate school. He always managed to find a way to inject levity into the lengthy process of completing this dissertation by regularly engaging me in conversations about restaurants and home cooking.

In my first semester as a GSI, I was assigned to Chuck's Introduction to American Politics course. He remains a paragon of professionalism, and I can only aspire to the lofty standard he set. Meetings with him at the dissertation stage were always equal parts daunting and productive: daunting because of his uncanny ability to identify the shortcomings in my work and productive because he would always offer suggestions for improving it.

Lastly, having read many of Liz's books and articles, I never thought I would have had a chance to work in any meaningful way with her. And yet, notwithstanding her busy schedule, Liz made time to meet at the early stages of the dissertation to discuss my half-baked ideas that thankfully were left at the cutting room floor.

In addition to my dissertation committee, Richard Anderson, David Cottrell, Jeremy Gelman, John Jackson, Trevor Johnston, Rachel Potter, Molly Reynolds, Chris Skovron and participants at the Interdisciplinary Workshop in American Politics (IWAP) read drafts, offered feedback and provided a healthy dose of laughs along the way.

Finally, I must thank my family for their support. My brother flew to Chicago for my first talk at a conference and—no matter how swamped he was with work—went to great lengths to make my bi-annual visits to Los Angeles memorable with sushi, tsukemen or hedonistic restaurant crawls. My mother, with whom I spoke almost daily throughout graduate school, has been subjected to thousands of hours of me complaining. No matter how deep a morass I was in, she found a way to allay my angst and encouraged me to embrace the comfort of simple pleasures, whether that was a long bike ride or an order of fried chicken at Zingerman's Roadhouse. She remains a steadfast pillar of support, for which I am endlessly grateful.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

ABSTRACT

Essays on Legislative Behavior: Evidence from State Legislatures

by

Andrew Feher

This dissertation focuses on the relationship between legislative institutions and legislative behavior, answering the following questions: (1) what impact do term limits policies have on the legislative output of state legislators? (2) how do term limits affect the participatory decisions of state legislators? and (3) what is the effect of ideology on a legislator's standing in the chamber?

CHAPTER I

Introduction

In recent years, scholars, pundits and political commentators have written extensively about the deficiencies in American legislatures, ranging from high reelection rates to ostensibly invulnerable incumbents and legislative intransigence (Mann and Ornstein, 2006, 2013; Samples and McDonald, 2006; Lessig, 2012; Will, 1992). Amid surveys suggesting voters want effective governance above all else, there has been a rise in the use of dilatory parliamentary procedures that keep agencies understaffed and in partisan conflict that has resulted in credit downgrades and even a government shutdown (Fiorina, Abrams and Pope, 2005).

Frustrated by these and related developments, reform-oriented groups and good government advocates have put forth proposals in an effort to remedy the aforementioned deficiencies. Given that those in office tend to benefit from the current state of affairs, reformers have bypassed governing institutions altogether, instead turning to external alternatives such as the initiative process and in some cases their efforts have paid off. Two of the more prominent institutional reforms imposed via the initiative have been the passage of term limits and primary reform. These measures endeavored to improve representation by removing "entrenched" incumbents from office and limiting the influence of ideologically extreme candidates through open primaries, respectively.

1.1 The Underpants Gnome Theory of Institutional Reform

The case for these reforms appears strong in the abstract, and in the eyes of proponents, they constitute sure-fire panaceas. What often gets glossed over, however, is how the proposed reforms will bring about their intended effects, evoking the disjointed logic of the underwear-stealing, profit-seeking South Park gnomes:[1]

1. Term Limits or Primary Reform
2. ???
3. Improved Representation!

Tacitly embedded in these initiatives is a set of assumptions about voters, legislative behavior and the role of political parties. In the case of primary reform, for example, rather than relying on party identification or newspaper endorsements as a heuristic to guide their candidate choice, voters must be able to successfully identify moderate candidates—in low information electoral settings no less—a challenge that has proven difficult in practice (Ahler, Citrin and Lenz, 2014). As for term limits, reformers expected that once long-serving incumbents were termed-out, a different type of officer-seeker would be incented to serve, one more attuned to voters and less likely to develop an inside-the-capitol perspective. Again, however, it seems this expectation failed to materialize as the evidence to date suggests that legislators in term-limited states are just as likely to see themselves as career politicians as those in non-term-limited states (Carey, Niemi and Powell, 2000).

One of the goals of this dissertation is to examine the relationship between legislative institutions and legislative behavior, and in so doing, to challenge the belief that layering new institutions atop existing ones represents a simple fix. Before simply accepting reformers' claims, it is important to question the need for sweeping changes when existing institutions—flawed though they may be—not only exist to combat the aforementioned maladies, but have demonstrated the capacity to do so.

[1] The gnomes' business plan was as follows: Collect Underpants $\Rightarrow$??? $\Rightarrow$ $Profit$!

Using the running examples of term limits and primary reform, in lieu of setting limits on the number of terms a legislator can serve, voters can use elections to either sanction unscrupulous legislators when they act contrary to their interests or to select high-quality representatives who are principled and share the constituency's preferences.[2] The empirical evidence from both state and federal elections suggests voters do indeed use elections for these purposes. As Hirano and Snyder (2012) find, incumbents involved in scandals are not only more likely to face primary challengers, but also have a much higher chance of losing their primary elections. Regarding selection, Hirano and Snyder (2014) present evidence that in House, Senate and gubernatorial races, primary voters are more likely to vote for high quality candidates (i.e. those with prior office-holding experience and with newspaper endorsements). And in lieu of primary reform, not only do (general) elections act as a bulwark against fringe legislators, but so too do internal legislative institutions, such as the committee assignment process, overseen by party leaders looking to maintain their party's brand (Erikson and Wright, 2000; Canes-Wrone, Brady and Cogan, 2002; Hall, 2015).

1.1.1 Elections and Representation

Upon observing high reelection rates and myriad incumbents across the country with a decade or more of experience in office, term limits advocates argued that representatives were out of touch with their districts, and that their extended incumbency was inimical to citizens' welfare. A more benign interpretation of repeated electoral success, consistent with a selection model of representation described below, is one of constituent satisfaction in which voters select a representative who remains constituency-oriented and competent and thus return her back to the legislature. In ballot pamphlets, however, sitting legislators were described as being far more interested in developing "cozy relationships" with "special interests" than with looking after their district's needs.

[2]To be clear, though, elections alone are insufficient. Ancillary institutions, from media outlets to issue advocacy groups, are also needed to help voters acquire information about candidates. If voters are to effectively select or "de-select" candidates, these ancillary institutions play a critical role, particularly in state legislative elections where few incumbents face serious challenges (Mansbridge, 2009).

And so, driven in part by a concern that incumbent legislators had developed a carapace of invulnerability, insulated from electoral defeat even while putatively neglecting their districts, twenty-one states passed some form of state legislative term limits with the goal of, inter alia, restoring accountability and responsiveness to the legislature. As of April 2015, fifteen states jointly housing 37% of the total U.S. population still limit state legislators' length of service.[3]

The voter-representative nexus is one of the key features of a functioning democracy. While scholars have developed different theories to understand the principal-agent relationship between voters and representatives, two prominent explanations center on sanctioning (Barro, 1973; Mayhew, 1974) and selection (Fearon, 1999; Mansbridge, 2009). The former is nicely encapsulated by Mayhew's (1974) pithy assumption that legislators are single-minded seekers of reelection. If legislators' primary motivation is reelection, what keeps them from shirking or acting contrary to their principal's interests is the threat of electoral defeat. Allowing for heterogeneity in legislators' motivations, a selection model of representation envisions elections being used to select competent, internally-motivated representatives who share the district's goals and will work to achieve those goals even when the threat of sanctions is absent.

These two conceptions of representation lead to manifestly different expectations about legislative behavior in the presence of term limits. In a sanctioning model, absent reelection incentives, representatives will shirk (i.e. reduce their effort or adopt out-of-step ideological positions) since their poor performance cannot be punished. In a selection model, by contrast, the absence of a reelection opportunity ought not alter the representative's behavior since a reelection incentive was not her primary motivation. The model that more accurately captures legislative behavior in practice is an empirical question, one explored in the first

[3]These states are Arizona, Arkansas, California, Colorado, Florida, Louisiana, Maine, Michigan, Missouri, Montana, Nebraska, Nevada, Ohio, Oklahoma, and South Dakota. In Idaho, Massachusetts, Oregon, Utah, Washington and Wyoming, term limit laws were initially passed but eventually repealed. See National Conference of State Legislatures, http://www.ncsl.org/research/about-state-legislatures/chart-of-term-limits-states.aspx, accessed April 19, 2015.

two substantive chapters.

To adjudicate between the two conceptions of representation, Chapter II offers a notable contribution by studying term limits experimentally in the Arkansas state senate. Examining the legislative behavior of Arkansas state senators is especially promising since the prospects for advancement beyond the state senate tend to be circumscribed in low professionalization legislatures (Maestas et al., 2006; Squire, 1988). Thus, if ever there were an opportunity for a representative to shirk, it would likely occur within these legislative environs. To provide additional insight into term-limited legislative behavior, Chapter III uses a modified research design proposed in Chapter II to unpack the strategic participatory decisions of term-limited legislators in the lower chamber of the Arkansas General Assembly where the electoral connection is not completely severed since these members are in a position to prolong their political careers by running for a seat in the upper chamber. Both chapters highlight systematic differences in the electoral fortunes of term-limited and non-term-limited legislators—specifically vote share and uncontested rates—suggesting that voters use elections to oust underperforming legislators.

1.1.2 Parties and Representation

As the Democratic and Republican parties have grown increasingly polarized, adjectives such as broken and dysfunctional have become synonymous with the nation's deliberative bodies (Mann and Ornstein, 2006; Bonica, 2013*b*). Ideologically extreme legislators are seen as the culprit for legislative paralysis, and in response, states have started to experiment with changes in electoral rules. In Alaska, California, Louisiana and Washington, voters have approved measures to adopt open primaries where the top-two vote getters, regardless of party, would appear on the general election ballot.

While supporters believed the reform would redound to moderates more amenable to compromise, the results to date have been far from sanguine, failing to uncover any evidence that open primaries reduce state-level polarization or that voters reward moderate candidates

(McGhee et al., 2014; Ahler, Citrin and Lenz, 2014). With the two political parties becoming more distinct not just in Congress but also across a majority of the states, one might expect to observe more extreme policy outcomes (Shor, 2013). As Lee (2015) explains, however, even with the rise of the Tea Party, replete with its hardline conservative members, it does not appear to be the case that polarization begets extremism in policy outcomes, which begs the question, why not?

Partisan theories of Congress describe the majority party as a cartel that controls key positions of power within the chamber as well as valuable resources that can aid its members' reelection efforts (Cox and McCubbins, 1993, 2005). Central to the cartel are party leaders intent on developing and maintaining a favorable brand name, which depends on a corpus of legislative accomplishments. Should rank and file members of any ideological persuasion undermine that objective, they can expect harsh repercussions. And party leaders have a number of tools at their disposal to discipline their members; one especially notable prerogative is their wide latitude over committees, not only determining their creation, jurisdiction, resources and membership but also influencing their policy outputs via control over which bills receive consideration on the floor. As studies of Congress have shown, majority party leaders are not shy about flexing their muscle, providing or withholding a variety of perquisites to members, including committee chair posts (Cox and McCubbins, 1993), targeted spending for a district (Carroll and Kim, 2010) and even campaign contributions from majority party leaders' PACs (Jenkins and Monroe, 2012).

I add to the sizable literature exploring how parties constrain member behavior at the state level. Chapter IV presents evidence that more ideologically extreme legislators in the North Carolina General Assembly are not only perceived to be less effective than their more moderate colleagues, but are also less likely to occupy key agenda-setting posts within their respective chambers. But it is not just internal legislative institutions that effectively reign in ideological extremism, for just as with term limits, elections—general elections, in particular—play a key role as well.

Extant empirical studies paint a clear picture, one in which voters favor moderates and levy a steep cost for extreme positioning. Examining roll-call data, Erikson and Wright (2000) find that more moderate roll-call voting records are associated with better electoral performance. In a similar vein, both Bender and Lott (1996) and Canes-Wrone, Brady and Cogan (2002) find that more ideologically extreme legislators tend to lose vote share in their next election. More recent studies investigating the relationship between candidate positions and electoral performance only serve to reaffirm earlier conclusions. Work by Hall and Snyder (2014) and Hall (2015) finds that nominating an extreme candidate, instead of a moderate, decreases the party's general election vote share, the probability that the party wins the seat and the share of contributions from PACs. Collectively, then, the findings in Chapter IV, in conjunction with the existing literature, spotlight two channels—both within and outside the legislature—wherein existing institutions mitigate the clout of ideologically extreme legislators, in turn raising doubts about the need for sweeping institutional reforms.

1.2 Plan for the Dissertation

In seeking to shed light on the relationship between legislative institutions and legislative behavior, this dissertation addresses the following questions:

- what impact do term limits policies have on the legislative output of state legislators?
- how do term limits affect the participatory decisions of state legislators?
- what is the effect of ideology on a legislator's standing within the chamber?

The remainder of this dissertation takes up these questions in turn. The next chapter, coauthored with Rocío Titiunik, presents experimental evidence on a question that has long been of interest to political scientists, namely, how does the absence of a reelection incentive affect legislative behavior. Empirical research on this question has produced a large body of evidence suggesting legislators engage in participatory shirking. We begin by providing

evidence that a commonly used observational research design comparing term-limited and non-term-limited legislators rests on an invalid counterfactual comparison due to the systematic differences between the two groups. Then drawing on several constitutional provisions in the Arkansas General Assembly, we experimentally examine the effect of term limits on legislative output. In contradistinction to the extant literature, across four outcomes we show that the output of term-limited and non-term-limited lawmakers is indistinguishable from zero. Failing to detect any evidence of a diminution of effort in the absence of a re-election incentive bolsters the case for the selection model of representation. Although the chapter examines a citizen legislature, the findings might carry implications for the six other low-professionalization legislatures that currently have term limits, which comprise almost half of the total number of states where term limits are currently in effect.

In Chapter III, I continue examining term-limited legislative behavior, though do so using observational data. I develop a career-concerns theory of legislative behavior, arguing that term limits alter the career incentives state legislators face; since legislators can no longer build long careers within a chamber, they must consider how their current decisions in office affect their future goals. Given the systematic differences between term-limited and non-term-limited legislators, I take a design-based approach to the study of term-limited legislative behavior, which endeavors to reduce unit-level heterogeneity. I use a panel of bill activity in the Arkansas state house to show that on proactive activities, such as bill introductions and constituency service, over which legislators have considerable discretion, term-limited legislators are significantly more active than their non-term-limited colleagues, but on reactive activities, such as roll-call votes, over which legislators have limited discretion, they are significantly less so.

Based on these seemingly contradictory findings, one obvious question is, why does shirking exist in one chamber but not the other? One possible explanation centers on differing incentives for legislators in the two chambers. Progressively ambitious term-limited legislators in the state house still retain a clear opportunity for upward mobility by pursuing a

state senate seat. But the pattern of behavior whereby term-limited legislators are active on some fronts but far less active on others is likely to be much less common among term-limited state senators in low professionalization bodies, since the prospects for advancement are generally limited, relative to their counterparts in high professionalization legislatures. As a result, many state legislators return to private life after 14 years in the state legislature (Maestas et al., 2006; Squire, 1988). This dynamic suggests that because most term-limited state senators do not go on to run for higher office, they do not face the same set of incentives to tailor their legislative activities in a manner that appeals to a broader constituency while also avoiding tough votes that might endanger their electoral prospects.

Chapter IV moves away from the subject of term limits, turning to the question of how ideology affects a legislator's standing in the chamber. I combine survey measures of legislative effectiveness and powerful committees with dynamic ideal point estimates to show that more ideologically extreme members of the North Carolina General Assembly are perceived as less effective and are also less likely to chair powerful committees. By showing that parties have the capacity to limit their members' legislative influence, the results have implications for states looking to overhaul their electoral rules. And Chapter V offers concluding remarks.

CHAPTER II

Term Limits and (the Absence of) Legislative Shirking: Experimental Evidence from the Arkansas Senate

Abstract: We analyze the impact of term-limit policies on the legislative output of state lawmakers. Several non-experimental studies have studied related questions, but inferences are complicated by the dynamic biases that may arise as a result of electoral defeat. We discuss and illustrate these methodological challenges with data from several states that adopted term limits, and show that paying close attention to research design clarifies the methodological obstacles. We then present an original experimental study that addresses many of these obstacles and is based on the random assignment of term length that occurs in the Arkansas Senate, which in turn induces the random assignment of term limits in 1997 and 2007. Across four measures of legislative output—bills introduced, bills passed, resolutions and abstention rates—we find no evidence of term limits effects. Since our experimental sample is small, we perform randomization-based inference, which is exact in finite samples. We also test the null hypothesis that term-limited lawmakers differ from their reelection-eligible counterparts, which allows us to assert that term limits decrease legislative output by at most a very small amount. Finally, we use bounds to address the fact that some senators in our original experimental sample are not observed.

2.1 Introduction

In the first Constitution of the United States, the Articles of Confederation, legislators were limited to no more than three out of every six years in office.[1] And during the Con-

[1]This paper is coauthored with Rocío Titiunik. We thank Steve Cook, the Chief Council of the Arkansas State Senate, for providing details regarding the drawing of lots following censuses; Blake Potts at the Arkansas Legislative Digest for providing the roll-call data; the Roy Pierce Fellowship at the Center for Political Studies, University of Michigan, for financial support; and Jeremy Gelman, Marjorie Sarbaugh-Thompson, Chuck Shipan, and participants at the 2014 State Politics and Policy Conference for thoughtful comments and discussions.

tinental Congress, Thomas Jefferson famously advocated for term limits to "prevent every danger which might arise to American freedom by continuing too long in office" (Jefferson, 1893, 61). By the time the subject resurfaced in the 1990s, when nearly half of the states passed term limits initiatives, term limits policies were seen in a less positive light as the goal of increased turnover was weighed against its potential costs— in particular, against the possibility that removing reelection incentives would lead legislators to decrease their effort or adopt 'out of step' ideological positions.

The behavior of representatives in their last term—that is, in the absence of reelection incentives—has been studied extensively, since its examination reveals important aspects about democratic representation. In particular, establishing whether last-term effects exist empirically and measuring their magnitude is important for several reasons. First, from a policy point of view, it is crucial to understand whether rules that allow representatives to govern without the prospect of future electoral accountability (e.g., term limits or lame-duck congressional sessions) result in systematic changes in their behavior and to consider those potential effects in the design and evaluation of policies. Second, from a theoretical point of view, the expectation that representatives will alter their behavior absent reelection incentives follows from an understanding of elections as a sanctioning mechanism in which the anticipation of punishment induces politicians to behave in particular ways. Thus, establishing the existence (or absence) of last-term effects has direct implications for our understanding of the role of elections in ensuring representation.

To date, a large amount of empirical evidence points to the dangers of severing electoral accountability. Studies of governors (Besley and Case, 1995; Alt, de Mesquita and Rose, 2011), members of Congress (Rothenberg and Sanders, 2000; Herrick, Moore and Hibbing, 1994) and state legislators (Carey et al., 2006; Powell, Niemi and Smith, 2007) have all found evidence consistent with last-period shirking. Besley and Case (1995), for example, find that term-limited governors are less proactive in responding to natural disasters.

Homing in more specifically on state legislatures, the existing literature suggests that

term limits adversely affect legislative behavior. For example, Carey et al. (2006) and Powell, Niemi and Smith (2007) find that term-limited state legislators devote less time to securing rewards for their district and helping constituents deal with government. In addition, Sarbaugh-Thompson et al. (2004) report that term-limited legislators turn their attention away from constituents and toward interest groups. The nonexperimental nature of the aforementioned studies, however, complicates the causal interpretation of these findings.

In this chapter, we study the impact of legislative term limits on state lawmakers' legislative output and participation, making two important contributions. First, we describe and document the methodological challenges that arise when this question is studied with non-experimental or quasi-experimental designs. In contrast to experimental designs, in natural experiments and non-experimental designs a valid comparison group may not be readily identifiable (Sekhon and Titiunik, 2012). We show that this methodological challenge arises in non-experimental studies of term limits effects. Any study of the impact of term limits must deal with the challenge that politicians serving their last term in office are often systematically different from those whose electoral horizons are longer, which complicates the ability to attribute last-term behavior to the lack of electoral incentives. This phenomenon is most evident when the decision to retire is entirely under the control of the individual politicians—e.g., legislators may retire preemptively when they anticipate that their poor past performance may result in a loss. But these inferential complications do not necessarily disappear when the occurrence of the last term is determined by an exogenous rule. As we discuss in detail below, while term limit laws prevent state legislators who wish to run indefinitely for reelection from doing so, they fail to eliminate the biases that may arise as a result of electoral defeat.

These methodological obstacles are the motivation for our second contribution, which is the study of term limits effects using an experimental strategy in the Arkansas Senate. The Arkansas Constitution includes two features—the random assignment of senators' term length in the first election after reapportionment and term limits—the combination of which

results in the random assignment of state senators to lame-duck status. In particular, senators randomly assigned four-year terms in the first election after reapportionment see one less session in office than those randomly assigned two-year terms, which gives us the unique opportunity to examine two legislative sessions in 1997 and 2007 where the group of legislators assigned four-year lots is term-limited while the group assigned two-year lots is still eligible for reelection. Drawing on these experiments, we empirically test hypotheses regarding the effects of term limits on several measures of legislative participation and output.

Our empirical analysis addresses several statistical challenges. Since the Arkansas Senate has only thirty-five members, the overall sample size in our experiment, which pools two cohorts, is relatively small. To address this issue, we use randomization-based inference techniques that are exact in finite samples. As we explain below, in the randomization-inference framework, the distribution of the test-statistic under the null hypothesis of no effect is entirely determined by the treatment assignment, which allows us to test the hypothesis of no effect with an exact finite-sample p-value instead of relying on large-sample approximations that may be inadequate with our small sample size.

Since we are unable to reject the null hypothesis of no term limits effect and are interested in asserting that term limits do not alter legislative output, we also test the hypothesis that the average outcomes of term-limited lawmakers differ from the average output of reelection-eligible lawmakers. In these tests, which are common in biomedical studies, the type I error rate is the probability of declaring that the two groups compared are equivalent when in fact they are not (Berger and Hsu, 1996). Thus, in setting the type I error rate at a given level, say 5%, we control the probability of declaring that term limits have no effect when in fact they do. For every outcome, we calculate the minimum discrepancy between the term-limited and non-term-limited group that leads to a rejection of the null hypothesis that both groups are different. As we show, these tests of equivalence allow us to assert with 95% confidence that term limits do not have large effects on legislative output and participation. In particular, at this level, we can rule out all *negative* term limit effects except for effects

of very small size.

Finally, since we have some attrition in our experimental sample, we estimate bounds on the average treatment effect for every outcome under the assumption that those lawmakers whose outcomes we do not get to observe would have been systematically high or low, potentially affecting our conclusions (see, e.g., Manski, 2003). This analysis indicates that non-random attrition does not seem to be driving our conclusions.

After accounting for these challenges, we find no evidence that term-limited legislators reduce their effort by introducing or passing fewer bills, providing less constituency service as proxied by resolutions, or abstaining at a higher rate on roll-call votes. Our findings are at odds with several observational analyses of state legislators' behavior in their last terms in office and have important policy implications for states with term limits.

The remainder of the paper is organized as follows. In the next section, we review existing research on the relationship between last terms and shirking and then develop theoretical expectations about last-term effects on legislative behavior that take into account the specific institutional constraints of state legislatures. From there, we discuss the difficulties faced by observational studies of legislative last-term effects and propose an improved observational design. We then present the details of our experimental research design. Next, we present our results, followed by a section that presents robustness checks based on bounds to address the issue of attrition in the original experimental sample. We conclude in the last section. Additional results are presented in the Appendix.

2.2 Existing Literature on Last Terms and Shirking

In some of the earliest work on the subject, scholars compared retiring members of Congress to those who would be returning to the chamber, finding myriad instances of shirking, including higher abstention rates and fewer bill introductions, amendments and trips back to one's district (Zupan, 1990; Herrick, Moore and Hibbing, 1994; Rothenberg

and Sanders, 2000).[2] More recently, institutional features of legislatures that potentially facilitate shirking have been well-documented, with two of the most prominent being lame-duck sessions in the U.S. Congress and the 17th amendment, which ended the indirect election of U.S. senators.

Treating the passage of the 17th amendment as a natural experiment, Bernhard and Sala (2006) examine patterns of roll-call voting of U.S. Senators and find that the shift in selectorates from state legislatures to state electorates corresponded with more moderate late-term NOMINATE scores. And Gailmard and Jenkins (2009) argue that bringing an end to monitoring and sanctioning by state legislators, who are relative political experts, and delegating those responsibilities to voters created an institutional context that placed fewer constraints on senators' behavior.

Scholars have also analyzed lame-duck U.S. Congress sessions, which occur after an election but prior to the start of a new Congress, and give outgoing legislators, who either lost or will retire, a few remaining weeks to legislate without facing electoral repercussions. In pre-20th amendment lame-duck sessions, Jenkins and Nokken (2008) find that the voting behavior of departing members of Congress moved farther away from the median party position during lame-duck sessions than it did for returning members of Congress.[3] And in more recent, post-20th amendment lame-duck sessions, Nokken (2013) finds that lame-duck legislators abstain on roll-call votes at rates twice as high as their returning colleagues. These studies, coupled with earlier work on retiring legislators, point to two principal ways in which legislators are likely to shirk: through participatory shirking in which the member reduces his or her effort and through ideological shirking in which the member changes his or her roll-call behavior. We focus on the former.

Term-limited legislatures provide another setting in which to examine shirking. One of the advantages of analyzing term limits-induced lame-duck status is that these members of

[2]See also the June 1993 issue of Public Choice dedicated entirely to shirking.

[3]The 20th amendment reduced the time between election day and the start of presidential and congressional terms, moving it up to January from March.

the legislature serve a full two- or four-year term, whereas lame-duck sessions tend to be hurried affairs that last a few weeks. Term limits advocates believed that because legislators would not be as preoccupied with reelection-centered activities, they would expend more effort on policymaking and related legislative activities (Will, 1992; Glazer and Wattenberg, 1996). However, based on surveys and case studies, the scholarly literature on term limit effects largely aligns with work at the congressional level and finds that lame-duck state legislators engaging in shirking.

Drawing on a 50-state survey of state legislators, both Carey et al. (2006) and Powell, Niemi and Smith (2007) note that legislators in term-limited states (i) spend less time keeping in touch with their constituents and handling casework than legislators in non-term-limited states and (ii) are more likely to prioritize their own conscience above the interests of the district and (iii) within term-limited chambers, the more terms of eligibility legislators have, the more effort they are likely to expend.

Case studies also reveal that term-limited legislators in Michigan's lower chamber attended interest group-sponsored events in the state capital at higher rates, suggesting they were more concerned with their future employment opportunities than with their current district and constituents (Sarbaugh-Thompson et al., 2004). Drawing on interviews with legislators, lobbyists and knowledgeable observers, Berman (2007) finds that lame-duck state legislators in Arizona acted in a manner they thought to be unpopular with their constituents due to the absence of electoral sanctions.

In short, there is an abundance of empirical research showing that lame-duck status adversely affects legislative behavior in different settings. Below we develop theoretical expectations about last-term effects that take into account the specific institutional constraints of our application, which centers on term limits in a state legislature.

2.3 Legislative Behavior Absent Reelection Incentives

The most fundamental mechanism by which last-term effects are expected to arise is the removal of reelection incentives. Under a sanctioning model of representation, voters incorporate politicians' past actions into their voting decisions and elections serve as an accountability mechanism that sanctions representatives' behavior (Mansbridge, 2009). In turn, the threat of punishment induces reelection-seeking politicians to behave in accordance with constituents' preferences and expectations. Under this model, the logical consequence of adopting term limits is to enable undesirable legislative behavior or shirking, as the threat of punishment is removed and legislators have no incentive to please the electorate.[4] Thus, if elections' main role is to serve as an accountability mechanism, removing the possibility of running for reelection should result in systematic changes in legislative behavior (Fearon, 1999, p. 63).[5]

There may also be mechanisms by which the removal of reelection incentives results in lower legislative participation and output that are not directly related to the removal of electoral accountability. One such mechanism is the potential opportunity costs of seeking future employment. If legislators harbor some degree of progressive ambition and hope to secure their next occupation—higher office or otherwise—before their tenure comes to an end, lame-duck legislators face a trade-off: continue to actively participate in legislative activities or curb some of that legislative participation in order to invest attention in surveying their future employment options. Given the time commitments associated with casework and with building the coalitions needed to successfully navigate bills through the legislature, we might expect those who are in their last term, and thus more likely to be in search of their next job, to reduce the effort they expend on constituency service and policymaking.

