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

Andrew Feher †
Department of Political Science
University of Michigan

Rocío Titiunik‡
Department of Political Science
University of Michigan

February 26, 2015

Abstract

We study whether lame-duck status affects legislative behavior. Observational empirical research on this question has produced a large body of evidence suggesting that legislators engage in participatory shirking, 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, 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. Our study relies on the random assignment of term length that occurred in the Arkansas Senate, which in turn induced the random assignment of term limits in 1997 and 2007. Across five measures of legislative behavior—bills introduced, co-sponsorship, bills passed, resolutions and abstention rates—we find no evidence that term limited senators engage in participatory shirking. We do not observe all senators in our original experimental sample, a challenge we address using bounds.

*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.

†Ph.D. Candidate, Department of Political Science, University of Michigan, <afeher@umich.edu>.
‡Assistant Professor, Department of Political Science, University of Michigan, <titiunik@umich.edu>, http://www.umich.edu/~titiunik, Department of Political Science, University of Michigan, 5700 Haven Hall, 505 South State St., Ann Arbor, MI 48109.
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. And during the Continental 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). But 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 at least two reasons. First, from a policy point of view, it is important 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.

A large amount of empirical evidence suggests that there are dangers to severing electoral accountability. Studies of governors (Besley and Case, 1995; Alt et al., 2011), members of Congress (Rothenberg and Sanders, 2000; Herrick et al., 1994) and state legislators (Carey et al., 2006; Powell et al., 2007) have all found evidence consistent with last-period shirking.
Sarbaugh-Thompson et al. (2004), for example, report that term limited legislators turn their attention away from constituents and toward interest groups, and Besley and Case (1995) find that term limited governors are less proactive in responding to natural disasters. Interpreting these findings causally, however, requires strong assumptions.

We contribute in two ways to the study of the relationship between lame-duck status and legislative behavior. 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 readily identifiable (Sekhon and Titiunik, 2012). We show that this methodological challenge arises in many of the previous studies of last-term effects. Any study of last-term effects 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 since, for example, representatives may decide to retire preemptively when they anticipate that their poor past performance will result in a loss. As a result, it may be difficult to find a group of legislators that provides a valid counterfactual to lame-duck incumbents.

Moreover, these inferential complications do not necessarily disappear when the occurrence of the last term is not under the control of politicians but rather determined by an exogenous rule. Most previous studies rely on non-experimental research strategies such as interrupted time-series and cross-state comparisons. In particular, studies of term limits typically rely on the “exogeneity” of term limit rules, i.e. on the fact that under term limits the maximum number of terms in office is determined externally and is, by construction, unrelated to the strategic decisions of individual legislators.

However, policies that impose a limit on the number of terms that a politician may stay in office do prevent those who wish to run indefinitely for reelection from doing so,
but they fail to eliminate the *dynamic* biases that may arise as a result of electoral defeat. For example, even under externally imposed term limits, some representatives are defeated before term limits take effect. Since lower quality politicians are less likely to get reelected, representatives who are reelected the maximum allowed number of times and are therefore term-limited are, on average, stronger candidates than those who have the ability to run for reelection. Moreover, in a legislature subject to term limits, the proportion of reelection-eligible legislators who are last elected in a competitive open-seat election is much higher than the similar proportion among term-limited legislators, since the last election faced by term-limited members is, by construction, not an open seat and as a result may occur in a less competitive environment. We provide empirical evidence that these dynamic biases are of concern using data from eight different states, and suggest ways to improve this type of research designs.

These methodological obstacles are the motivation for our second contribution, which is the study of last-term effects under term limits 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 participatory shirking.

The Arkansas General Assembly is a legislature that “has been revolutionized by term limits” and is the type of legislature particularly conducive to shirking (Smith, 2007, 194). Members in these “dead-end” legislative environments face low salaries, limited staff and poor advancement prospects (Squire, 1988, 72). It seems reasonable to expect, then, that in
the absence of both a reelection incentive and little hope of continuing their political career beyond their current office, lame-duck legislators will have little incentive to remain active participants in the chamber.