These two non-exclusive scenarios—shirking induced by the removal of reelection incentives and shirking induced by the opportunity costs of securing future employment—imply

[4] See also Barro (1973) and Ferejohn (1986) for models on the control of politicians via reelection incentives.

[5] We consider critiques of the sanctioning model in the conclusion.

that term-limited legislators will have less time or incentives to meet with staff to help draft legislation, to learn about the kinds of bills their colleagues are sponsoring, to jockey for support in committee and on the floor to ensure passage of those bills that they do introduce, to allocate attention to casework and to attend roll-call votes. Accordingly, we can hypothesize that at the level of the individual legislator, term-limited members of the chamber will reduce their effort, (1) introducing fewer bills, (2) passing fewer bills, (3) performing less constituency service and (4) abstaining on a greater proportion of roll-call votes than non-term-limited members.

Legislative term limits vary enormously in their severity (see review by Sarbaugh-Thompson, 2010). Some states only impose bans on consecutive years of service, allowing legislators to cycle back and forth between both legislative chambers with minimal restrictions. This alters but does not completely remove reelection incentives, as legislators may plan to come back to their old district once they spend the mandatory number of years out of office. In contrast, some states impose lifetime bans on reelection, completely eliminating the possibility of ever running for the same seat and inducing a much more drastic reduction of electoral incentives.

Last-term effects will also depend on the degree of the legislature's professionalism. Facing low pay, limited staff resources, poor advancement prospects and the absence of a reelection incentive, members serving in one of the twenty-four states with low-professionalization legislatures have few incentives to actively participate (Squire, 1988; Maestas, 2000). In these low-salary and limited-resource settings, state legislators must find time to negotiate the balance between outside careers, from which they derive their primary source of income, and their legislative obligations, potentially leaving them with scant time to devote to legislative tasks. By contrast, in professional legislatures, not only do staff subsidize the cost of policymaking, but these legislators earn salaries that permit them to devote all of their time to legislating. The incentives introduced by the removal of electoral accountability may therefore be amplified in less professional legislatures.

However, low professionalism could also attenuate last-term effects, in particular through

mitigating the need to secure future employment. Less professionalized legislatures tend to conduct the preponderance of legislative business in regular sessions held in odd-numbered years. Thus, the trade-off between legislating and searching for future employment may never arise; instead, most legislators might be able to make the required three-month commitment every other year while maintaining an alternative source of employment on which they can continue to rely after term limits are imposed. Below, we use our experimental study to investigate these hypotheses empirically.

2.4 Research Designs to Study Last-Term Effects

A possible strategy to study term limits effects is to compare the outcomes of term-limited legislators to the outcomes of those legislators who can still run for reelection. This type of observational research design is a common choice (see, e.g., Carey, Niemi and Powell, 1998; Carey et al., 2006; Powell, Niemi and Smith, 2007; Sarbaugh-Thompson et al., 2004). However, the two groups compared in this design may be systematically different and thus threaten the validity of the inferences. This may occur for at least two reasons.

First, legislators who are serving their last term because of term limit restrictions have by construction survived the highest possible number of elections that a legislator is allowed to contest before term limits come into effect. In contrast, legislators who are not yet term-limited are of two types: the "departor" type, composed of legislators who will be defeated and will depart before term limits are binding, and the "survivor" type, composed of legislators who will win all elections and serve the highest possible number of terms under the current term limit laws. The result is that, at any given point in time, the group of term-limited legislators is composed entirely of survivors, while the group of non-term-limited legislators is composed of both survivors and departors.

The differences between both groups will likely be systematic and related to the outcomes of interest. Since departors are candidates that will eventually be defeated, they are likely to be of lower average quality—i.e., less competent—than survivors, making the non-

term-limited group of lower average quality than the term-limited group (since the latter is composed exclusively of survivors).[6] Differential quality between both groups might lead to systematic differences in future performance for a number of reasons. For example, strong challengers may be deterred at higher rates in the term-limited group due to the higher quality of incumbents in this group. Naturally, it is not possible to distinguish both types in the group of legislators who are not yet term-limited, and thus these underlying differences cannot be "controlled for."[7]

Second, the group of term-limited legislators may differ systematically from the group of non-term-limited legislators due to the incumbency advantage and the phenomenon of strategic waiting by challengers. Term-limited legislators are, by definition, veteran incumbents. Thus, the incumbency status they enjoyed at the time of their last election could have translated into a less competitive election due to the advantages brought about by incumbency, including increased name recognition or possibly a weaker challenger. In contrast, depending on the specific term-limits restrictions, some or all legislators in the non-term-limited group will have been elected in open-seat races. These races will tend to be more competitive than incumbent races, which may translate into higher pressure to display good performance. Moreover, differences in the degree of competitiveness between term-limited and non-term-limited legislators could arise because challengers might prefer to wait until the seat becomes open instead of challenging an incumbent in his last eligible election. The result would be higher rates of uncontested races and as a result less competition among term-limited legislators.

In sum, either because of intrinsic differences in candidate quality due to the differ-

[6]The result that more competent legislators are more likely to survive reelection is standard in agency models of electoral selection (see, for example, Alt, de Mesquita and Rose, 2011, and references therein).

[7]This difficulty applies not only to studies of term limits, but also to other types of studies such as those that compare retiring and non-retiring Congress members in lame-duck sessions (see, e.g, Carey, 1998). Note, however, that the situation might be inverted in lame-duck congressional sessions: since there are no term limits in the U.S. Congress, those retiring might have served less terms on average than those who are returning to the chamber, which might result in returning members being of higher average quality than departing members. This is more likely to occur, for example, when retirements are due to anticipated bad performance than when they occur because the member has reached retirement age.

ent number of terms survived, differences in incumbency status, or strategic waiting by challengers, and likely due to the combination of all of these factors, non-experimental comparisons of term-limited versus non-term-limited legislators may not lead to valid estimates of term-limits effects. To provide evidence of these inferential problems, we compare the electoral fortunes of term-limited and non-term-limited legislators in their most recent election in eight state legislatures—four with consecutive service bans and four with lifetime service bans. We pool observations from several election cycles, comparing cross-sectionally the mean vote shares and the proportion of uncontested races between term-limited and non-term-limited legislators in their most recent election.

The first four rows in Table 2.1 report results from the lower and upper chambers in the Colorado, Ohio, Arizona and South Dakota legislatures, all of which have consecutive service bans. We see that there is a substantively large and statistically significant difference in vote share in all four states: term-limited (TL) legislators in these states enjoy vote shares between 4-5 percentage points higher than non-term-limited (NTL) legislators. In two states, term-limited legislators are also significantly less likely to face a challenger.

The remaining four rows in Table 2.1 report results from the lower chamber in the Arkansas legislature and the lower and upper chambers in the California, Michigan and Oregon legislatures, all of which have lifetime service bans. Again, term-limited representatives have significantly higher vote shares than non-term-limited representatives, though the size of the differences is more variable, ranging from roughly 2 to 10 percentage points depending on the state. Moreover, with the exception of Arkansas' lower chamber, term-limited legislators are not significantly less likely to face a challenger.

Overall, the results show that, on average, term-limited legislators are more electorally successful than reelection-eligible legislators, a result consistent with the aforementioned possible confounders. Moreover, these vote share differences are not solely explained by the higher rate of uncontested races among term-limited legislators. As we show in the Appendix, the differences in vote shares persist when uncontested races are excluded from

Table 2.1: Uncontested Rates and Difference-in-means for Vote Share Between Term-Limited and Non-Term-Limited State Legislators

	Vote Share			Uncontested Rates			
	Mean NTL	Mean TL	p-val	Mean NTL	Mean TL	p-val	N
States with Consecutive Bans							
Arizona	49.45	53.74	.09	10.97	16.30	.14	630
Colorado	67.41	71.51	.01	14.33	21.28	.04	741
Ohio	68.19	72.60	.00	11.40	17.39	.03	921
S. Dakota	47.20	52.02	.04	9.34	14.02	.14	728
States with Lifetime Bans							
Arkansas (House)	83.45	93.92	.00	57.67	82.40	.00	798
California	65.73	67.92	.01	4.60	4.48	.93	898
Michigan	66.71	68.31	.06	1.41	1.25	.85	1030
Oregon	66.54	71.85	.02	11.32	18.18	.17	225

Note: The election data come from the State Legislative Election Returns (1967-2010) ICPSR #34397 dataset. Column labeled 'Mean TL' reports the mean outcome for term-limited senators, and column labeled 'Mean NTL' reports the mean outcome for non-term-limited senators. The columns labeled 'p-val' report p-values from two-tailed t-tests of the null hypothesis that means are equal. Column labeled N reports the number of observations used in each row.

the analysis.

Below we present an experimental design that addresses most of these challenges, but first we consider a non-experimental research design that may mitigate these inferential problems in cases where an experimental design is not available. We propose a research strategy that restricts the comparison group to only those non-term-limited legislators who have successfully won at least one reelection bid. There are several reasons why eliminating freshman legislators from the comparison group might alleviate some of the biases. First, if most low-quality incumbents are defeated in their first reelection, restricting the comparison group to non-freshman incumbents may eliminate weaker incumbents and thus increase the average candidate closer to the average quality among survivors in the term-limited group. The systematic differences between the groups could also be alleviated by the fact that most candidates eliminated from the non-term-limited group will be incumbents elected in (relatively) competitive open seats. The veteran incumbents who stay in the non-term-

limited group after freshmen are eliminated are thus likely to have been elected, on average, in an environment more similar to the environment faced by non-term-limited incumbents.

Table 2.2: Uncontested Rates and Difference-in-means for Vote Share Between Term-Limited and Non-Term-Limited State Legislators, Excluding Freshman Legislators

	Vote Share			Uncontested Race			
	Mean NTL	Mean TL	p-val	Mean NTL	Mean TL	p-val	N
States with Consecutive Bans							
Arizona	54.32	53.74	.83	14.65	16.30	.69	447
Colorado (House)	71.29	74.01	.21	19.53	25.00	.29	373
Ohio (House)	71.66	73.14	.32	15.42	19.08	.32	559
S. Dakota	52.15	52.02	.96	12.56	14.02	.69	521
States with Lifetime Bans							
Arkansas (House)	91.19	93.92	.00	75.29	82.40	.05	495
California (House)	68.04	68.15	.93	5.79	3.66	.30	450
Michigan (House)	69.63	68.98	.54	2.08	1.59	.68	540
Oregon (House)	72.08	70.35	.59	20.63	16.67	.60	111

Note: The election data come from the State Legislative Election Returns (1967-2010) ICPSR #34397 dataset. Column labeled 'Mean TL' reports the mean outcome for legislators who are term-limited and column labeled 'Mean Non-TL' reports the mean outcome for legislators who are not term-limited. The columns labeled 'p-val' report p-values from two-tailed t-tests of the null hypothesis that means are equal. Column labeled N reports the number of observations used in each row.

Consistent with our expectations, once we exclude freshmen from the group of non-term-limited legislators, uncontested rates and vote shares between the two groups become considerably more similar. Table 2.2 presents uncontested rates and vote shares from the same eight states in Table 2.1 after removing freshman legislators. Because the state senates in Colorado, Ohio, California, Michigan and Oregon only permit two four-year terms, all non-term-limited senators are freshman, so we cannot apply this research design for these chambers. For this reason, in these states we only examine lower chamber election results. By contrast, in Arizona and South Dakota legislators in both chambers can serve up to four two-year terms, so we continue to examine both the lower and upper chambers for these two state legislatures. With the exception of Arkansas' lower chamber, there are no distinguishable differences in uncontested rates and vote shares between term-limited and

reelection-eligible legislators. This suggests that observational studies of term-limits effects might be improved by excluding freshman legislators from the comparison group.

2.5 A Research Design Based on Random Assignment

Motivated by the aforementioned methodological challenges in observational studies of term limits, we use an experimental design that relies on the random assignment of term length, which in turn induces the random assignment of term limits later in the decade. This design avoids many of the inferential problems mentioned above, as the group of term-limited and non-term-limited legislators are on average identical at baseline due to the initial random assignment.

Our research design is based on the random assignment of term length in the Arkansas Senate. Arkansas senators normally serve a term of four years and their terms are staggered, with (roughly) half of the 35 senate seats up for election every two years. However, Article 8, Section 6 of the state's constitution mandates that, in the first election following a decennial census and the corresponding redrawing of district boundaries, all 35 seats must be up for election.[8] Since the simultaneous election of all 35 seats breaks the staggering of terms, term lengths are randomly assigned to return the chamber to staggered terms.

Specifically, Section 6, Amendment 23, of the Arkansas Constitution instructs senate seats to be randomly divided into two classes of size 17 and 18 after each reapportionment. The pattern of term length differs by class: senators elected to a seat in the class of size 18 serve a two-year term immediately following reapportionment and remain eligible for two four-year terms thereafter, while senators elected to a seat in the other class serve two successive four-year terms immediately following redistricting and a two-year term at the end of the decade. Senators draw lots at the beginning of the first legislative session immediately after redistricting to determine the composition of each class of seats. This design, and

[8]The motivation for this provision is to ensure that all sitting senators have been elected by their new constituencies.

similar designs in Illinois and Texas, was used by Titiunik (2015) to study the effects of term length on legislative behavior.

In November 1992, 60 percent of Arkansas voters supported Amendment 73, a term limits initiative that was among the most stringent in the country. This amendment limited state representatives' service to a lifetime maximum of three two-year terms and state senators' service to a lifetime maximum of two four-year terms. A critical element of our research design is that two-year terms do not count against the two-term limit for state senators—only four-year terms do.

Our goal is to measure outcomes for term-limited and non-term-limited senators during the same legislative session, to avoid conflating time differences and genuine last-term effects. For this reason, we study two cohorts of senators for whom term limits become effective during the same legislative session: those first elected or reelected in 1992, and those first elected in 2000 or 2002. All Arkansas senators elected in November 1992, whether elected for the first time or reelected, served their last period either in 1996-2000 or 1998-2002 (if they did not retire or lose sooner). Senators elected in 1992 for a two-year term could run for reelection in 1994 for a four-year term (1994-1998),[9] and again in 1998 for another four-year term (1998-2002), because the first two-year term did not count towards term limits. By contrast, those elected in 1992 for a four-year term could run for reelection only once in 1996 to serve a second four-year term between 1996 and 2000. In other words, since 1992 is the "baseline" year when term limits are adopted, regardless of how many times senators in this cohort had been reelected prior to 1992, they would all serve their last allowed term at the same time, except for the 2-year discrepancy induced by the staggering.

The situation for later cohorts is different, because as some senators lose or retire before the maximum allowed number of terms, the newly elected senators' last allowed terms occur at different points in time. If a few new senators were entering every year, it would be

[9]Throughout, we refer to the length of terms by an interval from an even year to another even year, such as 1998-2002. The first year in the interval indicates the year when the election took place, and the last year in the interval the last year of the term. For example, 1998-2002 refers to the term, served between January 1999 and December 2002, for which a senator was elected in November 1998.

hard to study an additional cohort, as everyone would be term-limited at different times, invalidating our design. Luckily, there are only six senators who are elected for the first time between 1994 and 1998, with the remaining 29 first elected in either 2000 or 2002. These 29 senators constitute the second cohort in our analysis—since two-year terms do not count toward term limits, a first election in 2000 or 2002 leads to a final term during either 2006-2010 (if a four-year term is drawn in 2002) or 2008-2012 (if a two-year term is drawn in 2002).

Figures 2.1 and 2.2 illustrate the sequence senators experience based on whether they draw a two-year or a four-year term in the 1990s and 2000s, respectively. As the figures show, there are two legislative sessions in the Arkansas General Assembly, the 81st in 1997 and the 86th in 2007, during which senators randomly assigned four-year terms following reapportionment are ineligible to run for reelection and thus are *lame ducks*, while those assigned two-year terms are eligible for their last reelection in the following year.

Figure 2.1: Illustration of 1997 Research Design

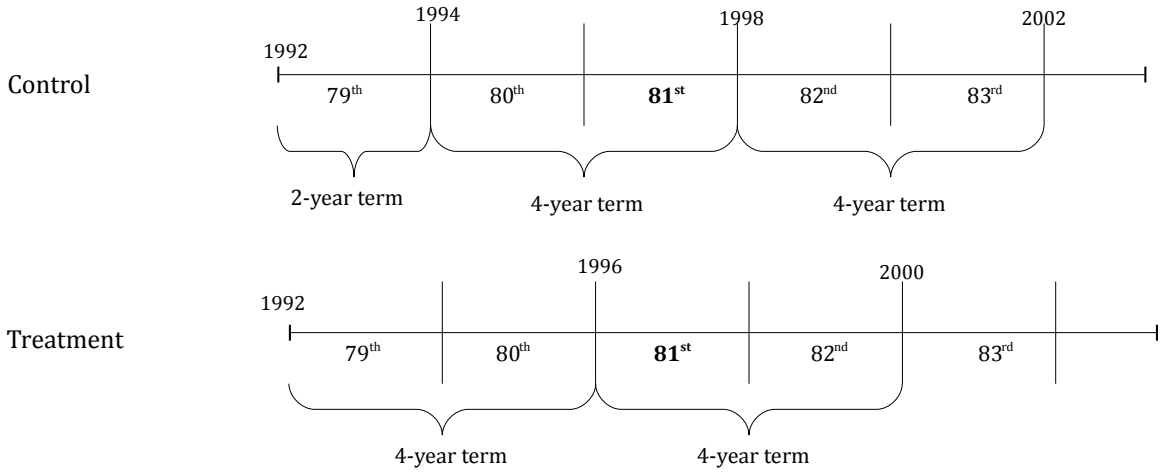
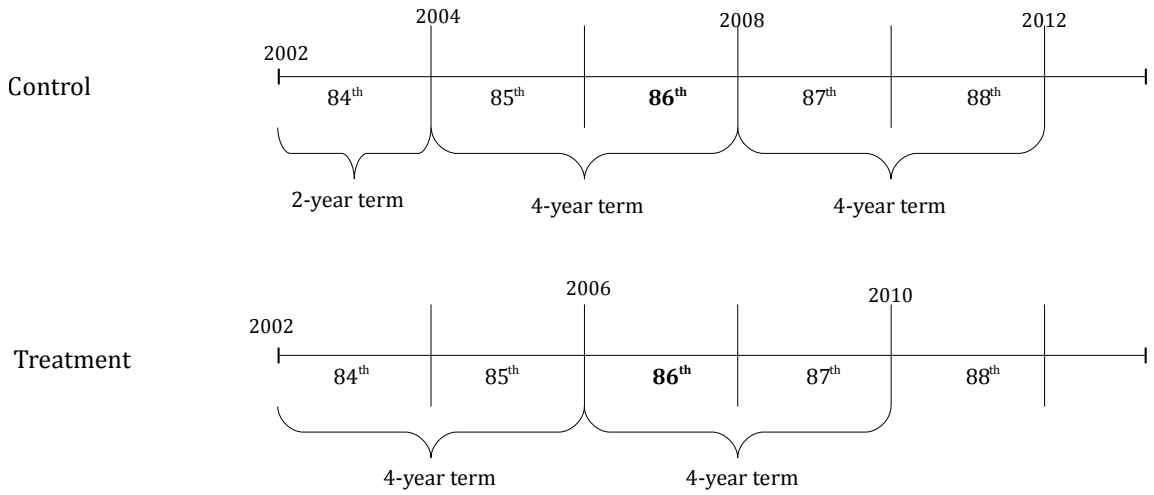


Figure 2.2: Illustration of 2007 Research Design



Consider two senators who entered the chamber in either 2000 or 2002 and won their races in 2002, one assigned a two-year term and one assigned a four-year term. Because a two-year term does not count toward the two-term lifetime limit, the senator assigned to serve a two-year term will stand for reelection in November 2004 and again in November 2008. By contrast, a state senator assigned a four-year lot in 2002 is already on the term-limit clock and will only stand for reelection one more time in November 2006. This in turn makes for a legislative session in 2007 (the 86th) where senators assigned four-year lots in 2002 are lame-ducks while senators assigned two-year lots in 2002 still face an election in 2008. The sequence is analogous in the 1990s.[10] For analysis, we pool both cohorts, totaling 64 senators (35 from 1992 cohort, 29 from 2000/2002 cohort), and study outcomes of interest during the 81st and 86th regular sessions.

Similar to observational studies, our design does require the Stable Unit Treatment Value Assumption (SUTVA) in order to yield effects that can be interpreted as the effects of term limits. When SUTVA holds, the outcome of every experimental unit is solely affected by

[10]Due to the confusion surrounding passage of Amendment 73, lots were drawn after the 79th session in October 1993 instead of at the beginning of the session. In the post-2000 reapportionment, by contrast, lots were drawn in December 2002, after the election but before the start of the legislative session.

the treatment received by that unit, regardless of the treatment status assigned to the rest of the units participating in the experiment (see,e.g., Rubin, 1990; Bowers, Fredrickson and Panagopoulos, 2013). In our research design, SUTVA requires that a legislator who is term-limited behave in the same way regardless of how many other legislators in the chamber are term-limited. This would restrict scenarios where, for example, non-term-limited legislators let term-limited legislators have a larger share of those resources that have a fixed budget (e.g. floor time) to help them take actions that will position them favorably in their quest for higher political office. However, given that term-limited legislators are not returning to the chamber, these agreements might be difficult to sustain in equilibrium (see Muthoo and Shepsle, 2010). Moreover, this kind of strategic coordination may be less likely to occur for outcomes that are not directly constrained by the actions of others (e.g. bill introductions). The fact that our results are consistent across abstention rates and bill introductions thus alleviates our concerns about possible SUTVA violations.

Validity of Experimental Research Design

We now provide evidence regarding the validity of the experimental research design just described. Under random assignment of term length, all predetermined characteristics at the senator level are identical in expectation between senators assigned a two-year term (henceforth '2-year senators') and senators assigned a four-year term (henceforth '4-year senators') following reapportionment. Thus, if we observed significant dissimilarities between our two samples, the validity of the randomization might be called into question.

We present the results from balance tests from our full sample of senators described above who drew lots in October 1993 and December 2002. Since the total sample size is 64, we tested the null hypothesis that there is no treatment effect on predetermined covariates using randomization inference instead of parametric tests.[11] We chose the difference-in-means between the term-limited and non-term-limited groups as a test-statistic, and used

[11]See Rosenbaum (2010) for an introduction to randomization inference methods.

the random assignment of state senators to either group to calculate the exact randomization distribution of this test-statistic under the null hypothesis of no treatment effect for any senator. Since this test is based on the known randomization of the treatment, it has the advantage that its distribution is exact even in small samples. In contrast, a t-test would require the validity of large-sample approximations, which may not be appropriate given our low sample size.[12]

As shown in the first four columns in Panel A of Table 2.3, both groups have similar means across seven pretreatment covariates, including vote share obtained in the previous election, party, and race. And using our randomization-based tests, we fail to reject the null hypothesis of no treatment effects for these covariates, as is expected in a successfully implemented experiment (minimum p-value across all covariates is 0.18). Given our relatively low sample size, however, our failure to reject the null hypothesis could be driven by a lack of power. For this reason, we also test the hypothesis that the term-limited and non-term-limited groups are different.

Letting μ_{TL} be the mean in the term-limited group and μ_{NTL} the mean in non-term-limited group, we test the null hypothesis that the discrepancy or dissimilarity between both means is larger than a positive number δ, that is, we test $H_0^\delta : |\mu_{TL} - \mu_{NTL}| > \delta$. These tests, sometimes referred to as equivalence tests, are commonly used in medical studies to establish bioequivalence between generic and brand-name drugs (Berger and Hsu, 1996). Whereas in the common balance tests reported above the null hypothesis is that there is no difference between term-limited and non-term-limited legislators, with equivalence tests we make the null hypothesis that the groups are different and only reject it when there is sufficient evidence that the two groups are similar. Whether the H_0^δ is rejected naturally

[12]Since the enumeration of all possible realizations of the treatment assignment is not feasible, we based our tests on 10,000 simulations—where each simulation takes one treatment assignment at random from all possible treatment assignments. Since the random assignment of terms was done separately for each cohort, in our simulations we separately drew the treatment assignment of each cohort, and then pooled both cohorts to compute the difference in means between our pooled treatment and control groups. This implementation respects the way in which the randomization was actually performed. However, analyzing all observations jointly ignoring that the two cohorts were separately assigned produces qualitatively similar results.

depends on the value of δ. We report the *minimum* value of δ for which H_0^δ is rejected at 5% level. To reject H_0^δ for a given $\delta_\star$ means that we can assert with 95% confident that the pretreatment covariates in the term-limited group differs from the pretreatment covariates in the reelection-eligible group by at most $\delta_\star$. Thus, if $\delta_\star$ is small, we can assert that the groups are similar.

We report the results from equivalence tests in the last two columns of Panel A in Table 2.3. We report the minimum value of δ for which H_0^δ is rejected at 5% in two ways: in raw units and in standard deviation units—i.e., the quantity $\delta_\star$ in raw units divided by each variable's standard deviation in the pooled sample. For example, the first row shows that we can be 95% confident that the difference between the vote share obtained by term-limited senators and the vote share obtained by reelection-eligible senators is no greater than 11.03 percentage points in either direction, an absolute difference that represents 0.8 pooled standard deviations. Given that the average vote share is so high in both groups—between 85% and 90%, a difference of at most 11 percentage points in either direction would still imply that senators in both groups were elected by a very large margin. We reach similar conclusions for the other covariates, although the minimum δ corresponding to the Democrat and Black variables, 0.32 and 0.21, respectively, is fairly high, a result we attribute to the low sample sizes. Taking the results from both the tests of no effect and the tests of equivalence, and considering the same tests separately by session—reported in the Appendix—we conclude that the random assignment of senators to groups was implemented successfully.

A final issue that we address is sample attrition. By 1997 and 2007, the years when the 81st and 86st Legislative Sessions begin, 15 senators in our sample of 64 had left the chamber: 12 senators left between 1993 and 1997 and 3 senators left between 2002 and 2007. This leaves us with a remaining sample of 49 senators, whom we call "compliers", which begets some complications to the design. Any time attrition occurs in an experimental setting it raises concerns that the remaining subjects no longer represent a random sample of the

Table 2.3: Covariate Balance Between Term-Limited and Non-Term-Limited Arkansas Senators, pooling 81st (1997-1998) and 86th (2007-2008) Legislative Sessions

	Means			Test of no effect	Test of $H_0^\delta : \lvert\mu_{TL} - \mu_{NTL}\rvert > \delta$	
	TL	NTL	Difference	p-value	Min δ for which H_0^δ is rejected Raw units	SD units
Vote Share	89.27	86.8	2.46	0.6	11.03	0.8
Married	0.88	0.88	0	1	0.17	0.7
Male	0.94	0.81	0.12	0.23	0.28	1.2
Democrat	0.91	0.75	0.16	0.18	0.32	1.2
Black	0.12	0.06	0.06	0.67	0.21	1
Attorney	0.25	0.34	-0.09	0.55	0.26	0.8
Age	50.66	52.78	-2.12	0.45	6.98	0.9
Sample size:	32 (TL group) and 32 (NTL group)					

Note: 'TL' refers to term-limited senators (assigned 4-year lot in 1992 or 2002), and 'NTL' refers to non-term-limited senators (assigned 2-year lot in 1992 or 2002). The test of no effect reports randomization-based p-values corresponding to the sharp null hypothesis that the treatment has no effect for any unit using the difference-in-means as test-statistic. In tests of the hypothesis H_0^δ reported in the last two columns, μ_{TL} refers to the mean outcome among term-limited senators and μ_{NTL} refers to the mean outcome among non-term-limited senators; these tests are also randomization-based, assuming a constant treatment effect model. SD units are expressed in terms of the pooled standard deviation across all senators.

original experimental sample, which in turn would lead to invalid inferences (Green and Gerber, 2012). This would occur in our case if a senator's defeat or retirement before term limits become binding is affected by the term length assigned after reapportionment—i.e., by our treatment of interest.

In this and future sections, we present arguments and evidence that suggest we can make meaningful inferences despite this methodological challenge. First, we note that attrition levels are comparable across treatment and control groups. As shown in the last row of Table 2.3, our initial sample size is 32 senators in each group, and after attrition there are 26 senators in the term-limited group and 23 senators in the non-term-limited group, samples sizes that are entirely consistent with a 1/2 probability of assignment to each group.[13] Moreover, this balance in attrition levels is also seen when we consider each legislative session

[13]The null hypothesis that the true probability of success is equal to 0.5 in 49 trials of a Bernoulli experiment cannot be rejected with 23 successes (p-value 0.7754). And the difference in compliance rates by group (23/32 and 26/32) are statistically indistinguishable (t-test p-value is 0.3840).

individually.[14]

Note also that, in addition to the initial randomization—which allows us to ensure comparability at baseline—a crucial aspect of our design is that both groups of senators have survived the same number of elections (one) when the outcomes are observed. As a result, the attrition that results from electoral defeat in the first reelection is likely to affect both groups equally, and the composition of both groups in terms of departors and survivors is thus likely to be similar at the moment when outcomes are measured. Moreover, this composition is equal (on average) at baseline due to the initial randomization.

Nonetheless, and despite losing roughly the same number of senators in each group, there could still be differences in the *type* of senators who drop out. For example, if the senators who drop out in the term-limited group are more productive on average than the senators who drop out in the control group, we might observe that the term-limited group has lower legislative output than the non-term-limited group, but we would be mistaken to attribute this difference to last-term effects. The assumption that there is no endogenous attrition is inherently unobservable, so we cannot directly provide evidence that it does not happen in our case. For this reason, in the results section we use bounds to explore whether our conclusions are sensitive to the possibility of endogenous attrition. In general, we find that even allowing for endogenous attrition term limits do not seem to lead to lower productivity.

2.6 Last Term Effects in the Arkansas Senate

We now study our main question of interest, whether term-limited state senators engage in lower levels of legislative activity than their reelection-eligible counterparts. To do so, we examine four dependent variables at the individual level: the number of bills introduced, the number of bills passed, the abstention rate on roll-call votes and the number of resolu-

[14]In the 1990s cohort, 8 senators assigned 2-year terms and 4 senators assigned 4-year terms drop out of the sample before 1997 (null hypothesis that true probability of success is 0.5 in 12 Bernoulli trials is not rejected, p-value 0.3877), and in the 2000s cohort 1 senator assigned a 2-year term and 2 senators assigned 4-year terms drop out of the sample before 2007.

tions filed, which we use as a proxy for constituency service. While we would prefer more conventional measures of constituency service (e.g. number of district staff, trips back to the district), they are not readily available. Instead, we use data on the resolutions that state senators file during the legislative session. As is typically the case with constituency service, these resolutions are devoid of ideological content; examples include recognizing the achievements of a citizen within their district or congratulating a local high school for its athletic accomplishments. We thus use the number of resolutions as an imperfect proxy for constituency service.

Theoretical Expectations

The framework developed above allows us to state specific expectations for the effects of term limits. In Squire's (2007) index of legislative professionalism, Arkansas ranked 39th in 1996 and 41st in 2003, and in the National Conference of State Legislature's (NCSL) Red-White-Blue trifurcation, Arkansas is considered a "White," or hybrid, legislature based on its intermediate-sized staff and salary, as legislators do not earn enough to make a living without having other sources of income (National Conference of State Legislatures, 2009). The General Assembly holds its regular session in odd-numbered years, meeting for approximately sixty days, and holds what are variously known as fiscal sessions or extraordinary sessions in even-numbered years.

During the 1980s, before term limits were adopted, legislative turnover in Arkansas was low: approximately half of the house seats and two-thirds of the senate seats were occupied by veteran legislators during this decade (Sarbaugh-Thompson, 2010, Table 1). Given this low turnover before term limits and stringent limits on length of service that followed, the effects of term limits on the Arkansas Senate should be higher than in most other states, where the institutional change induced by more lenient term limit policies did not represent such a drastic change (Sarbaugh-Thompson, 2010, p. 202). In addition, applying the framework discussed above, the relatively low professionalization of the Arkansas General Assembly is

likely to exacerbate these last-term effects even more.

On the other hand, the need to secure future employment is likely not a major factor in this setting. In Arkansas, English and Weberg (2007, 148) note, "the part-time nature of the Assembly provide[s] members with ample opportunity to earn a living in their primary vocation while also serving their constituencies as lawmakers." Occasionally, these legislatures convene for special shorter sessions in even-numbered election years, which Arkansas staff describe as "uneventful" and "pro forma" affairs that have historically involved simply rubber-stamping budgets.[15] This means that the opportunity cost of seeking future employment is not likely to induce additional changes in legislative behavior.

In sum, although the need for future employment is not likely to be a major factor in inducing last-term effects, the initial low levels of legislative turnover, the stringent term limit restrictions adopted in the early 1990s, and its low level of professionalization, make the Arkansas General Assembly an environment where the effects of removing reelection incentives should be detected.

Results

Our analysis of the outcome variables mirrors the analysis of covariates reported above. We start by testing the null hypothesis that the effect of term limits is zero for every senator with randomization inference, using the difference-in-means between term-limited senators (those randomly assigned four-year lots) and non-term-limited senators (those randomly assigned two-year lots) as the test-statistic. We also provide randomization-based confidence intervals for the term limits effect, assuming that this effect is constant.[16] As before, we pool observations across the two cohorts to maximize the number of observations. We report separate analyses for the 1997 (81st) and 2007 (86th) sessions in the Appendix, which lead

[15] Personal communication with Arkansas senate staff.

[16] We calculate this confidence interval by inverting a hypothesis test in a constant treatment effect model. That is, letting Y_{iTL} denote the potential outcome of senator i under term limits and Y_{iNTL} denote the potential outcome of senator i in the absence of term limits, we constructed a 95% confidence interval by adopting the model that $Y_{iTL} = Y_{iNTL} + \tau$, testing the null hypothesis that $\tau = \tau_0$ for all possible values of τ_0, and keeping the hypotheses that we failed to reject at 5% level.

to comparable results.

Table 2.4 presents results for our four dependent variables, and Figures 2.3-2.6 display the entire distributions using box-plots. We start by describing bill outcomes. The average number of bills introduced is remarkably similar between the two groups, with term-limited senators introducing about 1.5 more bills than the non-term limited senators. Using the difference-in-means as the test-statistic, we are far from rejecting the sharp null hypothesis of no effect (p-value 0.65), with a 95% confidence interval that ranges from -4.77 to 7.39 bills. A similar result is observed regarding the number of bills passed, with term-limited senators again marginally outperforming their non-term-limited colleagues, passing approximately 1 more bill during the session. Again, though, we are unable to reject the null that term limits have no effect, and the confidence interval is narrower than in the case of bill introductions.

Turning to our measure of constituency service, the third row of Table 2.4 shows that term-limited senators average a little over two resolutions while non-term-limited senators file just over 1.5 resolutions during the legislative session. This difference is small and, again, we fail to reject the null hypothesis of no effect. Moreover, the 95% confidence interval covers mostly positive effects. Figures 2.3-2.5 show that, looking at the entire distributions, there is no evidence that the term-limited group produces less legislative output according to these bill measures.

An additional measure of participatory shirking involves abstention rates on roll-call votes. The term-limited senators abstain on approximately one percent more roll-calls than non-term-limited senators, leading to yet another failure to reject the sharp null hypothesis. The confidence interval is not symmetric around zero, ranging from a small negative effect of -0.65 to a larger positive effect of 3.35. This disparity in the mean abstention rates between groups, which translates into a confidence interval that is shifted to the right of zero, is driven entirely by one senator in the treatment group, Steve Faris, who missed approximately 31 percent of roll-call votes during the 2007 session. As the fourth and fifth rows of Table 2.4 and Figure 2.6 indicate, when we drop this senator from the analysis, the mean differences in

Table 2.4: Effects of Term Limits on Legislative Behavior in Arkansas Senate, pooling 81st (1997-1998) and 86th (2007-2008) Legislative Sessions

	Means			Test of no effect	95% CI
	TL	NTL	Difference	p-value	(constant effect)
Bills introduced	22.88	21.48	1.41	0.65	[-4.77,7.39]
Bills passed	14.58	13.35	1.23	0.58	[-2.91,5.35]
Resolutions	2.23	1.57	0.67	0.22	[-0.36,1.7]
Abstention rate	2.36	1.14	1.22	0.49	[-0.65,3.35]
Abstention rate w/o Faris	1.21	1.14	0.07	0.85	[-0.68,0.83]
Sample size:	23 (TL group) and 26 (NTL group)				

Note: 'TL' refers to term-limited senators (assigned 4-year lot in 1992 or 2002), and 'NTL' refers to non-term-limited senators (assigned 2-year lot in 1992 or 2002). The test of no effect reports randomization-based p-values corresponding to the sharp null hypothesis that the treatment has no effect for any unit using the difference-in-means as test-statistic. Confidence interval calculated by inverting randomization-based hypothesis tests in a constant treatment effect model. Calculations for abstention rates excluding Steve Faris (in the TL group) use a total sample size of 48.

abstentions between both groups vanish almost entirely and the confidence interval becomes much shorter and approximately symmetric about zero.

Figure 2.3: Term limit effects on bill introductions in Arkansas Senate—81st (1997-1998) and 86th (2007-2008) Legislative Sessions

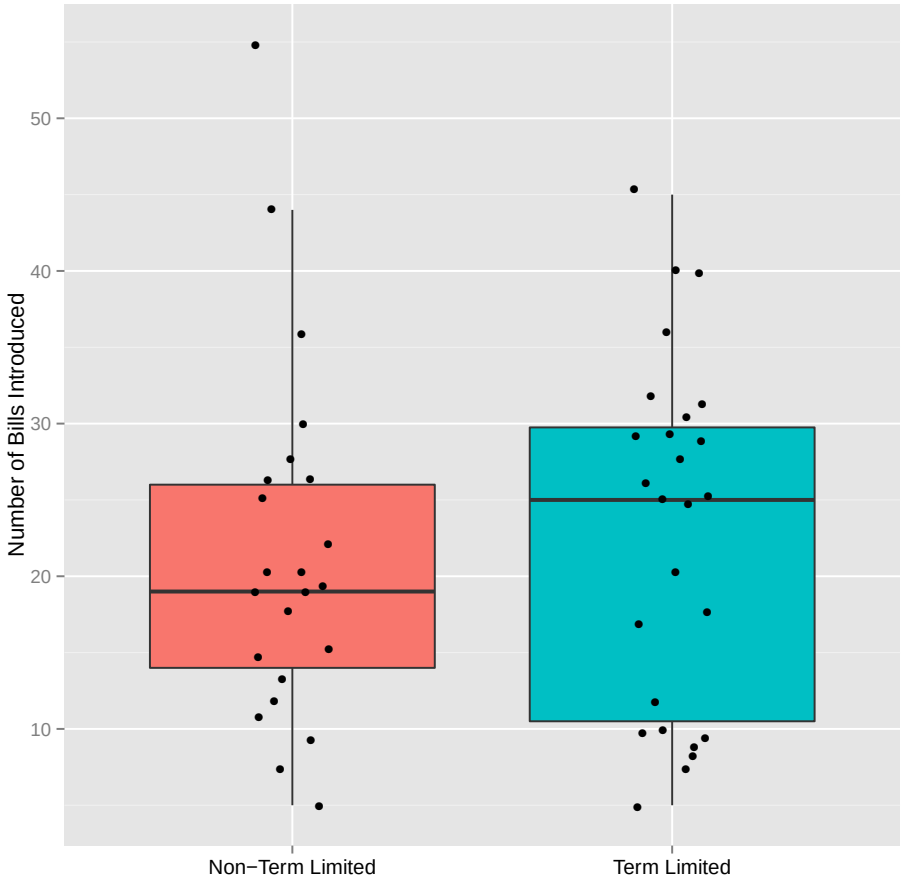


Figure 2.4: Term limit effects on passage in Arkansas Senate—81st (1997-1998) and 86th (2007-2008) Legislative Sessions

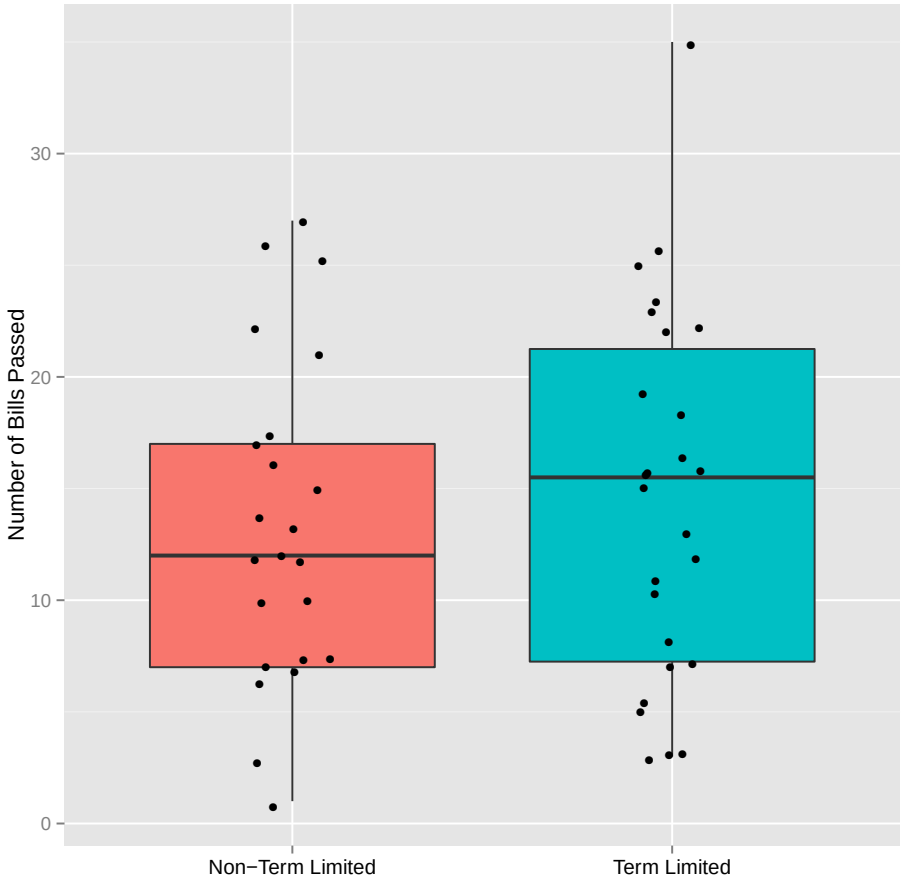
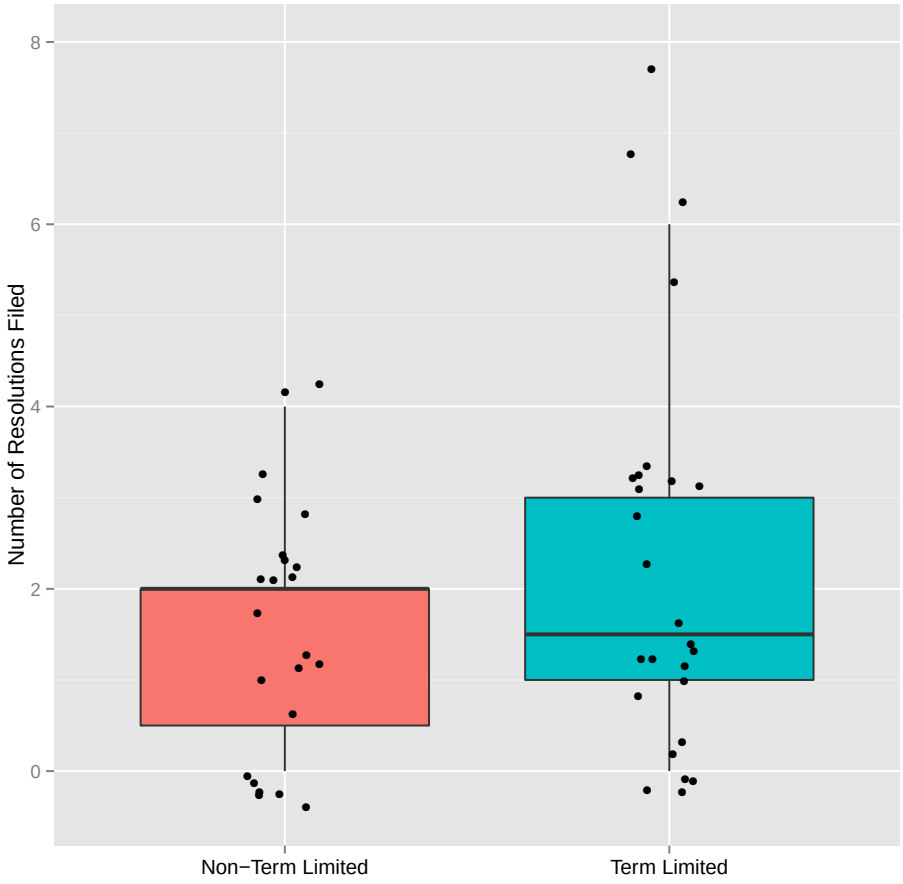


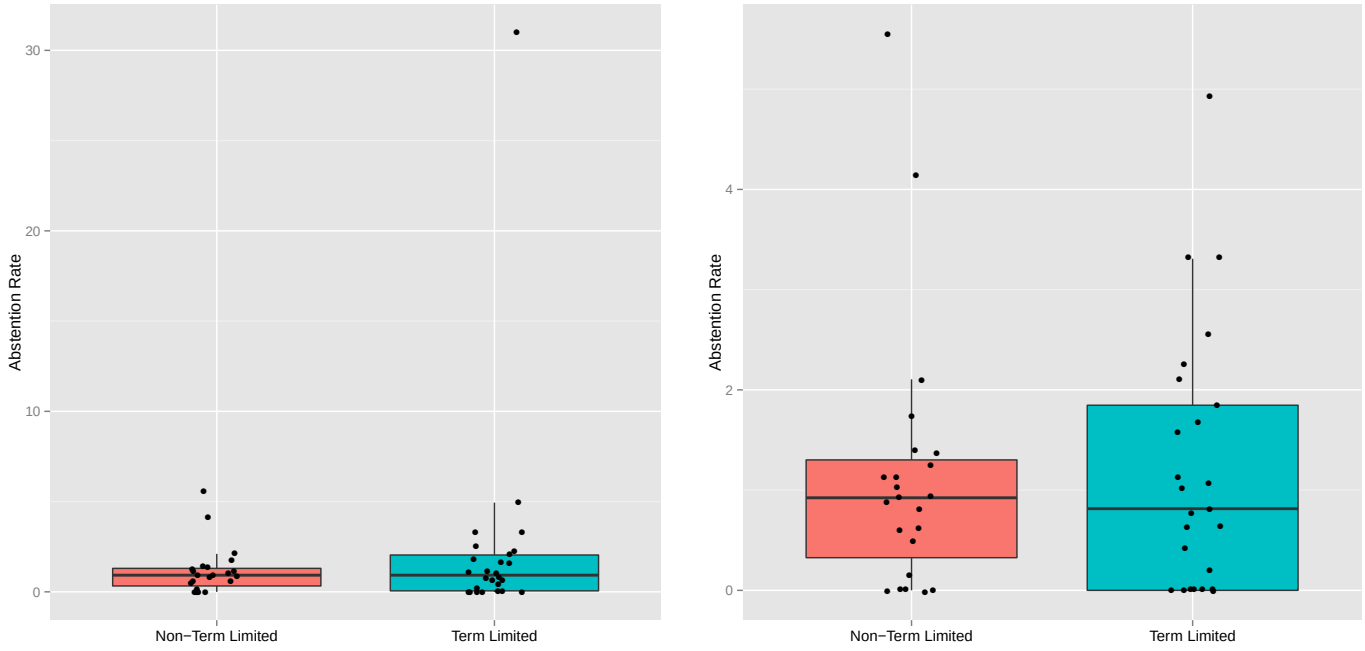
Figure 2.5: Term limit effects on symbolic bills in Arkansas Senate—81st (1997-1998) and 86th (2007-2008) Legislative Sessions



In sum, our randomization-based inferences suggest that there is no evidence of term limits effects on four measures of legislative output and participation. For three of our outcomes, the mean point estimates indicate that term-limited senators are more active, not less, than their non-term-limited counterparts, and the box-plots of the entire distributions equally show that there is no evidence of shirking. However, failing to reject the null hypothesis of no effect does not necessarily mean that we can be confident in asserting that the outcomes in the two groups are equivalent. This is true in every application, and it is a more pressing concern in our case due to the low sample size, which affects our ability to detect true differences.

We therefore examine the impact of term limits using tests of equivalence, as we did for

Figure 2.6: Term limit effects on abstention rates in Arkansas Senate—81st (1997-1998) and 86th (2007-2008) Legislative Sessions



the covariate balance tests. As explained above, in these tests, our null hypothesis is that the legislative output in the term-limited group is sufficiently different from the legislative output in the reelection-eligible group. The first two columns of Table 2.5 report how large of a disparity in legislative output, in either a positive or negative direction, would lead to a rejection of the null hypothesis of nonequivalence. Based on the absolute values in the first column, we can say with 95% confidence that the effect of term limits is at most about 7 bill introductions, 5 bills passed, a 3 percent abstention rate and 1.5 resolutions. In other words, these are the smallest differences between the groups that we can reject with a 5% level test.

Since many previous studies have found that term limits lead to lower effort or output—sometimes referred to as legislative shirking—we are particularly interested in whether we can rule out large *negative* term limits effects. To explore this issue, we test the hypothesis the term-limited and non-term-limited groups differ by a negative amount, that is, we test the

null hypothesis of shirking $H_0^{\delta,S} : \mu_{TL} - \mu_{NTL} < -\delta$, for a positive δ. Note that $\mu_{TL} - \mu_{NTL} < 0$ implies shirking only for the bill and resolution outcomes, but not for our abstention rates, since a *positive* effect of term limits on abstentions is a shirking effect. For this reason, when we report this one-sided test for abstention outcomes, we transform the outcome to be one minus the abstention rate, so that $\mu_{TL} - \mu_{NTL} < 0$ means that term-limited senators are voting less than their non-term-limited counterparts.

The results, which we report in columns 3 and 4 of Table 2.5, suggest that we can rule out large negative effects of term limits on legislative output and participation, as we can assert with 95% confidence that non-term-limited senators, at most, introduce three more bills, see two more bills passed, and file one-quarter more resolutions than their term-limited counterparts, effects that range from 0.2 to 0.4 standard deviations. In terms of abstentions, when Steve Faris is excluded, we can assert that the average rate of non-absent votes among non-term-limited senators is at most 0.74 percentage points higher than among term-limited senators, an effect of 0.8 standard deviations but small in absolute terms.

In sum, our experimental results from Arkansas fail to reject the null hypothesis that term limits have no effect on the measures of legislative output and participation we report, and allow us to rule out with 95% confidence that term limits have large or even moderate negative effects.

Robustness Check: Bounds

The previous section treated attrition as random. In this section, we explore whether our results survive patterns of retirement or defeat that are correlated with the initial assignment of term length. To address attrition in our experimental samples, we estimate upper and lower bounds on the average treatment effect following Manski (2003). In calculating the upper bound on the average treatment effect, we set the outcome values of those senators initially assigned to the treatment group (i.e., 4-year term length group) who were not present in the legislature during the 1997 and 2007 sessions equal to the 75th percentile of

Table 2.5: Tests of equivalence and negative effects (shirking), pooling 81st (1997-1998) and 86th (2007-2008) Legislative Sessions

	Test of equivalence at 95% $H_0^\delta : \lvert\mu_{TL} - \mu_{NTL}\rvert > \delta$		Test of shirking at 95% $H_0^{\delta,S} : \mu_{TL} - \mu_{NTL} < -\delta$	
	Min δ for which H_0^δ is rejected		Min δ for which H_0^S is rejected	
	Raw units	SD units	Raw units	SD units
Bills introduced	7.4	0.9	3.29	0.4
Bills passed	5.52	1	2.21	0.4
Abstention rate[a]	3.36	1.1	3.36	1.1
Abstention rate[a] w/o Faris	0.83	0.9	0.74	0.8
Resolutions	1.66	1.3	0.26	0.2
Sample size:	23 (TL group) and 26 (NTL group)			

Notes: 'TL' refers to term-limited senators (assigned 4-year lot in 1992 or 2002), and 'NTL' refers to non-term-limited senators (assigned 2-year lot in 1992 or 2002); μ_{TL} refers to the mean outcome among TL senators and μ_{NTL} refers to the mean outcome among NTL senators. Tests of the hypotheses H_0^δ and $H_0^{\delta,S}$ are performed using randomization inference, assuming a constant treatment effect model; SD units are expressed in terms of the pooled standard deviation across all senators. Calculations for abstention rates excluding Steve Faris (in the TL group) use a total sample size of 48. [a]In the last two columns where the test of $H_0^{\delta,S} : \mu_{TL} - \mu_{NTL} < -\delta$ is reported, the outcome used is one minus the abstention rate.

each of our four measures of legislative behavior in our pooled complier sample, and the missing outcome values of those senators initially assigned to the control group (i.e., 2-year term length group) equal to the 25th percentile value. To calculate the lower bound on the average treatment effect, we do the opposite, setting missing outcomes in the treatment group at the 25th percentile value of the observed outcomes, and missing outcomes in the control group at the 75th percentile value.

We believe this to be a plausible scenario to test the robustness of the results presented above. The bounds calculated under this scenario essentially recompute the average treatment effect assuming that the pattern of attrition is severely correlated with the initial term length assignment. Our lower bound assumes that all missing outcomes in the treatment group, if observed, would have been low (i.e., equal to the 25th percentile of the observed outcomes) while all the missing outcomes in the control group would have been high (i.e., equal to the 75th percentile of the observed outcomes). Analogously, our upper bound as-

sumes that all missing outcomes in the treatment group, if observed, would have been high (i.e., equal to the 75th percentile of the observed sample) while all the missing outcomes in the control group would have been low (i.e., equal to the 25th percentile of the observed sample).

Table 2.6 reports the bounds on the average treatment effect of term limits on legislative behavior across the four outcomes. The columns 'ATE Lower Bound' and 'ATE Upper Bound' report, respectively, the estimated lower and upper bounds of the difference-in-means between the term-limited and non-term-limited groups. Since our intention is to establish if the no-shirking results reported in the previous section are robust, we focus on the lower bound for bill and resolution outcomes and on the upper bound for abstention rates. The results in Table 2.6 show that for bill introductions, bill passage, resolutions and abstention rates, even a severely endogenous pattern of attrition would result in small shirking or negative term-limits effects. For example, for the number of bills introduced, assuming that all missing term-limited senators would have introduced just 12 bills (25th percentile) while all missing non-term-limited senators would have introduced 29 bills (75th percentile) would result in a lower bound for the difference-in-means of just -2.75 bills, showing that even under severely endogenous attrition we could rule out large last-term effects on this outcome.