Despite these expectations, we find no evidence that term-limited legislators reduce their effort by introducing, cosponsoring and 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 most conventional observational analyses of legislators’ behavior in their last terms in office, and have important policy implications for states with term limits. In approving term limits for its state legislators, voters do not appear to have magnified agency problems, and in that sense one potential cost of introducing term limits has not been realized. But the very finding that legislative behavior is not altered when reelection incentives disappear casts doubt on the assumption that elections’ main role is to serve as an accountability mechanism. If elections serve primarily as a selection mechanism instead, our findings suggest that, by adopting term limits, voters have placed themselves in the inauspicious position of restricting the set of high quality politicians that could represent them.

The remainder of the paper is organized as follows. In the next section, we 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 (Manski, 2003) to address the issue of attrition in the original experimental sample. We conclude in the last section. Additional results are presented in an online Supplemental Appendix.
2 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 an accountability model of representation, voters incorporate politicians’ past actions into their voting decisions and elections serve as an accountability mechanism that sanctions representatives’ behavior. 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 induce undesirable legislative behavior or shirking, as the threat of punishment is removed and legislators have no incentive to please the electorate.\(^1\) 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).\(^2\)

Moreover, there may be mechanisms by which the removal of reelection incentives may result 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 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) cosponsoring fewer bills (3) passing fewer bills, (4) performing less constituency service and (5) 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. 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 a “dead end” legislature 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, 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, it is natural to expect that most legislators will 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.

## 3 Research Designs to Study Last-Term Effects

A possible strategy to study last-term effects induced by term limits 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 et al., 1998, 2006; Powell et al., 2007; Sarbaugh-Thompson et al., 2004), and it is often the only one available. 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).³ 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”.

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 incum-
bents. 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 incum-
bency, including, possibly, a weaker challenger and increased name recognition. 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 might 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-
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 com-
parisons 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 elec-
tion 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 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 legislators in these states enjoy vote shares between 4-5 percentage points higher than non-term-limited legislators. In two states, term-limited legislators are also significantly less likely to face a challenger. The remaining four rows in Table 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.

[Insert Table 1 about here]

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 Section A2 of the Supplemental Appendix, the differences in vote shares persist and become even more pronounced when uncontested races are excluded from 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. Our proposed strategy 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.

Consistent with our expectations, once we exclude freshman from the group of non-term-limited legislators, uncontested rates and vote shares between the two groups become considerably more similar. Table 2 presents uncontested rates and vote shares from the same eight states in Table 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, we only examine lower chamber election results for these states. 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.

[Insert Table 2 about here]
4 A Research Design Based on Random Assignment

Motivated by the methodological challenges in observational studies of term limits effects, 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. 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 a four-year term 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 similar designs in Illinois and Texas, was used by Titiunik (2014) 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 very important element of our
research design is that two-year terms do not count against the two-term limit—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), and again in 1998 for another four-year term (1998-2002), because the first two-year term did not count towards term limits. In 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 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 1((a)) and 1((b)) 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 lame-ducks (i.e., will not see another election as state senator), while those assigned two-year terms are eligible for their last reelection in the following election.

[Insert Figures 1((a)) and 1((b)) about here]

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.\footnote{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.}

Moreover, similarly 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 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 et al., 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.

The first three columns in Table 3 present the results from balance tests from our full sample of the 64 senators described above who drew lots in October 1993 and December 2002. Both groups show strong balance across seven pretreatment covariates, including vote share obtained in the previous election, party, and race. The lowest p-value is a .10, suggesting that the randomization of terms was successfully implemented.

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, which we call “compliers”, and brings 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 original experimental sample, which in turn would render a difference-in-means comparison a biased estimator of the average treatment effect (see, e.g., Gerber and Green, 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.

[Insert Table 3 about here]

In this and further 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 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.\(^8\) Moreover, this balance in attrition levels is also seen when we consider each legislative session individually.\(^9\)

Of course, even though roughly the same number of senators drops out of our sample 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; instead, it would just be the result of endogenous attrition. 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. But we can show indirect evidence that suggests it is not likely to be a problem, which we do below, and we can also show how sensitive our conclusions are to the possibility of endogenous attrition—which we do in the following section.

The indirect evidence we present is covariate balance tests for our complier sample, i.e. the subset of senators who remained in the chamber in the 1997 and 2007 sessions. As shown in the last three columns of Table 3, we still see that balance is strong, as not one of the seven covariates reaches a p-value below the .05 level of statistical significance. The table seems to suggest that the groups in the complier subsample are systematically more dissimilar (i.e., the difference-in-means point estimates are larger in absolute value) than in the full sample. But
in Tables A3 and A4 of the Supplemental Appendix we show that this an artifact of pooling two small samples, as this pattern is not present when each cohort is analyzed individually. Moreover, in the Supplemental Appendix we also show that randomization inference-based p-values, which do not rely on distributional approximations, are larger than the p-values based on the t-distribution reported in Table 3.