A similar pattern is observed for bills passed, resolutions and abstention rates. The lower bound on the last-term effects on bill passage is just -1.78 bills, and this is assuming that missing term-limited senators would have passed just 7 bills while missing non-term-limited senators would have passed 19. The lower bound on resolutions is positive at 0.03, ruling out shirking. And the upper bound on last-term effects for abstention rates excluding Steve Faris is 0.43%, less than half of a percentage point, assuming that the missing abstention rates among term-limited senators would have been 1.68% while the abstention rate among missing non-term-limited senators would have been 0.16%, about ten times smaller—as shown, including Senator Faris increases this bound to 1.37. The 95% confidence intervals on the estimated bounds interval are naturally consistent with larger negative effects, which is ex-

Table 2.6: 75th and 25th Percentile Bounds on Last-Term Effects in Arkansas Senate, pooling 81st (1997-1998) and 86th (2007-2008) Legislative Sessions

Dependent Variable	ATE Lower Bound	ATE Upper Bound	95% CI
Bills introduced	-2.75	5.22	[-7.14, 9.49]
Bills passed	-1.78	3.84	[-4.82 ,6.75]
Resolutions	0.03	0.97	[-0.64, 1.65]
Abstention rate	0.65	1.37	[-0.61, 3.18]
Abstention rate w/o Steve Faris	-0.27	0.43	[-0.76, 0.92]
Sample size:	32 (TL group) and 32 (NTL group)		

Note: Columns labeled 'ATE Lower Bound' and 'ATE Upper Bound' report, respectively, the estimated lower and upper bounds of the difference-in-means between the term-limited (TL) and non-term-limited senator (NTL) groups. Upper bounds sets missing treated outcomes to 75th percentile of observed treated outcome and missing control outcomes to 25th percentile of observed control outcomes. Lower bound is analogous, using 25th percentile for missing treated and 75th percentile for missing control. Confidence intervals calculated with bootstrapping. Calculations for abstention rates excluding Steve Faris (in the TL group) use a total sample size of 63.

pected given the variability that stems from our low sample size. In sum, these results show that when it comes to abstention rates, bill introductions, bill passage and resolutions, our finding that there are no last-term effects on participatory shirking appears robust to endogenous attrition.

2.7 Conclusion

We have examined how the adoption of term limits affects legislative behavior, specifically the extent to which term-limited legislators produce less legislation and abstain at higher rates than their reelection-eligible counterparts. While other scholars have studied the effects of term limits in the U.S. states, ours is the first to experimentally examine the question. Our experimental design overcomes several of the static and dynamic methodological challenges that the nonrandom assignment of legislators to term-limited status has presented previous scholars. Leveraging two experiments in the Arkansas Senate that lead to the random assignment of term-limited status, we find no evidence that legislators slack off when the electoral connection is severed, placing our findings in contrast to previous studies of term

limits in the American states.

Although our results are necessarily limited in scope because they cover only one state, the fact that we fail to see even small differences in the participatory measures we examine casts some doubt on whether, in the United States, the essential role of elections is to sanction representatives to facilitate accountability. Several authors, (e.g., Fearon, 1999; Mansbridge, 2009), have advocated for a selection model of representation, where elections are seen primarily as mechanisms to select representatives that are self-motivated to act in the best interest of voters even in the absence of monitoring and sanctioning. Our results are consistent with this view, and suggest that an often overlooked cost of adopting term limits is to reduce the pool of politicians that can act as representatives.

Our results also have implications for public policy. The decision of citizens across nearly half of the U.S. states to limit the number of terms their representatives could spend in office promised to deliver sweeping changes in both the operation of state legislatures and in the behavior of state legislators. Critics of the initiative process, and of term limits in particular, argued that legislators would have carte blanche to act irresponsibly. Those fears of voters magnifying agency problems have not come to fruition, at least not in the Arkansas Senate.

Much remains to be learned about the dynamics of term limits in particular and the dynamics of last-term effects in general, and additional empirical work is needed to ensure that our conclusions hold for non-participatory outcomes and are generalizable beyond Arkansas. However, we believe that our findings may carry some implications for other states. As mentioned above, Arkansas' lifetime term limits are among the most stringent in the country, suggesting that the change in the incentives facing legislators as they begin to serve their last term is likely larger in Arkansas than in most other states. If even in Arkansas term limits show no effect on legislative participatory outcomes, it is reasonable to expect that similar null effects would be found in other states where the removal of electoral accountability is neither so complete nor severe.

On the other hand, the scarce legislative resources associated with the low level of pro-

fessionalization of the Arkansas Legislature might partly mediate the observed effects, in which case our results would not be immediately applicable to states with highly professional legislatures, where resources are abundant. But even in this case, our findings might carry implications for the six other low-professionalization or "dead-end" legislatures that currently have term limits, which comprise almost half of the total number of states where term limits are currently in effect.[17]

[17]In addition to Arkansas, the other dead-end legislatures include Arizona, Louisiana, Maine, Montana, Nebraska and South Dakota (Maestas, 2000).

2.8 Appendix

In this appendix, Section 2.8.1 presents additional non-experimental comparisons between term-limited and non-term-limited legislators for the eight states analyzed in Section 2.4 of the main paper. Section 2.8.2 presents balance tests separately for each cohort, and Section 2.8.3 presents results for the main outcomes separately for each cohort. In the tables reported below, 'TL' refers to term-limited senators and 'NTL' refers to non-term-limited senators.

2.8.1 Additional Results for Non-experimental Comparisons in Eight States

Table 2.H.1.: Difference-in-means for Vote Share Between Term-Limited and Non-Term-Limited State Legislators, Excluding Uncontested Races

	Vote Share			
	Mean NTL	Mean TL	p-val	N
States with Consecutive Bans				
Arizona	43.22	44.73	.39	556
Colorado	61.96	63.81	.07	625
Ohio	64.33	66.38	.02	805
S. Dakota	41.77	44.20	.14	655
States with Lifetime Bans				
Arkansas (House)	60.81	64.74	.03	279
California	64.08	66.42	.00	857
Michigan	66.24	67.91	.04	1016
Oregon	62.27	65.60	.04	195

Note: The election data come from the State Legislative Election Returns (1967-2010) ICPSR #34397 dataset. Column labeled 'Mean TL' reports the mean outcome for term-limited senators, and column labeled 'Mean NTL' reports the mean outcome for non-term-limited senators. The columns labeled 'p-val' report p-values from two-tailed t-tests of the null hypothesis that means are equal. Column labeled N reports the number of observations used in each row.

Table 2.H.2.: Difference-in-means for Vote Share Between Term-Limited and Non-Term-Limited State Legislators, Excluding Freshman Legislators and Uncontested Races

	Vote Share			
	NTL	TL	p-val	N
States with Consecutive Bans				
Arizona	46.48	44.73	.35	380
Colorado (House)	64.32	65.35	.48	296
Ohio (House)	66.50	66.80	.76	468
S. Dakota	45.28	44.20	.54	454
States with Lifetime Bans				
Arkansas (House)	63.83	64.74	.68	103
California (House)	66.07	66.93	.33	428
Michigan (House)	69.63	68.98	.54	540
Oregon (House)	64.83	64.42	.85	90

Note: The election data come from the State Legislative Election Returns (1967-2010) ICPSR #34397 dataset. Column labeled 'Mean TL' reports the mean outcome for term-limited senators, and column labeled 'Mean NTL' reports the mean outcome for non-term-limited senators. The columns labeled 'p-val' report p-values from two-tailed t-tests of the null hypothesis that means are equal. Column labeled N reports the number of observations used in each row.

2.8.2 Additional Balance Tests for Arkansas Experimental Design

In this section, we present the same information in Table 3 in the main body of the paper, disaggregated by cohort. Table 2.H.3. presents the results from balance tests for the group of senators who drew lots in October 1993. Table 2.H.4. presents the results from balance tests for the group of senators who drew lots in December 2002. Finally, Table 2.H.5. presents the same covariate balance tests in Table 3 in the main paper, but now uses t-tests of the null hypothesis that means in both groups are equal instead of randomization-inference tests of the sharp null hypothesis, and it also uses t-tests to perform tests of equivalence.

Table 2.H.3.: Covariate Balance Between Term-Limited and Non-Term-Limited Arkansas Senators, 1992 cohort

	Means			Test of no effect	Test of $H_0^\delta : \lvert\mu_{TL} - \mu_{NTL}\rvert > \delta$	
	TL	NTL	Difference	p-value	Min δ for which H_0^δ is rejected Raw units	SD units
Vote Share	97.29	90.78	6.51	0.27	15.87	1.5
Married	0.94	0.83	0.11	0.6	0.29	1.3
Male	0.94	1	-0.06	0.48	0.15	1.2
Democrat	0.94	0.83	0.11	0.59	0.29	1.3
Black	0.12	0.06	0.06	0.6	0.22	1.1
Attorney	0.47	0.5	-0.03	1	0.36	1
Age	50.71	50	0.71	0.85	6.96	0.9
Sample size:	17 (TL group) and 18 (NTL group)					

Note: Column labeled 'TL' refers to term-limited senators (assigned 4-year lot in 1992), and column labeled 'NTL' refers to non-term-limited senators (assigned 2-year lot in 1992). The test of no effect reports randomization-based p-values corresponding to the sharp null hypothesis that the treatment has no effect for any unit using the difference-in-means as test-statistic. In tests of the hypothesis H_0^δ reported in the last two columns, μ_{TL} refers to the mean outcome among term-limited senators and μ_{NTL} refers to the mean outcome among non-term-limited senators; these tests are also randomization-based, assuming a constant treatment effect model. SD units are expressed in terms of the pooled standard deviation across all senators.

Table 2.H.4.: Covariate Balance Between Term-Limited and Non-Term-Limited Arkansas Senators, 2002 cohort

	Means			Test of no effect	Test of $H_0^\delta : \lvert\mu_{TL} - \mu_{NTL}\rvert > \delta$	
	TL	NTL	Difference	p-value	Min δ for which H_0^δ is rejected Raw units	SD units
Vote Share	80.17	81.69	-1.52	0.85	15.58	1
Married	0.8	0.93	-0.13	0.6	0.34	1.4
Male	0.93	0.57	0.36	0.04	0.63	2.2
Democrat	0.87	0.64	0.22	0.22	0.52	1.7
Black	0.13	0.07	0.06	1	0.29	1.3
Attorney	0	0.14	-0.14	0.22	0.29	1.6
Age	50.6	56.36	-5.76	0.17	13.02	1.7
Sample size:	15 (TL group) and 14 (NTL group)					

Note: Column labeled 'TL' refers to term-limited senators (assigned 4-year lot in 2002), and column labeled 'NTL' refers to non-term-limited senators (assigned 2-year lot in 2002). The test of no effect reports randomization-based p-values corresponding to the sharp null hypothesis that the treatment has no effect for any unit using the difference-in-means as test-statistic. In tests of the hypothesis H_0^δ reported in the last two columns, μ_{TL} refers to the mean outcome among term-limited senators and μ_{NTL} refers to the mean outcome among non-term-limited senators; these tests are also randomization-based, assuming a constant treatment effect model. SD units are expressed in terms of the pooled standard deviation across all senators.

Table 2.H.5.: Covariate Balance Between Term-Limited and Non-Term-Limited Arkansas Senators, 1997 and 2007 cohorts pooled

	Means			t-test of equal means	Test of $H_0^\delta : \lvert\mu_{TL} - \mu_{NTL}\rvert > \delta$	
	TL	NTL	Difference	p-value	Min δ for which H_0^δ is rejected Raw units	SD units
Vote Share	89.27	86.8	2.46	0.61	11.03	0.8
Married	0.88	0.88	0	1	0.14	0.6
Male	0.94	0.81	0.12	0.14	0.28	1.2
Democrat	0.91	0.75	0.16	0.1	0.32	1.2
Black	0.12	0.06	0.06	0.4	0.19	0.9
Attorney	0.25	0.34	-0.09	0.42	0.29	0.9
Age	50.66	52.78	-2.12	0.44	6.98	0.9
Sample size:	32 (TL group) and 32 (NTL group)					

Note: Column labeled 'TL' refers to term-limited senators (assigned 4-year lot in 2002), and column labeled 'NTL' refers to non-term-limited senators (assigned 2-year lot in 2002). t-test of equality of means is two-sided. In tests of the hypothesis H_0^δ reported in the last two columns, μ_{TL} refers to the mean outcome among term-limited senators and μ_{NTL} refers to the mean outcome among non-term-limited senators; these tests are also based on t-tests. SD units are expressed in terms of the pooled standard deviation across all senators.

2.8.3 Additional Treatment Effect Results for Arkansas Experimental Design

We now present, in Tables 2.H.6. and 2.H.7., the effects of term limits separately for the 1992 and the 2002 cohorts, respectively. The outcomes for the 1992 cohort occur during the 1997 legislative session, and the outcomes of the 2002 cohort occur during the 2007 legislative session. In the main body of the paper we pool observations to address potential concerns about sample size; here we show that conducting separate analyses for the 1997 and 2007 sessions returns comparable results.

Table 2.H.6.: Effects of Term Limits on Legislative Behavior in Arkansas Senate, 1992 cohort

	Means			Test of no effect	95% CI
	TL	NTL	Difference	p-value	(constant effect)
Bills introduced	18.38	14.4	3.98	0.35	[-4.23,12.23]
Bills passed	12.08	8.3	3.78	0.28	[-2.58,10.38]
Abstention rate	1.1	1.02	0.08	0.91	[-1.25,1.36]
Resolutions	2.08	0.9	1.18	0.09	[-0.16,2.51]
Sample size:	13 (TL group) and 10 (NTL group)				

Note: 'TL' refers to term-limited senators (assigned 4-year lot in 1992), and 'NTL' refers to non-term-limited senators (assigned 2-year lot in 1992). The test of no effect reports randomization-based p-values corresponding to the sharp null hypothesis that the treatment has no effect for any unit using the difference-in-means as test-statistic. Confidence interval calculated by inverting randomization-based hypothesis tests in a constant treatment effect model.

Table 2.H.7.: Effects of Term Limits on Legislative Behavior in Arkansas Senate, 2002 cohort

	Means			Test of no effect	95% CI
	TL	NTL	Difference	p-value	(constant effect)
Bills introduced	27.38	26.92	0.46	0.94	[-8.55,9.4]
Bills passed	17.08	17.23	-0.15	0.98	[-5.45,5.2]
Abstention rate	3.62	1.23	2.39	0.39	[-0.64,6.63]
Abstention rate w/o Faris	1.34	1.23	0.11	0.78	[-0.76,0.99]
Resolutions	2.38	2.08	0.31	0.78	[-1.30,2.00]
Sample size:	13 (TL group) and 13 (NTL group)				

Note: 'TL' refers to term-limited senators (assigned 4-year lot in 2002), and 'NTL' refers to non-term-limited senators (assigned 2-year lot in 2002). The test of no effect reports randomization-based p-values corresponding to the sharp null hypothesis that the treatment has no effect for any unit using the difference-in-means as test-statistic. Confidence interval calculated by inverting randomization-based hypothesis tests in a constant treatment effect model. Calculations for abstention rates excluding Steve Faris (in the TL group) use a total sample size of 25.

CHAPTER III

Fashionably LATE: Local Average Treatment Effects of Term-Limited Legislative Behavior

Abstract: How do term limits affect legislative behavior? When ambitious legislators are proscribed from service, they must consider how their current decisions in office affect their future goals. Using a panel of bill activity among members of the Arkansas state house from 1999-2011, I show that term-limited legislators are significantly more active on proactive forms of participation (e.g. bill introductions) over which they have discretion but significantly less active on reactive forms of participation (e.g. roll-call votes) over which they have limited discretion, revealing the nuanced ways term limits affect state legislators' incentives.

3.1 Introduction

> The central assumption of ambition theory is that a politician's behavior is a response to his office goals. Or, to put it another way, the politician as office seeker engages in political acts and makes decisions appropriate to gaining office (Schlesinger, 1966, 6)

Since its passage across nearly half of the states in the 1990s, term limits have attracted considerable attention from state legislative scholars. Much of the existing literature has focused on aggregate effects, examining changes in electoral competition (Schraufnagel and Halperin, 2006; Masket and Lewis, 2007), inter-institutional influence (Berman, 2007;

Moncrief and Thompson, 2001), demographic composition (Carey et al., 2006; Sarbaugh-Thompson et al., 2004) and the partisan makeup of legislatures (Hall, 2014). Less attention, however, has been paid to individual-level effects such as how term limits affect legislative behavior, particularly the behavioral choices legislators make as they seek higher office.

A small body of research explores how term limits alter the opportunity structure in state legislatures and the corresponding calculus strategic officeholders face when considering a bid for higher office. Survey evidence, for instance, points to legislators in term-limited states reporting higher levels of progressive ambition and a greater likelihood of running for higher office than legislators in states without term limits (Herrick and Thomas, 2005). And observational data confirms survey responses, finding that termed-out legislators are more likely to run for U.S. House seats than legislators in non-term-limited states (Powell, 2000).

To capture a more holistic effect of term limits, it is important to look not only at election day, but also at the months and years preceding it. In that vein, I build on the aforementioned existing literature by exploring the participatory decisions term-limited legislators make as they seek higher office. Channeling Schlesinger (1966), in a legislature where ambitious legislators are proscribed from service, they must consider how their current decisions in office affect their future goals. Will casting an unpopular vote alienate an attentive public? Can a productive legislative portfolio garner local media attention and concomitantly boost name recognition?

In this paper I explain how term limits restructure individual ambition and in turn affect the legislative choices state legislators make. I propose a career-concerns theory of legislative behavior that identifies how ambition for higher office among term-limited legislators affects patterns of legislative activity. Since running for higher office involves gaining support from a larger constituency, we should expect term-limited legislators (i.e. those imminently positioned to run for higher office) to be more legislatively active than their non-term-limited colleagues, but not across all types of legislative behavior.

Importantly, I distinguish between proactive and reactive forms of participation (Burden, 2007). In short, I argue that on proactive forms of participation (e.g. bill sponsorships, constituency service) over which legislators generally have more discretion, term-limited legislators are likely to display higher levels of activity than their non-term-limited colleagues, but on reactive forms of participation (e.g. roll-call attendance), term-limited legislators are likely to display lower levels of activity so as to avoid casting a vote on an issue that might hurt them in their pursuit up the political ladder (Lee, 2009). To test the theory, I use a panel of bill activity among members of the Arkansas state house from 1999-2011. The empirical analysis supports the career-concerns theory, thereby revealing the nuanced ways term limits affect state legislators' incentives.

The remainder of the paper proceeds as follows. In the next section, I develop hypotheses regarding how term limits affects legislative behavior. From there, I discuss some of the methodological challenges of studying term-limited legislative behavior, focusing on the endogenous term limits indicator commonly used in the literature. Given that researchers are interested in assessing the reform's effects with observational data, I propose two estimation strategies—one based on matching and one based on instrumental variables regression—both of which restrict comparisons to more homogenous units in order to recover a local average treatment effect (LATE) for a subset of the study population. And the final section concludes.

3.2 Ambition and Legislative Behavior

Schlesinger (1966) describes three types of political ambition: (1) progressive ambition where the goal is to attain higher office, (2) static ambition where the goal is to build a long career in one's current office, and (3) discrete ambition where the goal is to retire after a period of service. Prior to the passage of term limits, many state legislators appeared content with static ambition, securing reelection and building seniority within a chamber. Across myriad states that eventually went on to pass term limits, veteran legislators represented

large majorities within chambers; in a number of cases 75-80% of the seats were occupied by veteran legislators (Sarbaugh-Thompson, 2010).

With the passage of term limits, however, the same state legislatures once populated by veteran legislators with decades of experience soon began to witness high levels of turnover. In 1998, for example, the lower houses of states with term limits experienced legislative turnover that was more than twice as high as states without term limits; during that election cycle, over half of the termed-out legislators ran for other offices (Powell, 2003). Moreover, the average length of service declined precipitously in a short period of time. In California, for example, the average length of service for an assembly member dropped from 8 years in 1988 down to an average of 2.7 years in 1998 (Gordon and Unmack, 2003).

There does not appear to be much support for the notion that term limits suddenly incented a different type of office-seeker to run.[1] Indeed, in their survey of state legislators, Carey, Niemi and Powell (2000) find that legislators in term-limited states were just as likely as legislators in states without term limits to see themselves as career politicians. The difference, of course, is the former operate in an institutional setting where they can no longer build long careers in a single chamber but instead need to engage in chamber-climbing or some other form of office-jumping to prolong their political careers.

Upending the opportunity structure in state legislatures is bound to affect legislators' incentives, and it appears to have done so through a corresponding shift from static to more progressive forms of ambition. Importantly, those legislators who harbor ambitions to prolong their political career must consider how their current decisions affect their future chances of winning seats to higher offices. And since running for higher office tends to involve gaining support from a larger constituency, it is reasonable to expect progressively ambitious term-limited (PATL) legislators, who are imminently faced with the decision to run for higher office, to be more active than their non-term-limited colleagues, but not across

[1]Survey evidence does, however, point to a difference between lower-house members in term-limited and non-term-limited states in their expressed intention to run for the upper house with legislators in the former being far more likely to say they intend to run for election to the upper house.

all types of legislative behavior. Accordingly, I develop a career-concerns theory of legislative behavior that distinguishes between proactive and reactive forms of legislative participation.

Proactive activities, such as bill introductions and constituency service, are not vigilantly policed by party leaders, allowing legislators to choose an issue portfolio that they find electorally valuable (Hall, 1996). In laying the groundwork for elective office beyond their current chamber, PATL legislators stand to gain in myriad ways from higher levels of proactive output, both among voters and among attentive publics. This greater discretion in bill sponsorships and casework enables legislators to cultivate a base of support by claiming credit for initiating policy changes and for remaining connected to district ongoings (Mayhew, 1974; Grimmer, 2013). Introducing a bill or resolution that spotlights the accomplishments of a district resident or high school can garner local media attention and help build name recognition, both of which aid in the legislator's pursuit of higher office.

Maintaining high levels of activity not only curries favor with voters, but also enables PATL legislators to appeal to attentive publics (Arnold, 1990). By guiding bills through the chamber and demonstrating an affinity for policymaking niceties, PATL legislators can convince donors to support their bid for higher office, avoiding any impression that they are slowing down and may not be worth the resource investment (Diermeier, Keane and Merlo, 2005). Based on the aforementioned ways in which legislative effort can result in favorable reactions from various constituencies, I test the following hypothesis:

> *Hypothesis 1: term-limited legislators will display higher levels of bill activity and attention to constituents than non-term-limited legislators on proactive modes of participation.*

Term limits is also likely to affect reactive forms of participation such as roll-call voting. In contradistinction to proactive participation, roll-call voting is militantly policed by party leaders and interest groups, offering legislators far less discretion. And since legislators generally have limited influence over which issues come up for roll-call votes on the floor, PATL legislators may be more likely to abstain so as to avoid casting a yea or a nay vote on

an issue that might hurt them in their pursuit up the political ladder. When preparing to make the move from the lower to the upper chamber, for example, abstaining on a significant majority of roll-calls could certainly generate a negative news story, but it is still likely to be less politically damaging than casting a vote that results in a fusillade of attacks ads or that alienates either supporters or a key segment of the electorate (Lee, 2009). This leads to the second hypothesis:

> *Hypothesis 2: term-limited legislators will display lower levels of activity than non-term-limited legislators on reactive modes of participation.*

Before turning to the empirical analysis, I briefly consider how (i) legislative professionalization, (ii) treatment heterogeneity and (iii) preference heterogeneity could affect the relationship between ambition and behavior.

3.2.1 The Role of Professionalization, Treatment Heterogeneity and Preference Heterogeneity

The general conception of progressive ambition holds that state legislators make upward moves in their elective careers, moving from state house to state senate to federal office. But it is entirely plausible to expect the characteristics of state legislatures, the nature of a state's term limits regime and career opportunities outside of government to affect the likelihood legislators will pursue higher office after service at the state level. Beginning with legislative professionalization, high professionalization bodies closely resemble Congress in terms of both institutional resources (e.g. large staff, high salaries, full-time job) and campaign experiences (Squire, 2007; Maestas, 2000; Maestas et al., 2006). In states like California and Michigan, legislating requires a year-round commitment with incumbents spending more time keeping in touch with constituents and responding to casework requests (Samples and McDonald, 2006). The full-time nature of the job, in turn, allows ambitious legislators to see politics as a career. In part-time, low professionalization legislatures, by contrast, lawmakers earn a

comparatively paltry salary and thus return to their primary occupation—from which they derive the preponderance of their income—at the conclusion of a three-month legislative session. Given the similarity of high professionalization legislatures to the U.S. Congress, mounting a bid for a U.S. House seat is likely to constitute a considerably lower barrier for PATL legislators in high professionalization bodies relative to their counterparts in low professionalization bodies.

While this paper's analysis is confined to the lower chamber in the Arkansas General Assembly and thus the intuition described above lies beyond the empirical scope of this paper, the expectation would be that in high professionalization bodies, among those pursuing higher office, one would observe behavioral differences between term-limited and non-term-limited legislators at both the state house and state senate levels. Low professionalization bodies, by contrast, are marked by poor advancement prospects; as a result, state legislators are more likely to return to private life after 14-16 years in office. Based on these features, one would primarily expect to observe behavioral differences between term-limited and non-term-limited legislators at the state house level where PATL office-seekers are still seeking an upper chamber seat. These differences in institutional resources and advancement opportunities between high and low professionalized legislatures could help explain why in Chapter II we see no difference in the legislative behavior of term-limited and non-term-limited Arkansas state senators.

In addition to a legislature's professionalization, another important consideration centers on the wide variation in term limit laws across the states: some states impose a lifetime ban while others impose a more lenient consecutive service ban. While scholars have often pooled legislators across states, there is little reason to believe such treatment heterogeneity would be irrelevant to the outcome under study.

If it is the case that the effect of term limits will depend on the version of treatment to which legislators are subjected, in states with consecutive service bans, we might expect to observe an attenuated relationship between term-limited status and legislative behavior, as

ambitious office-seekers can cycle between chambers, holding off on a run for statewide office until electoral conditions become most auspicious. Ambitious legislators pursuing higher office in states with lifetime service bans, by contrast, do not enjoy the same luxury and instead must continue their ascent up the political ladder or risk an interregnum in their office-holding.

A third element that could affect the relationship between term limits and legislative behavior involves career opportunities outside of government. In an effort to directly engage the existing literature, the previous pages have focused on the types of choices ambitious legislators make when their goal is higher office. But not all state legislators harbor aspirations for higher office (Maestas et al., 2006). Instead, some departing legislators might choose to monetize their time in government. One such way to do so would be to seek out lucrative private sector employment opportunities, such as lobbying, where legislative experience has proved to be valuable (Eggers and Hainmueller, 2009; Vidal, Draca and Fons-Rosen, 2012).[2]

Extant studies of state legislatures suggest that legislators in states with term limits have reason to be especially interested in working as a lobbyist. For example, in drawing on a survey of knowledgeable observers, Mooney (2007) reports that with the passage of term limits, lobbyists wield more influence in the legislative process. Using the 2002 State Legislative Survey, Miller, Nicholson-Crotty and Nicholson-Crotty (2011) find that lifetime service bans, which exist in Arkansas, California and Michigan, are associated with significantly greater interest group power. And so, those term-limited lawmakers with a preference for lucrative opportunities outside of government could use the knowledge of the legislative process they developed, as well as connections they established over the course of three terms in the state house, on behalf of an interest group or firm.

Unfortunately, the empirical strategy in this chapter, which simply counts legislative activities, is ill-equipped to tease out these niceties. A more profitable approach would

[2]There is at least some anecdotal support for the idea that state legislators entertain the idea of a career in lobbying. Arkansas staff describe state legislators returning to the capitol several days a week long after the legislative session's adjournment to speak with lobbyists.

center on collecting data on the career decisions of term-limited legislators following their term-limited session, explicitly modeling the set of options as a multinomial choice, and also examining the issue content of their legislative portfolios in an effort to detect evidence of currying favor with a prospective future employer.

3.3 Empirical Strategy

3.3.1 Data

To gauge the effect of term limits on proactive legislative behavior, I examine three dependent variables at the individual level: bill introductions, bills passed and resolutions filed.[3] And to test the effect of term limits on reactive legislative behavior, I examine abstention rates. While the panel for the first three outcomes—bill introductions, bills passed, and resolutions filed—spans regular legislative sessions from 1999-2011, for abstentions, data is only available for the 2005-2011 sessions.

3.3.2 The Methodological Challenge: Dealing with Noncompliance

Because term-limited legislators are proscribed from service after a set number of terms regardless of their prospects for electoral success, the reform was seen as a convenient setting in which to study legislative behavior. On this point, Mooney (2009, 217), for example, notes that "scholars can safely set up quasi-experiments [...] using state legislative term limits as an exogenous treatment." If true, it allows state politics researchers to simply compare the outcomes of term-limited legislators to the outcomes of reelection-eligible legislators since the the control group (i.e. non-term-limited legislators) would serve as a valid counterfactual for the treatment group (i.e. term-limited legislators). While intuitively appealing, this approach may be misguided if there are systematic differences between the two groups.

Importantly, Mooney's (2009) claim that term limits constitute an exogenous treatment

[3]See Chapter II for an explanation for the use of resolutions as a measure of constituency service.

rests on, inter alia, the tacit assumption that legislators follow similar electoral trajectories. But an often overlooked empirical challenge of studying term limits is that not all legislators are assured the electoral fortune of surviving to their term-limited session. As a result, non-experimental research designs—surveys, case studies, cross-state comparisons etc.—may be confounded by unmeasured differences between term-limited and non-term-limited legislators, resulting in biased estimates of the reform's effects.

While a number of reasons could account for systematic differences between term-limited and non-term-limited legislators, one plausible explanation centers on a straightforward model of electoral selection. As Chapter II explains, the group of term-limited legislators consists entirely of legislators who have survived the maximum number of elections before term limits come into effect, thereby revealing themselves to be high-quality politicians, whereas the group of non-term-limited legislators pools two types of legislators: the "survivor" type who will win the maximum number of elections before term limits come into effect, but also the "departor" type who will be defeated and will depart before term limits become binding.

Accordingly, we can hypothesize that this quality differential will be reflected in the electoral fortunes of the two groups. Specifically, one would expect that at the level of the individual legislator, term-limited legislators will on average enjoy higher levels of electoral success as measured by vote share and contestation rates. To test this hypothesis, I use a simple difference-in-means test to examine the pretreatment covariate distributions of the two groups of legislators.

Table 3.1. reports results from the balance tests. Examining the election cycles from 1996-2010, the results show imbalance in party affiliation with a significantly higher percentage of Democrats in the term-limited group. In addition, there is a substantively large and statistically significant difference in both uncontested rates and vote share in the hypothesized direction. Term-limited representatives see a 10 percentage point higher vote share than non-term-limited representatives and are also significantly less likely to face a

Table 3.1.: Balance Between Term-Limited and Non-Term-Limited Arkansas State Representatives, 1996-2010

	Non-Term-limited	Term-limited	p-value
Democrat	.72	.79	.03
Black	.09	.09	.78
Female	.20	.22	.59
Uncontested Rate	57.67	82.40	.00
Vote Share	83.82	93.92	.00
N	666	232	

Note: The election data come from the State Legislative Election Returns (1967-2010) ICPSR #34397 dataset. Reported p-values are from two-tailed tests.

challenger.

Detecting imbalances in observable characteristics, such as vote share and uncontested rates, is likely symptomatic of a broader problem of imbalance across a host of unmeasured attributes (e.g. work ethic, political skill etc.) that affect treatment status as well as outcomes. Collectively, the findings in Table 3.1. point to a clear quality differential between the two groups; higher quality legislators tend to comfortably win reelection, thereby surviving to their term-limited session, whereas the control group of non-term-limited legislators includes departors who are likely to be of lower average quality than survivors, which could drive the observed differences in electoral success we see.

3.3.3 Alternative Empirical Approaches: Matching and IV

While randomized experimental designs, such as the one used in Chapter II, are best able to provide credible estimates of causal effects, they are often difficult to implement or simply unavailable. The alternative empirical strategies proposed and tested in this section—one based on matching and one based on instrumental variables (IV) regression—are, to be clear, a second-best option relative to randomized experiments, in that they require a stronger set of assumptions.[4] In addition, both estimation strategies are better suited to provide a local

[4]For example, whereas random assignment ensures that treated and control groups will be similar in observable and unobservable covariates, in the case of matching, one must make an unverifiable selection-

average treatment effect, or LATE, for a more homogenous subpopulation of legislators as opposed to recovering an overall average treatment effect (Angrist, Imbens and Rubin, 1996).

To understand why, recall the intuition surrounding electoral selection from the previous section. Relative to survivors, departors in the control group are those politicians who are of lower-than-average quality. As a result, these legislators cannot be induced to serve to their term-limited session, and thus, one cannot identify the average effect of term limits for this group because one never observes their participation in the term-limited treatment. Therefore, the estimates from an IV regression, for example, will apply only to the subpopulation of compliers—that is, those who are haphazardly encouraged to serve through to their term-limited session.

As a first cut to reduce bias from confounds, such as electoral success, I use parametric and nonparametric matching estimators. One of the motivations for matching is to avoid linear extrapolations that accompany regression estimates by restricting comparisons to treated and control units with overlap in covariate distributions. For this particular application, one is able to restrict comparisons to a more homogenous subpopulation of term-limited and non-term-limited legislators who experience comparable levels of electoral success. Examining my first measure of proactive behavior, Table 3.2. reports estimates of the average treatment effect on the treated (ATT) from two different methods: propensity score and nearest neighbor matching. In each case, the estimates recover highly significant differences with term-limited legislators introducing approximately four more bills than non-term-limited legislators.

It is, of course, possible that term-limited legislators, who are most likely to harbor progressive ambition and endeavoring to appeal to broader constituencies, are active during this first stage of the legislative process, but subsequently take their foot off of the proverbial gas pedal after the bill gets filed (Butler and Revesz, 2012). The reasoning is straightforward: seeing a bill through to passage represents a significant investment in time; and because it is the result of chamber-wide action, passing a bill is also a noisier means by which to convey

on-observables assumption, namely that the researcher observes and conditions on all known confounds such that the only reason the treated and control groups differ is that the former received the treatment.

effort to constituents. To test for this possibility, Table 3.3. presents ATT estimates of bills passed. Here, again, the differences are highly significant at the 1 percent level and indicate that term-limited legislators pass roughly 2.5 more bills than their non-term-limited colleagues. As a third measure of proactive participation, Table 3.4. examines resolutions filed and again finds that term-limited legislators are significantly more active. Based on the matching estimates, then, there is consistent support for Hypothesis 1.

Table 3.2.: Effects of Term Limits on Bill Introductions, 1999-2011

	ATT	p-value	95% CI
Propensity Score	4.34	.00	[2.35, 6.33]
	(1.02)		
Nearest Neighbor	4.01	.00	[2.01, 6.01]
	(1.01)		
N	695		

Note: The ATT column reports matching estimates of the effect of term limits on bill introductions using propensity score and nearest neighbor matching. The specifications include all of the pretreatment covariates listed in Table 3.1. Robust standard errors are in parentheses.

Table 3.3.: Effects of Term Limits on Bill Passage, 1999-2011

	ATT	p-value	95% CI
Propensity Score	2.58	.00	[1.33, 3.84]
	(.64)		
Nearest Neighbor	2.54	.00	[1.26, 3.83]
	(.66)		
N	695		

Note: The ATT column reports matching estimates of the effect of term limits on bills passed using propensity score and nearest neighbor matching. The specifications include all of the pretreatment covariates listed in Table 3.1. Robust standard errors are in parentheses.

I now turn to Hypothesis 2, which concerns legislators' reactive behavior. Whereas the previous three outcomes were collected for seven legislative sessions, I am limited to abstention rates from four legislative sessions from 2005-2011. Table 3.5. presents the propensity

Table 3.4.: Effects of Term Limits on Resolutions, 1999-2011

	ATT	p-value	95% CI
Propensity Score	.39	.00	[.11, .67]
	(.14)		
Nearest Neighbor	.41	.00	[.14, .68]
	(.66)		
N	695		

Note: The ATT column reports matching estimates of the effect of term limits on resolutions filed using propensity score and nearest neighbor matching. The specifications include all of the pretreatment covariates listed in Table 3.1. Robust standard errors are in parentheses.

score and nearest neighbor matching estimates. Consistent with expectations, term-limited legislators abstain on approximately 4 percent more roll-call votes than non-term-limited legislators, a difference that is significant at the 1 percent level, thereby providing support for Hypothesis 2.

Table 3.5.: Effects of Term Limits on Abstentions, 2005-2011

	ATT	p-value	95% CI
Propensity Score	4.47	.00	[1.79, 7.15]
	(1.37)		
Nearest Neighbor	4.24	.00	[1.55, 6.93]
	(1.37)		
N	395		

Note: The ATT column reports matching estimates of the effect of term limits on abstention rates using propensity score and nearest neighbor matching. The specifications include all of the pretreatment covariates listed in Table 3.1. Robust standard errors are in parentheses.

To further explore concerns about endogeneity, an alternative approach to address biases from unobserved confounds is to find an instrument for treatment status. With respect to term limits, one would need an instrument that as-if randomly encourages legislators to become term-limited but has no direct effect on legislative participation. One such possibility to satisfy the exclusion restriction is the year a legislator is first elected. The reasoning centers

on the fact that members of the Arkansas House of Representatives are limited to three two-year terms, so the year in which they are first elected will strongly predict when they will be term-limited; less clear, however, is whether the instrument is unrelated to unobserved determinants of bill activity, a potential violation of the exclusion restriction that I consider briefly below.

Tables 3.6., 3.7., 3.8. and 3.9. report estimates from the IV regressions for the four outcome variables: bills introduced, bills passed, resolutions filed and abstention rates. The first column in each table reports the coefficient estimates from a model with random effects, which is useful insofar as providing insight into how time-invariant individual-level characteristics affect bill activity.[5] And the second column reports the coefficient estimates from a model with legislator fixed effects. Both models include year fixed effects to account for any unobserved time-varying factors.

The first thing to note is that our instrument is positively correlated with the endogenous term limits indicator at the 1 percent, satisfying the inclusion restriction. In addition, the instrument generally appears to be strong, yielding an F-statistic above 10 for two of the four behavioral outcomes. For resolutions and abstention rates, however, the F-statistics fall below 10, so the finite sample problems of weak instruments could present a problem in this setting.

Turning to the estimates from the IV regressions, we see that they are slightly larger than the estimates from the matching analysis for bill introductions and bills passed. Assuming the exclusion restriction holds, the IV analysis indicates that term-limited legislators are significantly more proactive, introducing 6-7 more bills and passing 4-5 more bills. And the IV estimates for resolutions and abstention rates largely support the matching analysis, though they are less precisely estimated; the results suggest term-limited legislators filed about .5 more resolutions than non-term-limited legislators and abstain on approximately 3-5 percent more roll-call votes. The only other coefficient that occasionally reaches conventional

[5]This depends on the restrictive assumption that there is zero correlation between the covariates and unobserved unit-level effects (e.g. ability, motivation etc.)

levels of statistical significance is vote share, indicating that as vote share increases so too does a legislator's bill introductions and bills passed.

As mentioned above, the instrument could prove to be invalid if year elected is related to other unmeasured causes of legislative activity. For example, a legislator who first wins during a Republican landslide could be affected ideologically by the experience, which in turn would affect their penchant for being an active legislator; omitting legislators' ideologies could potentially result in biased estimates of the LATE. To test for such a possibility, as a robustness check, the appendix reports IV regression estimates for the 2001-2011 sessions with the inclusion of CFScores, which provide a measure of legislators' ideologies based on Federal Election Commission (FEC) records of donors' campaign contributions (See Bonica (2013*a*) for details). The regressions return comparable point estimates. Overall, then, both the matching analysis and the instrumental variables regressions suggest that institutions, such as term limits, that incent progressive ambition can boost levels of proactive participation while dampening reactive forms of participation.

Table 3.6.: Instrumental Variables Estimates for Bill Introductions, 1999-2011

DV: Bill Introductions	Random Effects	Fixed Effects
Term-Limited	7.75*	6.44
	(.89)	(2.87)
Committee chair	-.22	-.50
	(1.12)	(1.29)
Vote share	.06*	.05
	(.02)	(.02)
Democrat	-.99	
	(1.06)	
Black	-.49	
	(1.56)	
Female	-.93	
	(1.23)	
Constant	3.40	1.64
	(1.74)	(2.79)
F-stat for instrument	19.06	21.31

Note: The term-limited dummy is instrumented with year first elected. Standard errors are in parentheses. The models includes year fixed effects. * denote statistical significance at the 1 percent level.

Table 3.7.: Instrumental Variables Estimates for Bills Passed, 1999-2011

DV: Bills Passed	Random Effects	Fixed Effects
Term-Limited	4.17*	5.53*
	(.57)	(2.02)
Committee chair	.81	.08
	(.71)	(.91)
Vote share	.04*	.03
	(.01)	(.02)
Democrat	.72	
	(.56)	
Black	-.77	
	(.82)	
Female	.04	
	(.62)	
Constant	.86	2.76
	(1.05)	(1.96)
F-stat for instrument	23.19	16.44

Note: The term-limited dummy is instrumented with year first elected. Standard errors are in parentheses. The models includes year fixed effects. * denote statistical significance at the 1 percent level.

Table 3.8.: Instrumental Variables Estimates for Resolutions Filed, 1999-2011

DV: Resolutions Filed	Random Effects	Fixed Effects
Term-Limited	.56*	.61
	(.14)	(.51)
Committee chair	.02	.11
	(.17)	(.23)
Vote share	.00	.00
	(.00)	(.00)
Democrat	-.06	
	(.12)	
Black	.15	
	(.17)	
Female	.62*	
	(.13)	
Constant	.21	.37
	(.24)	(.50)
F-stat for instrument	5.51	2.67

Note: The term-limited dummy is instrumented with year first elected. Standard errors are in parentheses. The models includes year fixed effects. * denote statistical significance at the 1 percent level.

Table 3.9.: Instrumental Variables Estimates for Abstention Rates, 2005-2011

DV: Abstention Rate	Random Effects	Fixed Effects
Term-Limited	5.01*	3.29
	(1.18)	(1.19)
Committee chair	-2.61	-2.16
	(1.42)	(1.49)
Vote share	-.01	.02
	(.02)	(.03)
Democrat	-1.16	
	(1.36)	
Black	2.31	
	(2.04)	
Female	-2.21	
	(1.53)	
Constant	9.60*	6.82*
	(2.33)	(2.42)
F-stat for instrument	5.82	5.77

Note: The term-limited dummy is instrumented with year first elected. Standard errors are in parentheses. The models includes year fixed effects. * denote statistical significance at the 1 percent level.

3.4 Discussion

This chapter addressed two questions. First, how do term limits affect the participatory decisions of legislators, and second, how can scholars effectively study term-limited legislative behavior using observational data. In reverse order, the principal challenge researchers face when examining the effects of a major reform like terms limits with observational data is dealing with nonrandom assignment to treatment. Using difference-in-means tests, I show that systematic differences exist between term-limited and non-term-limited state legislators, suggesting that simply comparing the outcomes of the two groups of legislators would result in questionable inferences due to invalid counterfactual comparisons.

In an effort to constructively advance the scholarship on term limits, I proposed two alternative estimation strategies—one based on matching and one based on instrumental variables regression—both of which attempt to address biases from confounds by focusing on more homogenous subpopulations of legislators, allowing one to recover more local average treatment effects. This in turn leads to the paper's second contribution, which is to describe the nuanced ways in which term limits affect legislative behavior. Specifically, using a panel of bill activity from the Arkansas General Assembly, the analysis indicates that on proactive forms of participation over which legislators have discretion, term-limited legislators display significantly higher levels of output, but on reactive forms of participation over which legislators have limited discretion, term-limited legislators display significantly lower levels of participation.

The theoretical explanation for this pattern of behavior centers on progressively ambitious office-seekers using bill introductions and nonideological appeals via resolutions to curry favor with constituents and to build name recognition as they prepare a run for higher office, while also abstaining on a higher proportion of roll-call votes so as to avoid alienating potential supporters. Notwithstanding the limited empirical scope, this study constitutes an important step toward understanding behavioral responses to changes in legislative institutions. While term limits advocates sought to put an end to careerism, it appears that

instead of forcing politicians from elective office, term limits incentivized them to run for other political positions.

Much remains to be learned about how term limits affect legislators' incentives, and this paper has only explored bill activity and floor participation. One possible extension would involve a more detailed examination of bill content. For example, to what extent do higher rates of turnover in term-limited legislatures affect the degree to which legislators specialize? Alternatively, one could shift the focus away from the legislature and toward the district to explore whether term-limited legislators are as actively involved in their district as their non-term-limited colleagues.

While the analysis in this chapter contributes to the existing empirical literature on progressive ambition and legislative behavior, it neglected to empirically address term-limited legislators' career goals aside from the pursuit of higher office. In that vein, another possible direction for future research alluded to in Section 3.2.1 would entail broadening the scope of ambition beyond higher office to include non-legislative employment as well. Addressing this question hinges on conceiving of progressive ambition as a multinomial choice where term-limited legislators consider not only bids for Congress but also private sector employment, lobbying or a return to their primary occupation. Those state legislators who do not view running for Congress as a viable option may instead find that they can use their policy and procedural knowledge as well as their personal connections to segue from government into a lucrative lobbying career. Ultimately, though, of principal import going forward is conducting studies with an emphasis on shoring up the internal validity of research designs, which will manifestly improve our understanding of the nexus between term limits and legislative behavior.

3.5 Appendix

This appendix presents estimates from instrumental variables regressions that use CFscores to measure a legislator's distance from the chamber median.

3.5.1 Additional Results

Tables 3.E.1., 3.E.2., 3.E.3. and 3.E.4. repeat the instrumental variables regressions from the main body of the paper but include CFscores to address one possible violation of the exclusion restriction, namely that the instrument, year elected, may be correlated with unmeasured determinants of bill activity, such as a legislator's ideology. CFscores are available for Arkansas state representatives for the 2001-2011 sessions, so we can examine our four outcome variables for six of the seven sessions. The results largely corroborate those reported in the paper, though the point estimates are more variable across model specifications. In the random effects models, for example, term-limited legislators still significantly outperform their non-term-limited counterparts, introducing more bills, passing more bills, filing more resolutions, while also abstaining on roll-calls at higher rates. In the fixed effects models, however, while the sign on the coefficients remains in the expected direction, the magnitude of the point estimates is notably smaller, often falling outside of conventional levels of statistical significance. Broadly, then, the IV regression results presented in the paper are generally robust to at least one potential concern about the instrument's validity.

Table 3.E.1.: Instrumental Variables Estimates for Bill Introductions, 1999-2011

DV: Bill Introductions	Random Effects	Fixed Effects
Term-Limited	7.86*	4.58
	(.99)	(6.87)
Committee chair	-.39	.38
	(1.22)	(1.36)
Vote share	.07*	.05
	(.02)	(.04)
Democrat	-3.08	
	(1.50)	
Black	-.23	
	(1.74)	
Female	-1.15	
	(1.37)	
CFScore	-3.00	2.46
	(1.59)	(2.05)
Constant	5.78	10.87
	(2.19)	(2.79)
F-stat for instrument	13.83	15.18

Note: The term-limited dummy is instrumented with year first elected. Standard errors are in parentheses. The models includes year fixed effects. * denote statistical significance at the 1 percent level.

Table 3.E.2.: Instrumental Variables Estimates for Bills Passed, 1999-2011

DV: Bills Passed	Random Effects	Fixed Effects
Term-Limited	4.03*	1.80
	(.62)	(4.24)
Committee chair	.33	.65
	(.77)	(.93)
Vote share	.04*	.02
	(.01)	(.02)
Democrat	-.08	
	(.83)	
Black	-.49	
	(.93)	
Female	.10	
	(.73)	
CFScore	-1.15	1.99
	(.97)	(1.40)
Constant	.15	5.63
	(1.30)	(8.15)
F-stat for instrument	17.78	12.27

Note: The term-limited dummy is instrumented with year first elected. Standard errors are in parentheses. The models includes year fixed effects. * denote statistical significance at the 1 percent level.

Table 3.E.3.: Instrumental Variables Estimates for Resolutions Filed, 1999-2011

DV: Resolutions Filed	Random Effects	Fixed Effects
Term-Limited	.63*	.04
	(.15)	(1.04)
Committee chair	-.01	.09
	(.18)	(.23)
Vote share	.01	.00
	(.00)	(.01)
Democrat	-.13	
	(.19)	
Black	.14	
	(.20)	
Female	.63*	
	(.16)	
CFScore	-.18	.15
	(.22)	(.34)
Constant	-.08	.91
	(.29)	(2.00)
F-stat for instrument	4.90	1.51

Note: The term-limited dummy is instrumented with year first elected. Standard errors are in parentheses. The models includes year fixed effects. * denote statistical significance at the 1 percent level.

Table 3.E.4.: Instrumental Variables Estimates for Abstention Rates, 2005-2011

DV: Abstention Rate	Random Effects	Fixed Effects
Term-Limited	5.72*	3.39
	(1.11)	(1.20)
Committee chair	-2.65	-2.13
	(1.37)	(1.52)
Vote share	-.01	.02
	(.02)	(.03)
Democrat	-1.16	
	(1.36)	
Black	2.31	
	(2.04)	
Female	-2.21	
	(1.53)	
CFScore	-.44	-1.85
	(1.76)	(2.52)
Constant	9.44*	6.63*
	(2.36)	(2.43)
F-stat for instrument	6.29	4.87

Note: The term-limited dummy is instrumented with year first elected. Standard errors are in parentheses. The models includes year fixed effects. * denote statistical significance at the 1 percent level.

CHAPTER IV

On the Consequences of Legislative Extremism: Evidence from the North Carolina General Assembly, 1996-2011

Abstract: What is the effect of ideology on a legislator's standing in the chamber? More specifically, are more ideologically extreme legislators as effective as their more moderate colleagues, and are they as likely to acquire prized positions on powerful committees? Applying both panel data models and propensity score matching, I examine these questions using a new source of data on state legislators' ideal points as well as survey-based measures of legislative effectiveness and powerful committees. The estimation strategies yield the same basic result: legislators farther from the chamber median are significantly less effective and are also less likely to chair or co-chair the most powerful committees in their respective chambers.

4.1 Introduction

In the aftermath of the 2010 Tea Party movement, political commenters were quick to criticize recalcitrant new members of Congress—particularly those who opposed raising the debt ceiling in the summer of 2011, which resulted in the credit rating agency Standard and Poor's (S&P) downgrading the U.S. federal government's credit rating to AA+ from AAA—questioning their commitment to legislating responsibly and effectively (Mann and Ornstein, 2013). Similar criticisms resurfaced during the budgetary stalemate in the fall of 2013. Led by Senator Ted Cruz, congressional Republicans initiated a government shutdown,

the result of which was an estimated 120,000 jobs lost along with a $24 billion hit to the economy and a credit downgrade of U.S. debt by China's largest credit rating agency (Council of Economic Advisers, 2013). The combination of these two events led to concerns that hard-line conservatives, driven by ideological purity, had seized control of the Republican Party.

Since 2010 conservative Republicans have swelled their ranks across state legislatures, too. These conservative Republican majorities have attempted to shift policy rightward on myriad issues from unemployment benefits and gun control to voting and abortion rights (Johnson, 2013). With these policies, they have attracted the ire of both the Justice Department and of discontented citizens, most notably in North Carolina (Reilly, 2013). Indeed, beginning in April 2013, thousands of North Carolina residents began to gather every week at the state capitol for Moral Monday protests, where nearly 1,000 people have been arrested (Biesecker, 2013). As the gap between Democrats and Republicans grows ever-wider across a majority of states (Shor, 2013), there are important questions about the effects an influx of ideologically extreme legislators will have on political parties and on legislatures. I examine one piece of this broader subject, asking (1) are more ideologically extreme legislators as effective as their more moderate colleagues and (2) are more ideologically extreme legislators as likely to acquire prized positions on powerful committees?

Devoid of ideal-point estimates for state legislators, scholars have heretofore been unable to answer the questions posed above. I examine these two questions using a survey collected by the North Carolina Center for Public Policy Research (NCCPPR) that gauges both legislative effectiveness and the most powerful committees in each chamber of the North Carolina General Assembly, and a new database of dynamic ideal point estimates for state legislators. At the end of each long session (i.e. odd-numbered years), the NCCPPR sends a survey to all legislators in both chambers, several hundred registered lobbyists in the state and approximately two dozen journalists, who cover the General Assembly, asking them to rank how effective each legislator was during that session and what the most powerful

committees were in each chamber.[1] And I use CFscores from Bonica's (2013*a*) Database on Ideology, Money in Politics, and Elections (DIME) to measure legislators' ideologies, which are based on Federal Election Commission records of donors' campaign contributions. I will discuss these measures in more detail later in the paper.

In part, the findings align with previous studies of legislative effectiveness in that (1) members of the majority party, women and members with backgrounds in law are perceived to be more effective legislators and (2) first-term members and black legislators are perceived to be less effective. But whereas the existing literature uncovers mixed findings on the relationship between ideology and effectiveness, I find that members farther from the chamber median are seen as significantly less effective legislators. Being perceived as less effective by colleagues, lobbyists and journalists is not terribly important if it doesn't have tangible consequences; in that vein, I also find that members farther from the chamber median are less likely to chair or co-chair the most powerful committees in their respective chambers, which in turn limits their ability to influence the legislative agenda.

Those findings do not, however, lend insight into why more extreme legislators are less effective and less likely to attain positions of power. In brief, the theoretical explanation presented elides individual-level reelection incentives with partisan theories of legislative organization. When electoral circumstances and policy preferences vary, legislators are likely to pursue different strategies to help maintain their individual incumbency. Channeling Mayhew, some members find it beneficial to prioritize position-taking on divisive issues, while other members pursue strategies oriented around credit-claiming, focusing their effort on comparatively less ideologically-charged legislative activities such as securing—or promising to secure—appropriations for their constituents, which helps them build reputations as effective legislators among colleagues and onlookers.

Unchecked, this individual-optimizing creates a collective action problem with which party leaders must deal. In seeking to protect and promote the party brand, party leaders

[1]Personal correspondence with NCCPPR staff who administer and collect the survey. Each group of respondents' rankings are weighted equally.

can structure the strategic environment to limit defections, in turn accumulating a body of legislative accomplishments that will buoy its party's brand as elections approach. One such way to do so is through committee assignments. In particular, on those committees whose actions are likely to affect the party's reputation, leaders can appoint members who are ideologically moderate.

The remainder of this paper proceeds as follows. In the next section, I provide an overview of the extant literature on legislative effectiveness at both the state level and at the congressional level. Then I lay out an explanation for why more ideologically extreme legislators can expect diminished policy fortunes along with a perception of being less effective legislators. From there I describe the North Carolina General Assembly and introduce the data sources. This is followed by a presentation of the empirical results. And the final section offers concluding remarks.

4.2 Existing Research on Legislative Effectiveness

Empirical studies of effectiveness have abounded in recent years. A large literature examines how legislature-wide conditions explain variation in effectiveness. At the state level, Kousser (2003) finds evidence that "batting averages"—the proportion of bills introduced that pass the chamber—are higher in more professionalized legislatures likely because of the additional staff available to members. Cain and Kousser (2004) and Kousser (2005) also use batting averages as their measure of effectiveness to examine the effects of term limits on committee gatekeeping and show that batting averages tend to be higher in term-limited legislatures. And in studies of Congress, scholars have amassed evidence that divided government tends to reduce legislative effectiveness (Binder, 1999; Howell et al., 2000).

A more recent set of studies explores the extent to which various legislator-specific characteristics influence effectiveness. Given some of the drawbacks associated with batting averages,[2] scholars have turned to alternative measures of legislative effectiveness, such as count

[2]Principally, batting averages do not account for the total number of bills a legislator introduces, which

models. One of the most consistent findings in the literature is that members of the majority party are significantly more effective at moving bills through the chamber due to, inter alia, the larger resource endowments majority status begets (Anderson, Box-Steffensmeier and Sinclair-Chapman, 2003; Cox and Terry, 2008). In addition, there is evidence that women tend to be more effective than male legislators (Anzia and Berry, 2011; Volden and Wiseman, 2013).

The aforementioned studies represent important contributions, but somewhat surprisingly—especially in light of the growing complaints about partisan conflict and congressional performance coupled with efforts to reform primary elections in a number of states—little has been done to estimate differences in effectiveness based on legislators' ideologies. To the extent that such work exists, it rarely emerges as the focal point of analysis, but instead enters as a control variable in the specification of a broader empirical model. Indeed, the lone empirical study by Cox and Terry (2008) produced inconclusive results; operationalizing effectiveness as the number of bills a legislator passes during a congressional session, the authors find that members of Congress whose ideal points are further from the House median see fewer bills pass the chamber, but in their random effects model, the variable is insignificant. Given the mixed findings across model specifications, the authors do not dwell on the results in their analysis. Extending this earlier line of research with a more definitive set of empirical tests can help clarify the ideology-effectiveness relationship.

What's more, whether examining legislature-wide effectiveness or individual-level effectiveness, attention is almost exclusively given to bill passage (see Haynie, 2002 and Miquel and Snyder, 2006 for exceptions). Missing is a measure of how good legislators are at the legislative process more broadly. The approach taken in this paper moves beyond the standard emphasis on bill progress through a chamber. To do so, what is needed are well-placed observers who can identify prime movers during a legislative session. The NCCPPR meets

can produce misleading results. For example, there can be a wide gap in the scope of achievement for a legislator who introduces one bill and sees it pass the chamber, thereby securing a batting average of 100 percent, and another legislator who introduces 15 bills, 10 of which pass the chamber, thereby securing a 66 percent batting average percentage.

that objective, administering a survey to journalists, legislators and lobbyists, who in concert are well positioned to observe how issues develop, whose views influence whom and how preferences get aggregated, thereby providing a more complete picture of legislative effectiveness than does data that focuses solely on bill passage (Hall, 1992). And because the NCCPPR has continuously administered the survey since 1977, I can take advantage of the data's panel structure to control for legislator-specific characteristics that may influence effectiveness.

Relative to objective measures such as batting averages or count models, there may be concern about the reliability of a subjective, survey-based measure. To allay these concerns, I examine the relationship between the NCCPPR's effectiveness scores and patterns of bills introduced and ratified. From 1997-2010, in North Carolina's lower chamber, the correlation between the NCCPPR effectiveness scores and bills introduced and bills ratified is .42 and .49, respectively; in the state senate, the correlations between the NCCPPR effectiveness scores and bills introduced and bills ratified are .60 and .62, respectively, suggesting that while the measures are related, the NCCPPR data also captures additional dimensions of legislative behavior beyond introducing and passing legislation. The next section lays out the theoretical relationship between legislators' ideologies and their standing in the chamber, describing why more ideologically extreme legislators are seen as less effective and why they are less likely to assume positions of power within legislative chambers.

4.3 Theory

In this section I draw on the existing literature to develop a theoretical explanation for why members on the ideological fringes are seen as less effective legislators and face diminished policy fortunes. Canonical work on legislative behavior in Congress posits that members have a mix of goals—reelection, policy changes, and influence in the chamber—and suggests several distinct strategies for them to achieve those goals (Fenno, 1978; Mayhew, 1974). Some members build a personal vote and avoid premature departure from a legislature

by emphasizing position-taking. For example, those who depend to a greater extent on the support of strong partisans to win elections are incented to engage in higher levels of partisan conflict, including more extreme voting records and the use of more vitriolic rhetoric (Canes-Wrone, Brady and Cogan, 2002; Harbridge and Malhotra, 2011). Indeed, recent work on Congress shows that ideological legislators place less emphasis on policymaking; in examining press releases, Grimmer (2013) finds that more ideologically extreme legislators tend to focus a considerable amount of their energy on position-taking, touting their views on divisive policy issues.

Other reelection-motivated legislators cultivate a base of support by claiming credit for policy changes, highlighting their efforts at securing—or promising to secure—appropriations for their constituents (Grimmer, Westwood and Messing, 2013). These legislators situate themselves deep in the policymaking weeds, building coalitions, striking deals and guiding their bills through the chamber to ensure porcine-laden bills make it back to their constituents. The implications of these two strategies for legislative effectiveness and influence in the chamber diverge, however; whereas appeals based on credit-claiming convey the impression of being able to get things done, thereby contributing to perceptions of effectiveness, the same cannot be said for those legislators who emphasize position-taking, for it is difficult to make the case that polarizing a policy debate or inducing gridlock constitute a valorized form of legislative effectiveness. From this logic that legislators adopt different strategies to cultivate a personal vote, I test the following hypothesis:

H1: as a legislator's distance to the chamber median increases, his or her legislative effectiveness decreases.

The preceding discussion of individual legislators' incentives gives rise to a collective dilemma. In the absence of any mechanism to rein in excesses, individual members can impose untold damage to the party brand through individual optimizing; legislators who primarily make use of position-taking will articulate one extreme position after another, promulgating a perception that the party is out of touch with key voting blocs, and legislators

who primarily make use of credit-claims will be tempted to propose a continuous stream of projects for their constituents to the point of fiscal profligacy. The party leadership represents one such mechanism to maintain discipline. As partisan models of legislative organization make clear, the party leadership wants to remain in the majority and doing so requires protecting the collective good of the party's brand, which is based in no small part on the legislative record of Congress. By cartelizing the legislative agenda, invoking both positive and negative agenda power, senior partners in the cartel—particularly floor leaders and committee chairs—can build a favorable corpus of accomplishments to buoy the electoral fortunes of its members when they face voters (Cox and McCubbins, 2005).

The provenance of any such policy accomplishments must emanate from jurisdiction-specific standing committees, the assignments of which are controlled by the party leadership. While cartel theory hypothesizes that committees serve as tools of the majority party, their decisions can also adversely affect non-committee members. One way to minimize any such negative externalities from occurring is for party leaders to vigilantly police committee activity to ensure that their efforts benefit party members. A more practical alternative, however, is to simply keep fringe members of the chamber from chairing important committees;[3] these members would be more likely to support extreme bills harmful to the party brand, which in turn would require larger payoffs to moderate members so as to prevent defections. By contrast, when more moderate members of the party hold key agenda-setting posts, the party can be reasonably confident that committees will prevent the passage of policies that harm the majority's brand and encourage the passage of policies that benefit its brand. With this in mind, the party leadership should be more likely to empower relative moderates, who will advance a credible partisan agenda that is electorally useful to most of its members.

[3]Along with their control over positions on sought-after committees and their ability to block undesirable bills from reaching the floor, senior partners can also dissuade members from undercutting the party's brand through targeted spending. For example, Carroll and Kim (2010) find that the majority party uses pork barrel spending to ensure members support the party on procedural votes. In a similar vein, Jenkins and Monroe (2012) argue another means in which the majority party sustains the cartel is through campaign contributions from leaders' political action committees.

Evidence abounds that legislatures operate along the lines described by the cartel model. In the U.S. Congress, for example, defiant members who oppose the leadership on important votes find themselves no longer sitting on prized committees (Rohde, 1991). More closely related to the North Carolina context, state legislators who try to advance bills likely to reflect poorly on the party's brand, such as legislation outlawing Islamic sharia law in state courts—not known to be an especially pressing public policy problem in North Carolina—have little hope of seeing their bills emerge from committee (Zucchino, 2013).

If it is the case that senior partners use their institutional authority to shape the legislative environment, and if ideologically extreme legislators are more likely to attempt to pass bills that would reflect poorly on their party's brand, one might also expect them to be less likely to hold high-ranking positions on important committees where they would be able to exercise agenda-setting powers. This leads to a second hypothesis:

> **H2**: as a legislator's distance to the chamber median increases, the probability that he or she chair (or co-chair) a powerful committee decreases.

The remainder of this article tests these hypotheses. Before turning to the data, the following section offers a brief description of the North Carolina General Assembly.

4.4 Background and Data

The North Carolina General Assembly is a non-term-limited legislature and has two chambers, a 50-seat Senate with four-year terms and a 120-seat House of Representatives with two-year terms. In Squire's (2007) index of legislative professionalism, North Carolina ranked 13th, and in the National Conference of State Legislature's (NCSL) Red-White-Blue trifurcation, North Carolina is considered a White, or hybrid, legislature based on its intermediate-sized staff and pay, as legislators do not earn enough to make a living without having other sources of income (National Conference of State Legislatures, 2009). Rank and file members make an annual salary of $13,951, whereas chambers leaders such as the Speaker of the House and President Pro Tempore of the Senate earn $38,151 (North Carolina

General Assembly, 2013).

Party composition in the General Assembly has seen both Democrats and Republicans in control. For the period under study, Republicans held a majority in the House from 1995 to 1998, before Democrats regained control up through 2002. The 2003-2004 session witnessed a 60-60 split, leading both parties to share the top leadership posts in the chamber with co-speakers of the House. From 2005-2010, Democrats were in the majority, a majority they maintained until the November 2010 Republican landslide, and Republicans remain in the majority as of the 2013-2014 session. In the Senate, Democrats held the majority of seats from 1995-2010, before Republicans ascended to the majority.

The General Assembly also operates along the lines described in cartel theory in a number of ways. For instance, it is the prerogative of the Speaker of the House and Pro Tempore of the Senate to appoint standing committee chairs in their respective chambers. Further, it is at the committee chair's discretion to hear and report bills out of committee. And finally, the Rules Committee in each chamber determines the floor calendar.[4] The combination of these features lends itself to a legislature in which the majority party and its leadership exercise significant control over the legislative agenda, the committees and the rules.

Dependent Variables

To examine the effects of ideology on legislative effectiveness and the acquisition of high-ranking positions on powerful committees in the North Carolina General Assembly, I use data from the North Carolina Center for Public Policy Research (NCCPPR), a nonpartisan organization in North Carolina. As mentioned in the introduction, the NCCPPR administers its survey at the end of the General Assembly's long legislative session, soliciting the perspectives from currently serving senators and representatives, several hundred registered lobbyists in the state and one to two dozen state capital news correspondents with a sample size between 237-308 and an average response rate of 47 percent from 1996-2011, a rate con-

[4]Personal correspondence with General Assembly staff.

siderably higher than recent surveys of elites. In its biannual report, the NCCPPR explains,

> "The survey's purpose was to identify the most effective legislators in the General Assembly. The Center asked respondents to base their ratings on legislators' participation in committee work, their skill at guiding bills through committee and floor debate, their general knowledge and expertise in special fields, the respect they command from peers, his or her ethics, the enthusiasm with which they execute various legislative responsibilities, the political power they hold (by virtue of office, longevity, or personal skills), their ability to sway the opinions of fellow legislators, and their aptitude for the overall legislative process."

The quote above makes it clear that the NCCPPR's survey-based measure of legislative effectiveness is not solely about introducing and passing bills, but rather attempts to capture a panoply of different skills legislators may possess. In an effort to deal with the multifaceted nature of legislative effectiveness, in the empirical results section, I report estimates from models with legislator fixed effects, which will help mitigate the threat of omitted variable bias that can arise from unmeasured factors that vary across individual legislators (i.e. work ethic, aptitude, interest, etc.) but are constant over time. The results from the survey are then used to create an ordinal scale of legislative effectiveness, ranking representatives from 1 to 120 and senators from 1 to 50 with one indicating the most effective legislator in the chamber. Most NCCPPR reports include all 170 members of the General Assembly, though there are exceptions when members either resign or pass away during the session and, thus, are not assigned an effectiveness rating. For the legislative effectiveness dependent variable to test H1, I invert the scale so that higher numbers indicate more effective legislators.

There may be both substantive and methodological concerns about the dependent variable; regarding the former, in its bluntest form, one might be tempted to ask, who cares if a legislator is seen as effective or not? Put tersely, legislators do. A state senator at the time, Kay Hagan noted on her website that she was named "one of North Carolina's 'Ten Most Effective Senators' three terms in a row by the non-partisan North Carolina Center for

Public Policy Research." And the import attached to effectiveness ratings is not confined to the state senate, as Representative Deborah Ross lists her ranking as "the 9th most effective member of the NC House of Representatives in 2010" among her professional accomplishments on her website.

Methodologically, one might worry about possible response bias resulting from the evaluations of legislators, lobbyists and journalists; for example, journalists' partisan sensibilities may seep into their evaluations of legislators, leading them to systematically rate one party more favorably than the other. These concerns can be allayed for two reasons. First, in terms survey design, quantitative questions tend to engender less partisan bias from respondents (Ansolabehere, Meredith and Snowberg, 2012). More specifically, within the context of the NCCPPR, from 1977-1992, when the NCCPPR reported the average ranking that each legislator was given from the three types of respondents, Miquel and Snyder (2006) find strong correlations between respondents: .93 between legislators and lobbyists, .89 between legislators and journalists and .91 between lobbyists and journalists

The second dependent variable used to test H2 is an indicator for whether the member is the chair or co-chair of one of the most powerful committees in the General Assembly. In addition to being asked to rank legislators based on their effectiveness during the session, respondents are also asked to identify the most powerful committees in each chamber. Across the sessions from 1997-2009, the list typically includes Appropriations, Finance, Rules, Judiciary, Commerce and Education. In 2001, following the decennial census, the Legislative Redistricting committee made the list in both chambers.

Covariates

The key independent variable of interest is legislator ideology. I test whether more extreme legislators are less effective and less likely to chair or co-chair powerful committees by taking the absolute value of the distance between a legislator's ideal point and the median ideal point in the chamber for that session. While several scholars have produced state

Table 4.1.: Description of Ideal Point Estimates for North Carolina General Assembly

	Shor	Bonica
Years Covered	1995-2010	1996-2011
Recovery Method	Roll Call Votes	Campaign Contributions
House Median	.10	.05
Senate Median	-.55	-.04

legislative ideal points, I use Bonica's (2013*a*) dynamic CFscores, which are based on Federal Election Commission (FEC) campaign contribution records. Because they are available for each election cycle from 1996 to 2010, these ideal points are best equipped for handling the data's panel structure.

Another option would be Shor and McCarty's (2013) common-space ideal points used in (Shor and McCarty, 2011). These estimates are based on state legislative roll-call votes from the mid-1990s to the late 2000s. Table 4.1. provides a brief description of each data source along with chamber medians. Similar to NOMINATE, ideal point estimates less than zero indicate more liberal legislatures and scores greater than zero indicate more conservative legislatures. Averaging across the time-series, the Shor and Bonica estimates return comparable results, both indicating a slightly conservative lower chamber and a liberal upper chamber.[5]

Both sources are in common space, but the main difference and advantage of the Bonica estimates is that legislators' ideal points vary over time. By contrast, the Shor estimates provide a single ideal point for each legislator, save for those legislators who either switched parties or who came to represent a different district.[6] The justification for providing a single ideal point over the course of a member's career is a reference to Poole (1998) who finds that members of Congress typically die in their ideological boots.

[5]In calculating the median from the Shor data, twenty of the 378 observations are omitted as those legislators served in both chambers.

[6]In addition to the invariant nature of Shor's NPAT scores, several legislators with the same last name are elided, thereby requiring the researcher to dispose of a dozen or more observations, depending on the length of the time series one elects to use.

At both the national and the state level, though, evidence exists that weakens the sinews of such claims. For example, Clinton, Jackman and Rivers (2004) find that with the unexpected party switch of Senator Jim Jeffords in 2001, which catapulated Tom Daschle into the majority leader position and demoted Trent Lott to minority leader, both Daschle and Lott became more conservative, with Daschle in particular moving approximately 30 places rightward. And at the state level, Kousser, Lewis and Masket (2007) also help make the case for why time-variant estimates of ideal points are so valuable; they show how Democratic state legislators in California ideologically adapt to the changing electoral landscape, attempting to fend off defeat by becoming more conservative as Republicans gained electoral momentum during the 2003 gubernatorial recall.

The validity of CFscores emanates from their recovery of a liberal-conservative dimension as well as their ability to correctly classify roll-call voting outcomes at rates comparable to DW-NOMINATE (Bonica, 2013*b*). And more specifically with respect to North Carolina, the correlation between CFscores and Shor's NPAT scores is over .80. Figure 4.1. plots the CFscores and NPAT scores for North Carolina legislators. Both density plots reveal a clear bimodality, bifurcated along party lines. The distributions do differ in that the CFscores are more tightly ideologically dispersed than the NPAT scores (sd=.27 vs. sd=.44 for Democrats and sd=.25 vs. sd=.46 for Republicans).

In addition to ideology, there are many other potential factors that may influence a member's effectiveness. The analysis, thus, controls for several variables that previous research suggests might influence legislative effectiveness, beginning with an indicator for majority party status, anticipating that members of the majority party are better positioned, by virtue of additional resources and staff, to see their legislative priorities shepherded through the chamber, thereby gaining recognition as effective legislators. I include an indicator for members who hold top leadership positions, coded 1 for those members in the House who serve as speaker, majority and minority leader, and majority and minority whip; similarly, in the Senate, the indicator is coded 1 for those who serve as president pro tempore, majority

Figure 4.1.: Distribution of Ideal Points

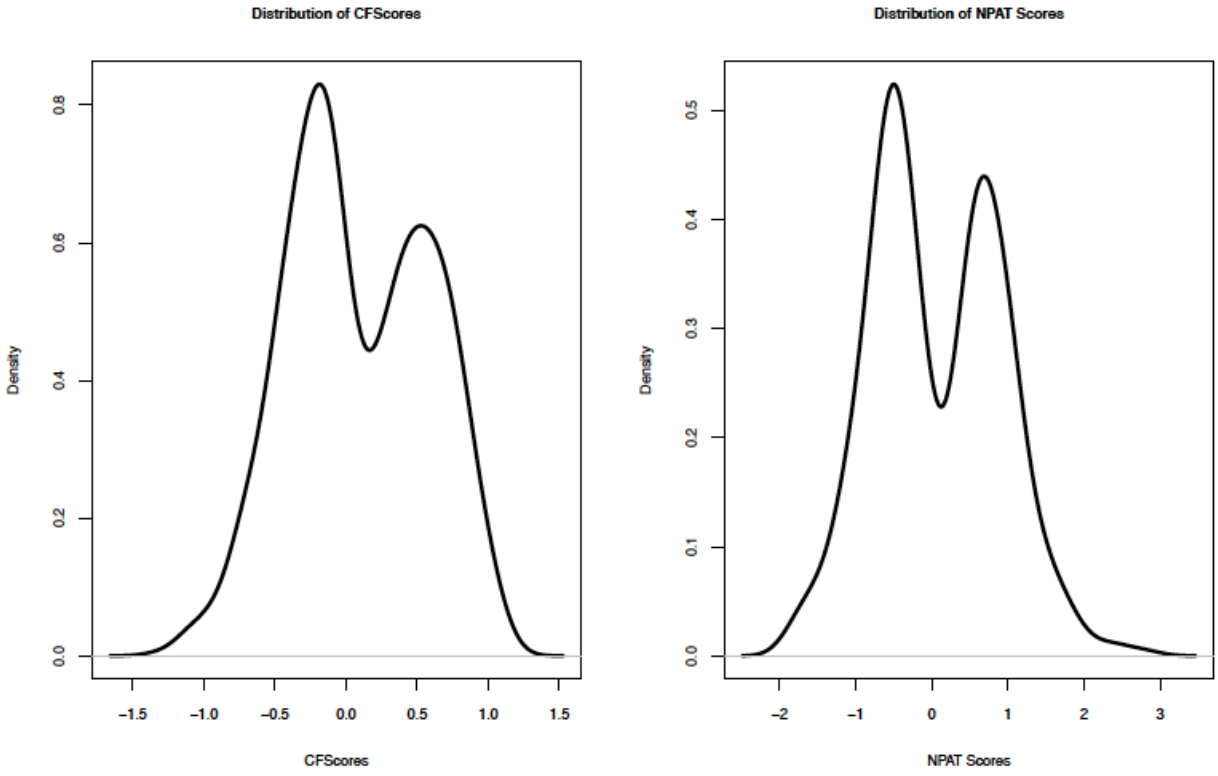


Table 4.2.: Summary Statistics for the NC House

Variable	Mean	SD	Min	Max
House Effectiveness	60.73	34.56	1	120
Majority Party	.53	.49	0	1
Senior Partner	.06	.25	0	1
Freshman	.15	.35	0	1
Distance from Median	.42	.26	.00	1.37
Terms Served	4.51	3.24	.5	19
Age	58.26	11.29	24	91
Black	.15	.36	0	1
Female	.25	.43	0	1
Lawyer	.16	.36	0	1

and minority leader, and majority and minority whip.

And in light of work that suggests longer-serving members are more effective, I include a number of sessions served variable. I also include indicators for black legislators, freshman, women and those legislators who list lawyer or attorney as their occupation. Based on prior research, I expect black and freshman legislators to be less effective, while women and legislators with backgrounds in law ought to be more effective (Haynie, 2002; Miquel and Snyder, 2006; Anzia and Berry, 2011). And along with the aforementioned controls, I include year fixed effects throughout the analysis to account for any unobserved time-varying factors. A table of descriptive statistics for each chamber is presented in Table 4.2. and 4.3..

4.5 Empirical Results

Testing H1

In this section, I begin by testing H1, which predicts that more extreme legislators will see lower levels of effectiveness. The model takes the following form:

$$Effectiveness_{it} = \alpha_i + \beta(\iota_{it}) + \psi X_{it} + \delta_i + \gamma_t + \epsilon_{it} \tag{4.1}$$

Table 4.3.: Summary Statistics for the NC Senate

Variable	Mean	SD	Min	Max
Senate Effectiveness	25.69	14.34	1	50
Majority Party	.62	.49	0	1
Senior Partner	.10	.30	0	1
Freshman	.14	.35	0	1
Distance from Median	.40	.30	.00	1.52
Terms Served	4.10	2.99	.5	17
Age	59.36	11.41	28	89
Black	.14	.35	0	1
Female	.13	.34	0	1
Lawyer	.33	.47	0	1

where subscript i denotes individual legislators and t denotes time. The variable of interest is ι_{it}, which is the absolute value of a legislator's distance from the chamber median. The vector X includes the broad set of control variables described in the previous section, δ_i captures time-invariant individual effects, γ_t represents year fixed effects and ϵ_{it} is an idiosyncratic error term.

Before discussing the findings, it is important to note that the panel is unbalanced, which raises potential sample selection issues that I attempt to address in the following section. Table 4.G.6. presents regression results for the North Carolina House from 1996-2011. The first column reports the coefficient estimates from the fixed effects regression, and the second column reports the coefficient estimates from the random effects regression. The latter are useful insofar as providing insight into how time-invariant individual-level characteristics affect the dependent variable—in this case, race, gender and occupation—but they should be interpreted with some caution, as a Hausman specification test comparing the fixed and random effects estimates rejects the null at the 1 percent level, suggesting use of the fixed effects model.[7]

[7]The shortcomings of the Hausman test are well-documented. One cannot compare coefficients on time-invariant variables, nor can one include coefficients on variables that only change over time (Wooldridge, 2010). And recent work based on Monte Carlo simulations by Clark and Linzer (2013) argues that the Hausman test is nether a necessary nor sufficient test statistic for choosing between fixed and random

Table 4.4.: Analysis of Legislative Effectiveness in the NC House, 1996-2011

DV: Effectiveness	(1) Fixed Effects	(2) Random Effects
Distance from Median	-14.51***	-16.36***
	(5.06)	(3.97)
Majority Party	33.12***	33.45***
	(2.15)	(1.88)
Senior Partner	15.90***	18.64***
	(3.24)	(3.03)
Freshman	-19.58***	-20.10***
	(2.52)	(2.25)
Terms Served	0.46	2.60***
	(1.63)	(0.45)
Age	1.78**	-0.42***
	(0.85)	(0.12)
Black		-20.23***
		(3.56)
Female		5.99**
		(2.99)
Lawyer		12.54***
		(3.59)
N	892	892

Note: Standard errors are in parentheses. All models include year fixed effects, and column (1) includes legislator fixed effects as well. The dependent variable is the inverted effectiveness score for House members. * ** *** denote statistical at the 10 percent, 5 percent and 1 percent levels, respectively.

While controlling for both member-specific and institutional factors, the results demonstrate that distance from the median is statistically significant at the 1 percent level and negatively correlated with legislative effectiveness. This indicates that legislators farther from the median—in other words, more ideologically extreme members—tend to be perceived as less effective, providing evidence in support of H1. In particular, moving one-and-a-half standard deviations below the median to one-and-a-half standard deviations above the median reduces a legislator's effectiveness rating by about twelve positions in the fixed effects model and about thirteen positions in the random effects model. In the appendix, I report

effects, but rather the decision should be based on the structure of the panel data one uses, particularly the number of observations per unit and the correlation between the covariates and unit-level effects.

additional regressions, including results bifurcated by party as well as results that consider possible omitted variables. Figure 4.2. plots this relationship over the full range of values for distance from the median.

Figure 4.2.: Predicted relationship between distance from the median and effectiveness in the NC House

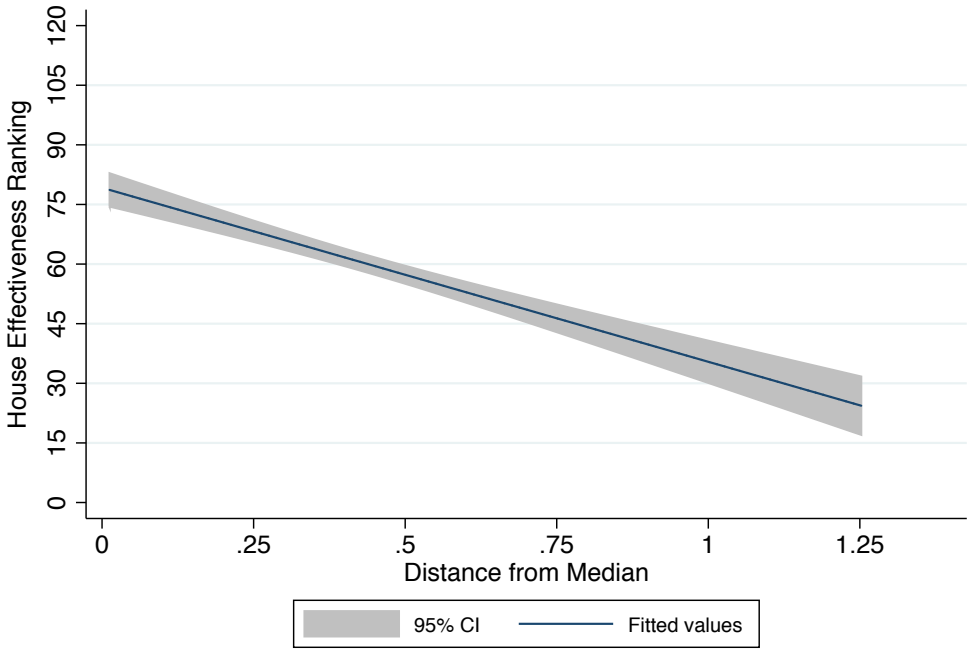


Figure 2: Predicted relationship between distance from the median and effectiveness in the NC House

The remaining variables behave as expected. The coefficient on the senior partner variable is large, positive and statistically significant, indicating that those who hold top leadership positions in the state house are seen as more effective legislators. And consistent with previous studies, members of the majority party, women and members with backgrounds in law are significantly more effective, while freshman and black legislators are significantly less effective. The random effects estimates indicate that being a freshman or black legislator, for example, decreases a legislator's effectiveness ranking by approximately twenty positions. Longer-serving legislators are seen as more effective, though the effects of tenure are only significant in the random effects model.

Table 4.5. presents results from the North Carolina senate. The differences between the

fixed and random effects models are larger, but again, the Hausman test suggests using the former. Substantively, the results are similar to those in the lower chamber with the signs of all the variables pointing in the same direction, but the coefficient on women legislators is no longer statistically distinguishable from zero. Distance from the chamber median continues to have a negative effect on effectiveness and is statistically significant at the 5 percent level in the fixed effects model and at the 1 percent level in the random effects model. The estimated effects for distance from the median suggest that moving one standard deviation below the median to one standard deviation above the median reduces a legislator's effectiveness rating by approximately three to six positions (fixed and random effects estimates, respectively), roughly the equivalent of going from a chamber leader to a rank and file member. Figure 4.3. plots this relationship over the full range of values for distance from the median. In sum, the results of these empirical tests are consistent with H1. What is notable, moreover, is that while the lone existing study by Cox and Terry (2008) is inconclusive on the ideology-effectiveness relationship, the results across models and across chambers reach the same conclusion: ideological extremity reduces legislative effectiveness.

Table 4.5.: Analysis of Legislative Effectiveness in the NC Senate, 1996-2011

DV: Effectiveness	(1) Fixed Effects	(2) Random Effects
Distance from Median	-5.54**	-9.70***
	(2.63)	(2.19)
Majority Party	12.87***	10.59***
	(1.78)	(1.44)
Senior Partner	5.23***	5.66***
	(1.52)	(1.39)
Freshman	-6.19***	-5.76***
	(1.27)	(1.17)
Terms Served	0.63	1.03***
	(1.40)	(0.27)
Age	0.31	-0.06
	(0.70)	(0.07)
Black		-4.52
		(2.14)
Female		2.73
		(2.20)
Lawyer		7.61***
		(1.78)
N	384	384

Note: Standard errors clustered by legislator are in parentheses. All models include year fixed effects, and column (1) includes legislator fixed effects as well. The dependent variable is the inverted effectiveness score for Senate members. * ** *** denote statistical at the 10 percent, 5 percent and 1 percent levels, respectively.

Figure 4.3.: Predicted relationship between distance from the median and effectiveness in the NC Senate

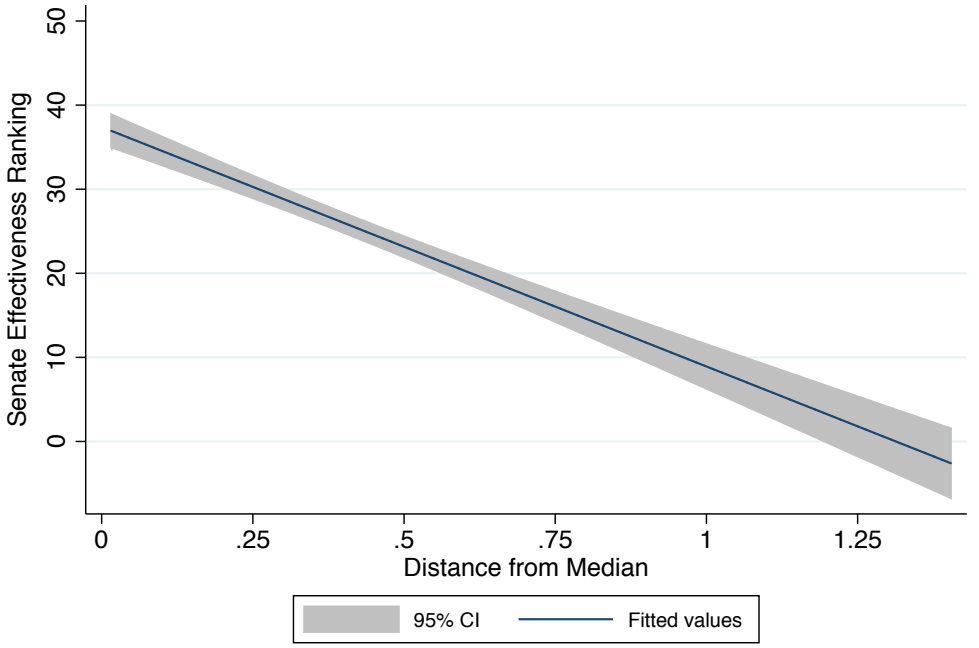


Figure 3: Predicted relationship between distance from the median and effectiveness in the NC Senate

Robustness Check: Propensity Score Matching

While the models above provide strong support for the hypothesis that more extreme legislators are less effective, there might be concerns that the individual fixed effects may be insufficient if there is imbalance or a lack of overlap that limits comparability. As a robustness check I use propensity score matching, a selection-on-observables strategy for observational data; the basic idea is to align the distribution of observables as closely as possible for a sample of units such that one compares units that are similar, save for treatment receipt. The propensity score, or conditional probability of assignment to treatment given a vector of observed covariates, simplifies this process of matching on potential pre-treatment confounds by creating a one-number summary that predicts the probability that a given unit receives the treatment (Rosenbaum and Rubin, 1983). One can then estimate the treatment effect of interest—in this case, the effect of ideology on effectiveness—by comparing the difference

in the outcome across relevant subgroups.

To more firmly ground the estimation strategy, I use potential outcomes notation, where each unit has two potential outcomes: $Yi(1)$ denotes the potential outcome in the treatment condition for individual i, and $Yi(0)$ denotes the potential outcome in the control condition for individual i. The treatment effect is the difference in legislative effectiveness for a legislator in extreme and moderate conditions; for this binary case, a treatment variable, z, takes the value 1 for those units exposed to the treatment condition and 0 for those units in the control condition.[8] Below I focus on average treatment effect on the treated (ATT).

$$ATT = E[Y1|z = 1] - E[Y0|z = 1] \tag{4.2}$$

The estimand provides a sense of the average treatment effect of ideology on legislative effectiveness for those who typically adopt extreme ideologies. Put another way, the ATT is the expected difference in legislative effectiveness for a randomly selected extreme legislator in both moderate and extreme states.

Dichotomizing CFscores into treated and control groups requires the creation of a cutoff for what denotes an extreme legislator.[9] Legislators with CFscores one standard deviation above the mean during a legislative session are placed in the treatment group of extreme.[10] In estimating the propensity score equation predicting extremism, I use a probit regression and include five pre-treatment covariates—party affiliation, race, gender, age and occupation—as well as year fixed effects and the lagged effectiveness ranking, which improved balance (Imbens, 2004).[11]

[8]While I use a binary conceptualization of ideologically extreme legislators, with a sufficiently large data set there is nothing to prevent researchers from using multi-valued treatments (Imai and Dyk, 2004).

[9]Bifurcating legislators into ideological groups is not uncommon; Jenkins and Monroe (2012), for instance, create an indicator for moderate legislators to examine whether they receive a disproportionate share of contributions from party leaders' PACs as compensation for policy losses.

[10]To assess the sensitivity of the results, I also conducted the analysis using one standard deviation above the median, and the results were similar.

[11]The intuition for using lagged effectiveness is that lagged outcomes are a function of both observable and unobservable covariates, and, thus, by conditioning on a lagged outcome, one can indirectly condition

Table 4.6.: Balance Between Treated and Control Group Means NC House

	Unmatched		Matched	
	Treated	Control	Treated	Control
Democrat	.39	.56***	.39	.34
Black	.10	.16	.10	.09
Lawyer	.17	.15	.17	.18
Female	.35	.23**	.35	.25
Lagged Effectiveness	54.69	63.89***	54.69	48.65
Age	55.80	59.26***	55.80	56.47

Note: ** *** denote statistical significance at the 5 percent and 1 percent levels, respectively, of t-tests evaluating difference in means.

Table 4.6 presents the means of the covariates for moderate and extreme legislators in the lower chamber for both the unmatched and matched samples. Prior to the propensity score matching, there was considerable imbalance. For example, moderate legislators were significantly more likely to be Democratic, older and as expected also enjoy significantly higher effectiveness rankings. Matching reduced the difference in sample means across all of the variables included in the calculation of the propensity score. Column 1 of Table 4.7 reports the ATT, indicating that extreme legislators in the state house can expect to be about 12 positions less effective, which mirrors the estimates from the panel data models.

Shifting to the senate, Table 4.8. presents the means of the covariates for moderate and extreme legislators for both the unmatched and matched samples. First, it should be noted that three observations in the treated group are dropped from the analysis because they are off common support; in other words, the matching estimator was unable to find a counterfactual observation. Again, as was the case in the lower chamber, there was considerable imbalance in the senate prior to matching. Among other differences, the moderate group of legislators is significantly more Democratic and includes significantly more black legislators. In the matched sample, though, no significant differences exist between the treated and control groups. Column 2 of Table 4.7 reports the ATT and shows that extreme senators can expect to be about 10 positions less effective. In sum, then, the results from matching

on unobservables (Keele, 2015).

Table 4.7.: Average Treatment Effect on the Treated (ATT) NC House and Senate

DV: Effectiveness	(1) House	(2) Senate
Extremism	-12.29***	-10.49***
	(4.62)	(3.82)
Treated(N)	127	47
Control(N)	654	219
N	781	266

Note: Standard error in parentheses. *** denotes statistical significance at the 1 percent level.

Table 4.8.: Balance Between Treated and Control Group Means NC Senate

	Unmatched		Matched	
	Treated	Control	Treated	Control
Democrat	.24	.68***	.26	.30
Black	.04	.15**	.04	.11
Lawyer	.38	.33	.40	.28
Female	.28	.08***	.23	.26
Lagged Effectiveness	14.92	28.82***	14.87	13.11
Age	59.22	60.56	59.38	62.02

Note: ** *** denote statistical significance at the 5 percent and 1 percent levels, respectively, of t-tests evaluating difference in means.

are consistent with the patterns observed in the panel data models in that more extreme legislators can expect to be perceived as less effective.

Ideology and Powerful Committees

Testing H2

Simply being perceived as less effective by colleagues, lobbyists and journalists is unlikely to keep such legislators up at night. But it may be a cause for concern if it tangibly undermines their influence in the legislature. In this section, I test H2, which predicts that more extreme legislators will be less likely to chair powerful committees. Given the dichotomous nature of the dependent variable, I estimate probit regressions with year fixed effects. Table

Table 4.9.: Analysis of Powerful Committees in the NC General Assembly, 1997-2009

DV: Pr(Powerful Chair)	(1) House	(2) Senate
Distance from Median	-.80*	-3.95***
	(.45)	(1.89)
Majority Party	1.38***	2.92**
	(.29)	(1.28)
Senior Partner	.14	1.35
	(.30)	(1.04)
Freshman	-1.11**	
	(.53)	
Terms Served	.15***	.41**
	(.04)	(.16)
Age	-.02*	.02
	(.01)	(.04)
Black	-.33	-1.71
	(.29)	(1.13)
Female	.32	1.93
	(.25)	(1.14)
Lawyer	.14	1.46
	(.28)	(.96)
N	817	346

Note: Standard errors are in parentheses. The models show probit coefficients and include year fixed effects. The dependent variable is the probability of chairing or co-chairing a powerful committee. * ** *** denote statistical significance at the 10 percent, 5 percent and 1 percent levels, respectively.

4.9 presents the results for both the lower and upper chambers.

As expected, in both chambers, longer-serving legislators and being in the majority party significantly increases the probability of chairing or co-chairing a powerful committee. As for the independent variable of interest, distance from the median, the results provide support for H2. The coefficients for distance from the median are negative, indicating that as distance from the chamber median increases, the probability that a legislator chairs or co-chairs a powerful committee decreases. In the senate, the coefficient is significant at the 5 percent level, while in the house, the coefficient is significant at the 10 percent level.

Figures 4.4 and 4.5 present predicted probabilities for the lower and upper chambers,

respectively. In the lower chamber, when continuous variables are held at their mean values, an increase in distance from one-and-a-half standard deviations below the median to one-and-a-half standard deviations above the median decreases the probability that a rank and file member of the majority party chairs or co-chairs a powerful committee by 8 percent, from 28 percent down to 20 percent. In the state senate the prospects of an extreme legislator chairing or co-chairing a powerful committee are even starker; with continuous variables held at their mean values, an increase in distance from one standard deviation below the median to one standard deviation above the median decreases the probability that a rank and file member of the majority party chairs or co-chairs a powerful committee by 34 percent, from 45 percent down to 11 percent.

Figure 4.4.: Predicted Probability of Chairing a Powerful Committee in the NC House

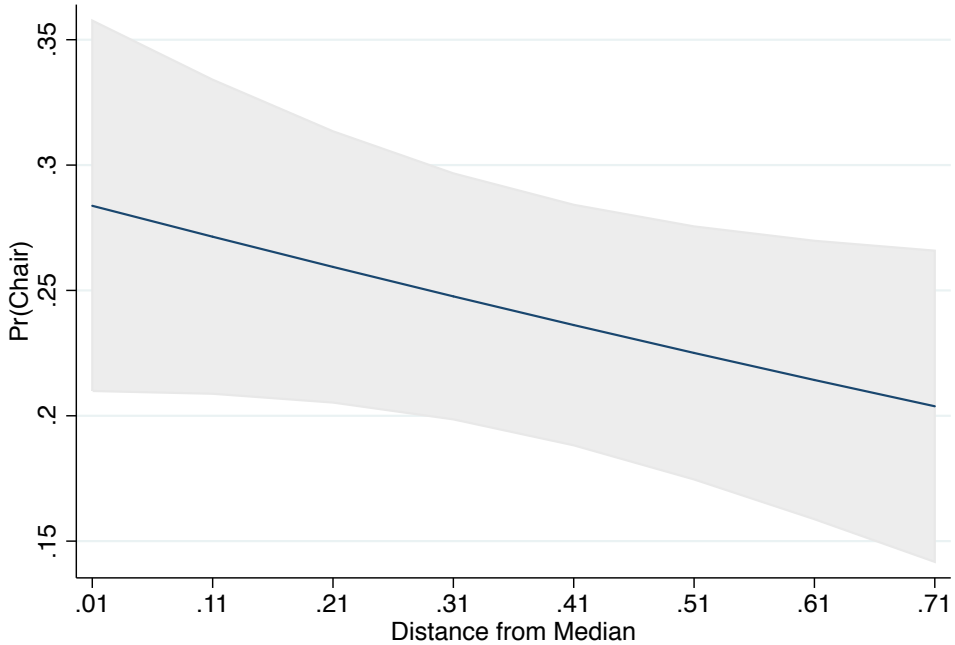


Figure 4: Predicted Probability of Chairing a Powerful Committee in the NC House

These results provide insight into some of the consequences of legislative extremism. Even after controlling for common member-specific and institutional factors, the results consistently support the notion that a legislator's ideology is associated with both their effectiveness and their likelihood of holding positions on prized committees.

Figure 4.5.: Predicted Probability of Chairing a Powerful Committee in the NC Senate

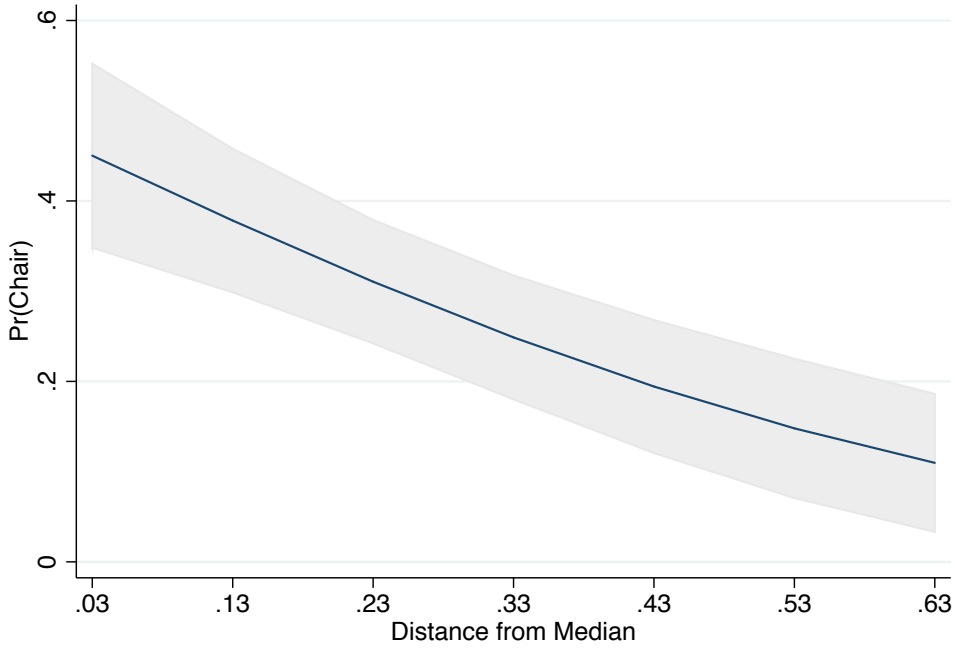


Figure 5: Predicted Probability of Chairing a Powerful Committee in the NC Senate

4.6 Discussion

In this chapter I examined whether more ideologically extreme legislators are as effective as their more moderate colleagues and as likely to chair or co-chair the most powerful committees in their respective chambers. While confirming several findings from previous studies—members of the majority party, women, and members with backgrounds in law are significantly more effective legislators and first-term members and black legislators are significantly less effective—the main takeaway across models and across chambers is that legislators farther from the chamber median are less effective and less likely to chair or co-chair the most powerful committees in the North Carolina General Assembly.

These results add to the scholarly literature on legislative effectiveness by moving beyond standard measures that center entirely on bill progress through a chamber as well as to research on legislative organization. The fact that more moderate legislators are more likely to hold powerful positions could be seen as somewhat reassuring news. These committees have

jurisdiction over crucial policies—taxes, agriculture, spending priorities and constitutional issues—where imprudent decisions and poorly drafted legislation can impose serious costs on citizens.

But because the analysis ends in 2011, it has only begun to scratch the surface of how legislatures are dealing with the post-Tea Party influx of ideologically extreme legislators. Clearly, more research is needed, both in North Carolina and in other states. In North Carolina, with successive wave elections in 2010 and 2012, the Republican Party not only regained the governor's mansion after two decades of Democratic control, but also captured large majorities in both chambers of the General Assembly for the first time since 1870. Consistent with the predictions of partisan models such as conditional party government, as the number of conservative Republican lawmakers increased, consensus emerged on a wide range of economic and social issues, resulting in significant partisan effects on policy (Rohde, 1991). This increased homogeneity encouraged party and committee leaders to pursue legislative action that has engendered so much consternation in the state.

As Rohde (2013) points out, the length of time a party has been out of power has important consequences for agenda control. In particular, the longer a party has been out of power, the more important positive agenda control becomes relative to negative agenda control. That dynamic played out with the 1994 Republican takeover of Congress after four decades in the minority, and it is currently playing out in North Carolina. In both cases, Republicans wanted to move policy in a conservative direction, which in turn meant that gatekeeping to maintain the status quo proved to be nowhere near as important as invoking positive agenda control.

With the ideological distance between congressional Democrats and Republicans at historic highs, reaching compromise has become more challenging than ever (Binder and Lee, 2014). One extension would involve taking advantage of the variation in polarization across the states to see if some state legislatures have remained islands of compromise amid a sea of gridlock. Future research might also examine other instances of how legislators' ideology af-

fects legislative performance. Are more polarized state legislatures as effective in bargaining with the governor? Can they adequately conduct oversight of the bureaucracy? Exploring these and a host of other empirical questions will add to our understanding of representation in legislatures and accountability in political institutions.

4.7 Appendix

This appendix begins by presenting estimates from fixed and random effects regressions for each party separately. Then, I present estimates of models that include an indicator for legislators to who hold a committee chair post to address possible concerns about omitted variable bias as well as an interaction between committee chair and distance from the median. And finally, to further explore the extent to which relative moderates occupy key agenda-setting posts, I compare the median ideology of powerful committee chairs and co-chairs to their party medians.

4.7.1 Results by Party

In light of work at the congressional level by Volden and Wiseman (2013), which suggests legislative effectiveness could vary by party, Tables 4.G.1. and 4.G.2. examine H1, which centers on the relationship between a legislator's distance from the median and his or her effectiveness, for the North Carolina state house, repeating the panel data models from the main body of the paper for each party separately. While in some instances the point estimates on the variable of interest—distance from the median—are marginally smaller, the main results still obtain: distance from the median remains negatively correlated with legislative effectiveness.

For Democrats, moving one-and-a-half standard deviations below the median to one-and-a-half standard deviations above median reduces a legislator's effectiveness rating by about ten positions in the fixed effects model and about twelve positions in the random effects model. And for Republicans, moving one-and-a-half standard deviations below the median to one-and-a-half standard deviations above median reduces a legislator's effectiveness rating by about nine positions in the fixed effects model and about ten positions in the random effects model.

Turning to the state senate, the estimates in Tables 4.G.3. and 4.G.4. again corroborate the point estimates reported in the main body of the paper, but in the fixed effects model,

Table 4.G.1.: Analysis of Legislative Effectiveness in the NC House, Democratic Party, 1996-2011

DV: Effectiveness	(1) Fixed Effects	(2) Random Effects
Distance from Median	-13.09**	-13.81**
	(6.60)	(5.52)
Majority Party	43.47***	12.86***
	(5.43)	(5.13)
Senior Partner	9.49**	12.09***
	(4.15)	(3.62)
Freshman	-21.90***	-23.48***
	(3.21)	(2.99)
Terms Served	1.45***	3.25***
	(.48)	(0.45)
Age	2.76***	-0.31**
	(0.43)	(0.16)
Black		-19.74***
		(3.30)
Female		-1.96
		(3.46)
Lawyer		12.16***
		(4.06)
N	469	469

Note: Standard errors are in parentheses. All models include year fixed effects, and column (1) includes legislator fixed effects as well. The dependent variable is the inverted effectiveness score for House members. * ** *** denote statistical at the 10 percent, 5 percent and 1 percent levels, respectively.

Table 4.G.2.: Analysis of Legislative Effectiveness in the NC House, Republican Party, 1996-2011

DV: Effectiveness	(1) Fixed Effects	(2) Random Effects
Distance from Median	-11.21	-12.52**
	(7.04)	(5.73)
Majority Party	29.11***	25.66***
	(2.72)	(4.78)
Senior Partner	21.84***	25.23***
	(7.96)	(7.49)
Freshman	-17.29***	-17.78***
	(3.01)	(2.78)
Terms Served	1.45***	2.15***
	(.48)	(0.77)
Age	2.76***	-0.60***
	(0.43)	(0.17)
Black		-20.38*
		(11.46)
Female		15.50***
		(3.94)
Lawyer		14.42*
		(8.86)
N	423	423

Note: Standard errors are in parentheses. All models include year fixed effects, and column (1) includes legislator fixed effects as well. The dependent variable is the inverted effectiveness score for House members. * ** *** denote statistical at the 10 percent, 5 percent and 1 percent levels, respectively.

Table 4.G.3.: Analysis of Legislative Effectiveness in the NC Senate, Democratic Party, 1996-2011

DV: Effectiveness	(1) Fixed Effects	(2) Random Effects
Distance from Median	-5.10	-10.40**
	(4.01)	(4.46)
Majority Party		20.81***
		(6.65)
Senior Partner	1.12	4.18
	(2.97)	(2.62)
Freshman	-7.50***	-8.13***
	(1.90)	(1.90)
Terms Served	.04	1.46***
	(.95)	(0.42)
Age	-.11***	.13
	(0.28)	(0.11)
Black		-5.32*
		(2.76)
Female		1.40
		(3.42)
Lawyer		9.81***
		(2.37)
N	228	228

Note: Standard errors are in parentheses. All models include year fixed effects, and column (1) includes legislator fixed effects as well. The dependent variable is the inverted effectiveness score for House members. * ** *** denote statistical at the 10 percent, 5 percent and 1 percent levels, respectively.

the coefficients on distance from the median are no longer statistically significant. For both Democrats and Republicans, the estimated effects for distance from the median suggest that moving one standard deviation below the median to one standard deviation above the median reduces a legislator's effectiveness rating by approximately three to seven positions in the fixed and random effects models, respectively.

Table 4.G.4.: Analysis of Legislative Effectiveness in the NC Senate, Republican Party, 1996-2011

DV: Effectiveness	(1) Fixed Effects	(2) Random Effects
Distance from Median	-4.82	-11.16***
	(3.22)	(3.82)
Majority Party		14.99***
		(3.05)
Senior Partner	9.16	7.84
	(1.20)	(1.48)
Freshman	-5.42***	-6.20***
	(1.95)	(1.72)
Terms Served	.47	.42
	(1.65)	(0.56)
Age	1.32***	.01
	(0.24)	(0.09)
Black		
		(2.76)
Female		4.15
		(4.50)
Lawyer		9.19**
		(3.62)
N	156	156

Note: Standard errors are in parentheses. All models include year fixed effects, and column (1) includes legislator fixed effects as well. The dependent variable is the inverted effectiveness score for House members. * ** *** denote statistical at the 10 percent, 5 percent and 1 percent levels, respectively.

4.7.2 Considering Omitted Variables

In the main body of the paper I incorporate institutional position by including indicators for legislators in the majority party and for legislators who hold one of the leadership posts in their respective chamber as key predictors of a legislator's effectiveness. Work by Volden and Wiseman (2009, 2014), however, finds that committee chairs tend to be significantly more effective than rank and file members. So to address potential concerns about omitted variable bias, I estimate regressions that include an indicator for legislators to who hold a committee chair post as well as an interaction between the committee chair indicator and distance from the median.

While the interaction is indistinguishable from zero in every specification for both chambers, the committee chair indicator was clearly an omitted variable: it has a large, positive effect on legislative effectiveness in all of the models, and its inclusion marginally reduces the magnitude on the distance from the median variable. In the house, being a committee chair increases a legislator's effectiveness ranking by about fourteen to seventeen positions in the fixed and random effects models, respectively. And in the senate, being a committee chair increases a legislator's effectiveness ranking by about six to eight positions in the fixed and random effects models, respectively.

4.7.3 Comparing Ideologies: Committee Chairs and Party Medians

When testing my hypotheses in the main body of the paper, I focus on a legislator's distance from the chamber median, ignoring the party median. To remedy that lacuna, in this section, I descriptively compare the median ideology of powerful committee chairs and co-chairs to their party medians. This subject is far from terra incognita, as a number of scholars have explored the ideological orientations of party leaders in Congress; using DW-NOMINATE scores, some of these studies find support for the middleperson hypothesis, whereby party leaders are often ideologically proximate to the median member of their

party,[12] while other studies find that party leaders tend to be extremists, drawn from the fringes of their party (Grofman, Koetzle and McGann, 2002; Harris and Nelson, 2008; Jessee and Malhotra, 2010).

To add to this literature, the first and second rows in Table 4.G.7. provide the median CFscores for the Democratic Party's chairs and co-chairs of powerful committees in the North Carolina state house and state senate, respectively, both of which are pooled across the entire panel from 1996 to 2011. Table 4.G.8. provides the same measures for the Republican Party. Recall that CFscores less than zero indicate more liberal ideologies, and CFscores greater than zero indicate more conservative ideologies. What is notable is the tendency for powerful committee chairs to come from a particular side of the party median; while still generally proximate to their respective party's median, thereby lending some support for the middleperson hypothesis, for both parties, committee chairs in the North Carolina General Assembly tend to hold a more centrist ideology.

[12]The middleperson hypothesis is consistent with partisan theories of legislative organization, as adumbrated in the main body of the paper, in that party leaders—to whom agenda setting powers are delegated—remain ideologically proximate to a majority of their party and thus are well-positioned both to advance legislation that reflects favorably on the party's brand and to keep the party from getting rolled.

Table 4.G.5.: Analysis of Legislative Effectiveness in the NC House, 1996-2011

DV: Effectiveness	(1) Fixed Effects	(2) Random Effects
Distance from Median	-13.74***	-15.07***
	(4.85)	(3.72)
Majority Party	28.77***	29.65***
	(2.28)	(1.92)
Senior Partner	15.14***	17.79***
	(4.07)	(3.67)
Freshman	-19.49***	-19.74***
	(2.25)	(2.09)
Terms Served	0.19	2.12***
	(0.47)	(0.43)
Age	1.48**	-0.38***
	(0.34)	(0.11)
Black		-19.06***
		(3.14)
Female		5.43**
		(2.45)
Lawyer		12.53***
		(3.70)
Committee chair	14.36****	17.31***
	(4.04)	(3.85)
Comm chair * Distance	5.07	-0.20
	(10.05)	(10.29)
N	892	892

Note: Standard errors are in parentheses. All models include year fixed effects, and column (1) includes legislator fixed effects as well. The dependent variable is the inverted effectiveness score for House members. * ** *** denote statistical at the 10 percent, 5 percent and 1 percent levels, respectively.

Table 4.G.6.: Analysis of Legislative Effectiveness in the NC Senate, 1996-2011

DV: Effectiveness	(1) Fixed Effects	(2) Random Effects
Distance from Median	-4.79	-9.77***
	(2.97)	(2.70)
Majority Party	10.57***	7.71***
	(1.99)	(1.76)
Senior Partner	5.01**	6.12***
	(1.99)	(1.70)
Freshman	-6.11***	-6.50***
	(1.44)	(1.23)
Terms Served	0.60	0.80***
	(0.72)	(0.30)
Age	0.46***	0.02
	(0.16)	(0.06)
Black		-3.58
		(2.22)
Female		1.77
		(2.30)
Lawyer		8.68***
		(1.72)
Committee chair	6.61****	8.63***
	(2.44)	(2.06)
Comm chair * Distance	-5.17	-5.87
	(8.32)	(6.65)
N	384	384

Note: Standard errors are in parentheses. All models include year fixed effects, and column (1) includes legislator fixed effects as well. The dependent variable is the inverted effectiveness score for House members. * ** *** denote statistical at the 10 percent, 5 percent and 1 percent levels, respectively.

Table 4.G.7.: Median Ideology of Power Committee Chairs and Party Medians, Democratic Party, 1996-2011

Chamber	Committee Chair/Co-chair Median	Party Median
House	-.22	-.27
Senate	-.17	-.28

Table 4.G.8.: Median Ideology of Power Committee Chairs and Party Medians, Republican Party, 1996-2011

Chamber	Committee Chair/Co-chair Median	Party Median
House	.41	.56
Senate	.49	.55

CHAPTER V

Conclusion

This dissertation contributes to the study of legislative behavior and state legislative institutions, and in so doing, has a bearing on lively debates about the need for, and effects of, government reforms. One of the challenges state politics scholars face when analyzing the relationship between legislative institutions and legislative behavior is overcoming the selection problems that make causal inference so difficult. Questions such as how do term limits affect legislators' participation or how does a legislator's ideology affect their standing within a chamber are undoubtedly important insofar as their implications for accountability and representation, and each of the preceding chapters provided new evidence to answer these questions while taking to care address the challenges that arise when treatments are non-randomly assigned.

The adoption of term limits across nearly half of the states between 1990 and 2000 has presented researchers with a valuable opportunity to study legislative behavior in the absence of reelection incentives. Although I am not the first to examine how term limits affects legislative behavior, I approach the question with a focus on research designs needed to overcome often overlooked methodological obstacles. Both Chapters II and III outlined research designs to more credibly study term-limited behavior. Chapter II, coauthored with Rocío Titiunik, constitutes the first experimental examination of last-term effects. We took advantage of two constitutional provisions in Arkansas—the random assignment of term

length following reapportionment and term limits—the combination of which results in the random assignment of state senators to lame-duck status. A popular conception of the role of elections is that they incentivize politicians to exert effort (Barro, 1973; Ferejohn, 1986; Mayhew, 1974). The expectation, then, is that when this incentive is removed via term limits, voters can no longer sanction representatives for infelicitous behavior, thereby enabling various forms of shirking (e.g. reducing effort, adopting out-of-step ideological positions etc.).

Buoying this intuition is the extant literature at the state legislative level, which has routinely found evidence of last-term effects (Carey et al., 2006; Sarbaugh-Thompson et al., 2004; Wright, 2007). We first provide evidence that a commonly used observational research design comparing term-limited and non-term-limited legislators rests on an invalid counterfactual comparison due to the systematic differences between the two groups. And then, to explore if legislators shirk when the electoral connection to constituents is severed, using our experimental design we examine four outcomes: bill introductions, bills passed, resolutions filed and abstention rates. In contradistinction to a number of observational studies of term-limited legislative behavior, our results did not uncover any evidence that the absence of a reelection incentive induces lower levels of effort. We interpret the results as being consistent with a selection model of representation in which voters select competent, internally-motivated representatives who share the district's goals and work to achieve those goals even in the absence of a reelection incentive.

The results have implications for states considering relaxing their term limit laws. Showing systematic differences in the electoral fortunes of term-limited and non-term-limited legislators across eight state legislatures suggests that term limits may not be necessary if elections already tend to oust underperforming legislators. And failing to uncover any evidence of participatory shirking among lame-duck senators points to some of the costs associated with term limits. After having elected scrupulous legislators, voters have hamstrung themselves by proscribing such legislators from continuing to serve in the chamber.

In addition, with term limits, voters have a narrower window with which to build a relationship with their representative, and the shorter tenure in office means the representative has less time to develop substantive expertise across different policy domains before drafting legislation.

Chapter III continued the study of term limits, though did so in an observational context using a modified version of a research design proposed in Chapter II. Given the aforementioned differences between term-limited and non-term-limited legislators, I draw on two estimation strategies—one based on matching and one based on instrumental variables regression—both of which restrict comparisons to more homogenous units in order to recover a local average treatment effect (LATE) for a subset of the study population. The evidence from a panel of bill activity in the lower chamber of the Arkansas General Assembly suggests that term-limited legislators are aware of the dangers of position-taking amid their pursuit up the political ladder. More specifically, term-limited legislators display significantly higher levels of participation on proactive activities (e.g. bill introductions, constituency service) over which they have a fair amount of discretion and significantly lower levels of participation on reactive activities (e.g. roll-call votes) over which they have limited discretion, revealing the nuanced effects term limits can have on legislators' participatory decisions.

In addition to providing scholars with a more credible methodological approach to analyzing term-limited legislative behavior, these results show that there may be differences in the effects of term limits depending on the chamber under study. Term limits represent one of the clearest examples of an institution that affects legislators' incentives to seek higher office. And Chapter III demonstrates that when the prospect for chamber-climbing remains an option—as is the case for those serving in state houses—progressive ambition can motivate legislators to take a strategic approach to the forms of participation they undertake when preparing to make a run for the upper chamber.

The pattern of behavior whereby progressively ambitious term-limited legislators are active on some fronts but far less active on others is likely to be much less common in state

senates, which often mark the end of political careers, especially in dead-end legislatures where the prospects for advancement are limited (Maestas, 2000). For example, among the 64 Arkansas senators examined in Chapter II, only four went on to run in a congressional race. Since nearly half of the U.S.'s state legislatures are considered dead-end bodies, most state senators may not need to worry about casting votes or taking positions that will come back to haunt them as most will not go on to run for higher office.

Finally, Chapter IV focused a different question, examining how legislators' ideologies affect their standing in their respective chambers. More specifically, I examine the relationship between a legislator's distance from the chamber median and their effectiveness ranking as well as their likelihood of chairing a powerful committee. Using two estimation strategies—panel data models and propensity score matching—I find that legislators further from the chamber median are perceived to be less effective and are also less likely to chair a powerful committee in the North Carolina General Assembly.

These results dovetail nicely with recent empirical work at both the congressional and the state legislative level that has found that not only do more ideologically extreme legislators pay a price at the ballot box via a diminution in vote share, but more effective legislators tend to enjoy higher levels of electoral success (e.g. less likely to face challengers and more likely to win reelection) (Canes-Wrone, Brady and Cogan, 2002; Anzia and Berry, 2011; Miquel and Snyder, 2006). The combination of elections, which beget positive sorting of legislators, and internal legislative institutions, which endeavor to overcome collective action problems by keeping certain legislators from accessing key agenda-setting posts, should be seen as encouraging. For amid recent efforts to overhaul electoral institutions via primary reform, there is evidence that existing institutions can to some extent curb the influence of ideologically extreme legislators.

BIBLIOGRAPHY

BIBLIOGRAPHY

Ahler, Douglas, Jack Citrin and Gabriel Lenz. 2014. "Do Open Primaries Improve Representation? An Experimental Test of California's 2012 Top-Two Primary." Forthcoming at Legislative Studies Quarterly.

Alt, James, Ethan Bueno de Mesquita and Shanna Rose. 2011. "Disentangling Accountability and Competence in Elections: Evidence from U.S. Term Limits." *The Journal of Politics* 73(1):171–186.

Anderson, William, Janet Box-Steffensmeier and Valeria Sinclair-Chapman. 2003. "The Keys to Legislative Success in the U.S. House of Representatives." *Legislative Studies Quarterly* 28(3):357–386.

Angrist, Joshua, Guido Imbens and Donald Rubin. 1996. "Identification of Causal Effects Using Instrumental Variables." *Journal of the American Statistical Association* 91(434):444–455.

Ansolabehere, Stephen, Marc Meredith and Eric Snowberg. 2012. "Asking About Numbers: How and Why?" Working Paper.

Anzia, Sarah and Christopher Berry. 2011. "The Jackie (and Jill) Robinson Effect: Why Do Congresswomen Outperform Congressmen?" *American Journal of Political Science* 55(3):478–493.

Arnold, Douglas. 1990. *The Logic of Congressional Action*. New Haven, CT: Yale University Press.

Barro, Robert J. 1973. "The control of politicians: an economic model." *Public choice* 14(1):19–42.

Bender, Bruce and John Lott. 1996. "Legislator Voting and Shirking: A Critical Review of the Literature." *Public Choice* 87(1-2):67–100.

Berger, Roger and Jason Hsu. 1996. "Bioequivalence trials, intersection-union tests and equivalence confidence sets." *Statistical Science* 11(4):283–319.

Berman, David. 2007. The Effects of Legislative Term Limits in Arizona: More Churning, More Chaos, and a Diminished Institutional Role for Legislators. In *Legislating Without Experience: Case Studies in State Legislative Term Limits*, ed. Rick Farmer, Christopher Mooney, Richard Powell and John Green. Lanham, Maryland: Lexington Books.

Bernhard, William and Brian Sala. 2006. "The Remaking of an American Senate: The 17th Amendment and Ideological Responsiveness." *The Journal of Politics* 68(2):345–357.

Besley, Timothy and Anne Case. 1995. "Does Electoral Accountability Affect Economic Policy Choices? Evidence from Gubernatorial Term Limits." *The Quarterly Journal of Economics* 110(3):769–798.

Biesecker, Michael. 2013. "NAACP Leader, 11 Others Guilty In North Carolina 'Moral Monday' Protests." *The Huffington Post* . December 4, 2013.

Binder, Sarah. 1999. "The Dynamics of Legislative Gridlock, 1947-96." *American Political Science Review* 93(3):519–33.

Binder, Sarah and Frances Lee. 2014. Making Deals in Congress. In *APSA Task Force on Negotiating Agreement in Politics*, ed. Jane Mansbridge and Cathie Jo Martin.

Bonica, Adam. 2013*a*. "Database on Ideology, Money in Politics, and Elections.". Accessed on July 7, 2013.
URL: *http://data.stanford.edu/dime*

Bonica, Adam. 2013*b*. "Ideology and Interests in the Political Marketplace." *American Journal of Political Science* 57(2):294–311.

Bowers, Jake, Mark M Fredrickson and Costas Panagopoulos. 2013. "Reasoning about interference between units: a general framework." *Political Analysis* 21(1):97–124.

Burden, Barry. 2007. *Personal Roots of Representation*. Princeton, NJ: Princeton University Press.

Butler, Daniel and Joshua Revesz. 2012. "The Electoral Connection and Legislative Policy Proposals." Working Paper.

Cain, Bruce and Thad Kousser. 2004. *Adapting to Term Limits: Recent Experiences and New Directions*. San Francisco, CA: Public Policy Institute of California.

Canes-Wrone, Brandice, David Brady and John Cogan. 2002. "Out of Step, out of Office: Electoral Accountability and House Members' Voting." *American Political Science Review* 96(1):127–140.

Carey, John M. 1998. *Term limits and legislative representation*. Cambridge University Press.

Carey, John M., Richard G. Niemi and Lynda W. Powell. 1998. "The effects of term limits on state legislatures." *Legislative Studies Quarterly* 23(2):271–300.

Carey, John, Richard Niemi and Lynda Powell. 2000. *Term Limits in the State Legislatures*. Ann Arbor, MI: The University of Michigan Press.

Carey, John, Richard Niemi, Lynda Powell and Gary Moncrief. 2006. "The Effects of Term Limits on State Legislatures: A New Survey of the 50 States." *Legislative Studies Quarterly* 31(1):105–134.

Carroll, Royce and Henry Kim. 2010. "Party Government and the 'Cohesive Power of Public Plunder'." *American Journal of Political Science* 54(1):34–44.

Clark, Tom and Drew Linzer. 2013. "Should I Use Fixed or Random Effects?" Working Paper.

Clinton, Joshua, Simon Jackman and Douglas Rivers. 2004. "The Statistical Analysis of Roll Call Data." *American Political Science Review* 98(2):355–370.

Council of Economic Advisers. 2013. "Impacts and Costs of the October 2013 Federal Government Shutdown." November 2013.

Cox, Gary and Mathew McCubbins. 1993. *Legislative Leviathan: Party Government in the House*. Berkeley, CA: University of California Press.

Cox, Gary and Mathew McCubbins. 2005. *Setting the Agenda: Responsible Party Government in the U.S. House of Representatives*. New York: Cambridge University Press.

Cox, Gary and William Terry. 2008. "Legislative Productivity in the 93d-105th Congresses." *Legislative Studies Quarterly* 33(4):603–618.

Diermeier, Daniel, Michael Keane and Antonio Merlo. 2005. "A Political Economy Model of Congressional Careers." *The American Economic Review* 95(1):347–373.

Eggers, Andrew and Jens Hainmueller. 2009. "MPs for Sale? Returns to Office in Postwar British Politics." *American Political Science Review* 103(4):513–533.

English, Art and Brian Weberg. 2007. Term Limits in the Arkansas General Assembly: A Citizen Legislature Responds. In *Legislating Without Experience: Case Studies in State Legislative Term Limits*, ed. Rick Farmer, Christopher Mooney, Richard Powell and John Green. Lanham, Maryland: Lexington Books.

Erikson, Robert and Gerald Wright. 2000. Representation and Constituency Ideology in Congress. In *Continuity and Change in House Elections*, ed. David Brady, John Cogan and Morris Fiorina. Stanford, CA: Stanford University Press.

Fearon, James D. 1999. Electoral Accountability and the Control of Politicians: Selecting Good Types Versus Sanctioning Poor Performance. In *Democracy, Accountability, and Representation*, ed. Bernard Manin, Adam Przeworski and Susan Stokes. New York: Cambridge University Press.

Fenno, Richard. 1978. *Home Style: House Members in Their Districts*. Boston, MA: Addison Wesley.

Ferejohn, John. 1986. "Incumbent performance and electoral control." *Public choice* 50(1):5–25.

Fiorina, Morris, Samuel J. Abrams and Jeremy Pope. 2005. *Culture War? The Myth of a Polarized America*. New York: Pearson Education.

Gailmard, Sean and Jeffrey Jenkins. 2009. "Agency Problems, the 17th Amendment, and Representation in the Senate." *American Journal of Political Science* 53(2):324–342.

Glazer, Amihai and Martin Wattenberg. 1996. How Will Term Limits Affect Legislative Work? In *Legislative Term Limits: Public Choice Perspectives*, ed. Bernard Grofman. Boston, MA: Kluwer Academic Publishers.

Gordon, Stacy and Cynthia Unmack. 2003. "The Effect of Term Limits on Corporate PAC Allocation Patterns: The More Things Change ..." *State and Local Government Review* 35:26–37.

Green, Donald and Alan Gerber. 2012. *Field Experiments: Design, Analysis, and Interpretation*. New York: W.W. Norton and Company Inc.

Grimmer, Justin. 2013. "Appropriators not Position Takers: The Distorting Effects of Electoral Incentives on Congressional Representation." *American Journal of Political Science* 57(3):624–642.

Grimmer, Justin, Sean Westwood and Solomon Messing. 2013. "The Impression of Influence: How Legislator Communication and Government Spending Cultivate a Personal Vote.".

Grofman, Bernard, William Koetzle and Anthony McGann. 2002. "Congressional Leadership 1965-96: A New Look at the Extremism versus Centrality Debate." *Legislative Studies Quarterly* 27:87–105.

Hall, Andrew. 2014. "Partisan Effects of Term Limits." *Legislative Studies Quarterly* .

Hall, Andrew. 2015. "What Happens When Extremists Win Primaries?" *American Political Science Review* 109(1):18–42.

Hall, Andrew and James Snyder. 2014. "Candidate Ideology and Electoral Success." Working Paper.

Hall, Richard. 1992. "Measuring Legislative Influence." *Legislative Studies Quarterly* 17(2):205–231.

Hall, Richard. 1996. *Participation in Congress*. New Haven, CT: Yale University Press.

Harbridge, Laurel and Neil Malhotra. 2011. "Electoral Incentives and Partisan Conflict in Congress: Evidence from Survey Experiments." *American Journal of Political Science* 55(3):494–510.

Harris, Douglas and Garrison Nelson. 2008. "Middlemen No More? Emergent Patterns in Congressional Leadership Selection." *PS: Political Science and Politics* 41:49–55.

Haynie, Kerry. 2002. "The Color of Their Skin or the Content of Their Behavior? Race and Perceptions of African American Legislators." *Legislative Studies Quarterly* 27(2):295–314.

Herrick, Rebekah, Michael Moore and John Hibbing. 1994. "Unfastening the Electoral Connection: The Behavior of U.S. Representatives when Reelection Is No Longer a Factor." *The Journal of Politics* 56(1):214–227.

Herrick, Rebekah and Sue Thomas. 2005. "Do Term Limits Make a Difference? Ambition and Motivations Among U.S. State Legislators." *American Politics Research* 33(5):726–747.

Hirano, Shigeo and James Snyder. 2012. "What Happens to Incumbents in Scandals?" *Quarterly Journal of Political Science* 7:447–456.

Hirano, Shigeo and James Snyder. 2014. "Primary Elections and the Quality of Elected Officials." *Quarterly Journal of Political Science* 9:473–500.

Howell, William, Scott Adler, Charles Cameron and Charles Riemann. 2000. "Divided Government and the Legislative Productivity of Congress, 1945-94." *Legislative Studies Quarterly* 25(2):285–312.

Imai, Kosuke and David Van Dyk. 2004. "Causal Inference with General Treatment Regimes: Generalizing the Propensity Score." *Journal of the American Statistical Association* 99(467):854–866.

Imbens, Guido. 2004. "Nonparametric Estimation of Average Treatment Effects Under Exogeneity: A Review." *The Review of Economics and Statistics* 86(1):4–29.

Jefferson, Thomas. 1893. Resolution for Rotation of Members of Continental Congress. In *The Writings of Thomas Jefferson*, ed. Paul Leicester Ford. Vol. II New York: G.P. Putnam's Sons.

Jenkins, Jeffrey and Nathan Monroe. 2012. "Buying Negative Agenda Control in the U.S. House." *American Journal of Political Science* 56(4):897–912.

Jenkins, Jeffrey and Timothy Nokken. 2008. "Partisanship, the Electoral Connection and Lame-Duck Sessions of Congress, 1877-2006." *The Journal of Politics* 70(2):450–465.

Jessee, Stephen and Neil Malhotra. 2010. "Are Congresional Leaders Middlepersons or Extremists? Yes." *Legislative Studies Quarterly* 35(3):361–392.

Johnson, Luke. 2013. "North Carolina Voter ID Bill Signed Into Law By Gov. Pat McCrory, Sparking Lawsuits." *The Huffington Post* . August 12, 2013.

Keele, Luke. 2015. "The Statistics of Causal Inference: A View from Political Methodology." *Political Analysis* 23(3):313–335.

Kousser, Thad. 2003. Term Limits and Legislator Performance. In *The Test of Time: Coping with Legislative Term Limits*, ed. Rick Farmer, John David Rausch Jr. and John Green. Lanham, Maryland: Lexington Books.

Kousser, Thad. 2005. *Term Limits and the Dismantling of State Legislative Professionalism*. New York: Cambridge University Press.

Kousser, Thad, Jeffrey Lewis and Seth Masket. 2007. "Ideological Adaptation? The Survival Instinct of Threatened Legislators." *The Journal of Politics* 69(3):828–843.

Lee, Frances. 2009. *Beyond Ideology: Politics, Principles, and Partisanship in the U. S. Senate*. Chicago, IL: University of Chicago Press.

Lee, Frances. 2015. "How Party Polarization Affects Governance." *Annual Review of Political Science* 18:261–82.

Lessig, Lawrence. 2012. *Republic, Lost: How Money Corrupts Congress and a Plan to Stop It*. New York: Twelve.

Maestas, Cherie. 2000. "Professional Legislatures and Ambitious Politicians: Policy Responsiveness of State Institutions." *Legislative Studies Quarterly* 25(4):663–690.

Maestas, Cherie, Sarah Fulton, L. Sandy Maisel and Walter Stone. 2006. "When to Risk It? Institutions, Ambitions, and the Decision to Run for the U.S. House." *American Political Science Review* 100(2):195–208.

Mann, Thomas and Norman Ornstein. 2006. *The Broken Branch: How Congress Is Failing America and How to Get It Back on Track*. New York: Oxford University Press.

Mann, Thomas and Norman Ornstein. 2013. *It's Even Worse Than It Looks: How the American Constitutional System Collided With the New Politics of Extremism*. New York: Basic Books.

Mansbridge, Jane. 2009. "A Selection Model of Representation." *Journal of Political Philosophy* 17(4):369–398.

Manski, Charles. 2003. *Partial Identification of Probability Distributions*. New York: Springer-Verlag.

Masket, Seth and Jeffrey Lewis. 2007. "A Return to Normalcy? Revisiting the Effects of Term Limits on Competitiveness and Spending in California Assembly Elections." *State Politics and Policy Quarterly* 7(1):20–38.

Mayhew, David. 1974. *Congress: The Electoral Connection*. New Haven, CT: Yale University Press.

McGhee, Eric, Seth Masket, Boris Shor, Stephen Rogers and Nolan McCarty. 2014. "A Primary Cause of Partisanship?" *American Journal of Political Science* 58(2):337–351.

Miller, Susan, Jill Nicholson-Crotty and Sean Nicholson-Crotty. 2011. "Reexamining the Institutional Effects of Term Limits in U.S. State Legislature." *Legislative Studies Quarterly* 36(1):71–97.

Miquel, Gerard Padro I and James Snyder. 2006. "Legislative Effectiveness and Legislative Careers." *Legislative Studies Quarterly* 31(3):347–381.

Moncrief, Gary and Joel Thompson. 2001. "On the Outside Looking In: Lobbyists' Perspectives on the Effects of State Legislative Term Limits." *State Politics and Policy Quarterly* 1(4):394–411.

Mooney, Christopher. 2007. Lobbyists and Interest Groups. In *Institutional Change in American Politics: The Case of Term Limits*, ed. Karl Kurtz, Bruce Cain and Richard Niemi. Ann Arbor, MI: University of Michigan Press.

Mooney, Christopher. 2009. "Term Limits as a Boon to Legislative Scholarship: A Review." *State Politics and Policy Quarterly* 9(2):204–228.

Muthoo, Abhinay and Keneth A. Shepsle. 2010. "Information, institutions and constitutional arrangements." *Public choice* 144(1):1–36.

National Conference of State Legislatures. 2009. "Full and Part-Time Legislatures." Accessed on August 13, 2013.
URL: *http://www.ncsl.org/legislatures-elections/legislatures/full-and-part-time-legislatures.aspx*

Nokken, Timothy. 2013. "Uncertainty and Roll-Call Voting in Lame-Duck Sessions of the U.S. House, 1969-2010." *Legislative Studies Quarterly* 38(4):571–591.

North Carolina General Assembly. 2013. "How much do Legislators get paid?" Accessed on August 8, 2013.
URL: *http://www.ncleg.net/ncgainfo/educational/funfacts/legpay.html*

Poole, Keith. 1998. "Changing Minds? Not in Congress!" Working Paper.

Powell, Lynda, Richard Niemi and Michael Smith. 2007. Constituent Attention and Interest Representation. In *Institutional Change in American Politics: The Case of Term Limits*, ed. Karl Kurtz, Bruce Cain and Richard Niemi. Ann Arbor, MI: University of Michigan Press.

Powell, Richard. 2000. "The Impact of Term Limits on the Candidacy Decisions of State Legislators in U.S. House Elections." *Legislative Studies Quarterly* 25(4):645–661.

Powell, Richard. 2003. The Unintended Effects of Term Limits on the Career Paths of State Legislators. In *The Test of Time: Coping with Legislative Term Limits*, ed. John David Rausch Jr. Rick Farmer and John Green. Lanham, Maryland: Lexington Books.

Reilly, Ryan. 2013. "DOJ To Sue North Carolina Over Voter ID Law, Voting Restrictions." *The Huffington Post* . September 30, 2013.

Rohde, David. 1991. *Parties and Leaders in the Postreform House*. Chicago, Ill: University of Chicago Press.

Rohde, David. 2013. "Reflections on the Practice of Theorizing: Conditional Party Government in the Twenty-First Century." *The Journal of Politics* 75(4):849–864.

Rosenbaum, Paul and Donald Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70(1):41–55.

Rosenbaum, Paul R. 2010. *Design of Observational Studies*. New York: Springer-Verlag.

Rothenberg, Lawrence and Mitchell Sanders. 2000. "Severing the Electoral Connection: Shirking in the Contemporary Congress." *American Journal of Political Science* 44(2):310–319.

Rubin, Donald B. 1990. "Comment: Neyman (1923) and Causal Inference in Experiments and Observational Studies." *Statistical Science* 5(4):472–480.

Samples, John and Michael McDonald. 2006. *The Marketplace of Democracy: Electoral Competition and American Politics*. Washington, DC: Brookings Institution Press.

Sarbaugh-Thompson, Marjorie. 2010. "Measuring 'Term Limitedness' in U.S. Multi-State Research." *State Politics and Policy Quarterly* 10(2):199–217.

Sarbaugh-Thompson, Marjorie, Luke Thompson, Charles Elder, John Strate and Richard Elling. 2004. *The Political and Institutional Effects of Term Limits*. New York: Palgrave MacMillan.

Schlesinger, Joseph. 1966. *Ambition and Politics: Political Careers in the United States*. Chicago, IL: Rand McNally.

Schraufnagel, Scott and Karen Halperin. 2006. "Term Limits, Electoral Competition and Representational Diversity: The Case of Florida." *State Politics and Policy Quarterly* 6(4):448–462.

Sekhon, Jasjeet and Rocio Titiunik. 2012. "When Natural Experiments Are Neither Natural Nor Experiments." *American Political Science Review* 106(1):35–57.

Shor, Boris. 2013. "Polarization Trends in American State Legislatures by Chamber." Accessed on July 27, 2013.
URL: *http://americanlegislatures.com/2013/07/26/polarization-trends-in-american-state-legislatures-by-chamber/*

Shor, Boris and Nolan McCarty. 2011. "The Ideological Mapping of American Legislatures." *American Political Science Review* 105(3):530–551.

Shor, Boris and Nolan McCarty. 2013. "Individual State Legislator Shor-McCarty Ideology Data.". Accessed May 20, 2013.
URL: *http://hdl.handle.net/1902.1/21509 UNF:5:txTFz7Zl5bcjDqQsPR/i/w== Shor, Boris [Distributor] V1 [Version]*

Squire, Peverill. 1988. "Career Opportunities and Membership Stability in Legislatures." *Legislative Studies Quarterly* 13(1):65–82.

Squire, Peverill. 2007. "Measuring State Legislative Professionalism: The Squire Index Revisited." *State Politics and Policy Quarterly* 7(2):211–227.

Titiunik, Rocio. 2015. "Drawing Your Senator From a Jar: Term Length and Legislative Behavior." *Political Science Research and Methods* . Forthcoming.

Vidal, Jordi Blanes I, Mirko Draca and Christian Fons-Rosen. 2012. "Revolving Door Lobbyists." *American Economic Review* 102(7):3731–3748.

Volden, Craig and Alan Wiseman. 2009. "Legislative Effectiveness in Congress." Working Paper.

Volden, Craig and Alan Wiseman. 2013. "When Are Women More Effective Lawmakers Than Men." *American Journal of Political Science* 57(2):326–341.

Volden, Craig and Alan Wiseman. 2014. *Legislative Effectiveness in the United States Congress: The Lawmakers*. New York: Cambridge University Press.

Will, George. 1992. *Restoration: Congress, Term Limits and the Recovery of Deliberative Democracy*. New York: The Free Press.

Wooldridge, Jeffrey. 2010. *Econometric Analysis of Cross Section and Panel Data*. Second edition ed. Cambridge, MA: MIT Press.

Wright, Gerald C. 2007. "Do term limits affect legislative roll call voting? Representation, polarization, and participation." *State Politics & Policy Quarterly* 7(3):256–280.

Zucchino, David. 2013. "N.C. Legislature draws protests; Clergy, others welcome arrest at 'Moral Monday' demonstrations against GOP policies." *Los Angeles Times* . June 13, 2013.

Zupan, Mark. 1990. "The Last Period Problem in Politics: Do Congressional Representatives Not Subject to a Reelection Constraint Alter Their Voting Behavior?" *Public Choice* 65(2):167–180.