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, because we cannot entirely rule out that significant differences in unobservable characteristics remain between both groups in the complier subsample, we present robustness checks based on bounds.

5 Last Term Effects in the Arkansas Senate

We now study our main question of interest, whether term-limited state senators engage in participatory shirking. To do so, we examine five dependent variables at the individual level: the number of bills introduced, the number of bills cosponsored, the number of bills passed, the abstention rate on roll-call votes and resolutions as a proxy for constituency service. While we would prefer a more conventional measure of constituency service (e.g. number of district staff, trips back to the district), they are not available. Instead, we use data on the resolutions that state senators file. 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 State Legislature 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.\textsuperscript{10} 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 State Legislature an environment where the effects of removing reelection incentives, if real, should be detected. As we discuss below, the fact that we find null effects suggests that the accountability view of elections may not entirely appropriate for this particular setting.

\textbf{Results}

All of the analyses are based on t-tests that test the null hypothesis that the difference-in-means between term-limited (those randomly assigned four-year lots) and non-term-limited senators (those randomly assigned two-year lots) is zero. As mentioned above, to maximize statistical power, we pool observations across the two natural experiments. However, as we show in Tables A3 and A4 in the Supplemental Appendix, conducting separate analyses for the 1997 and 2007 sessions returns comparable results.

Table 4 provides the results from the difference-in-means t-tests across our five dependent variables, and Figures 2 and 3 display the entire distributions using box-plots. We start by describing bill outcomes. The average number of bills introduced is remarkably similar between both groups, with term-limited senators introducing about 1.5 more bills than the non-term limited senators, a difference that is far from statistical significance (p-value 0.67). A similar result is observed regarding the mean 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, the difference does not allow us to reject the null. Figures 2(a) and 2(b) show that this mean analysis is not masking
any evidence of shirking: looking at the entire distribution, there is no evidence that the
term-limited group engages in participatory shirking according to these measures.

[Insert Table 4 about here]

In the third row of Table 4 we show results for bills cosponsored, an outcome for which
we only have data for the 2007 session. There is a roughly 4-bill difference between term-
limited and non-term-limited state senators, as the former cosponsored about 36 pieces of
legislation while the latter cosponsored about 40 pieces of legislation. Again, this difference
is not statistically significant. Turning to our measure of constituency service, term-limited
senators average a little over two resolutions while non-term-limited senators file just over
1.5 resolutions, a difference that falls short of statistical significance.

[Insert Figure 2 about here]

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, yielding yet another insignificant difference, but that disparity in the mean point estimate 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 4 and Figures 3 indicate, when we drop him from the analysis, all differences in abstention between both groups vanish almost entirely.

[Insert Figure 3 about here]

In sum, we find no evidence of last-term effects on five measures of participatory shirking. Importantly, for four of these measures, bills introduced, bills passed, resolutions filed and abstention rates, our inability to detect an effect is not driven by the potential low statistical power of our relatively small sample. 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 confirm that there is no evidence of shirking. The situation is different for the cosponsorship outcome, since in this
case the term-limited group does engage in less cosponsorships, and we cannot rule out that low statistical power is the reason behind our inability to reject the null, particularly given that we only have data for one session and our sample size is reduced from 49 to 26. But even if this cosponsorship effect were real, considering the lack of last-term effects in the other outcomes, we would still conclude that the evidence of last-term effects on participatory shirking is weak at best.

**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 each of our five measures of legislative behavior among our treated 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 among our control complier sample. 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 among the observed treated and missing outcomes in the control group at the 75th percentile value among the observed controls.

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

Table 5 reports the bounds on the average treatment effect of term limits on legislative behavior across the five 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 5 show that for bill introductions, bill passage, resolutions and abstention rates, even a severely endogenous pattern of attrition would result in small shirking 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 is 0.44%, 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. The 95% confidence intervals on
the estimated bounds interval are naturally consistent with larger negative effects, which is expected 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 seems robust to even severe endogenous attrition.

The same cannot be said for cosponsorship. The last-term effect on cosponsorship has an upper bound of -1.81, essentially ruling out a non-negative effect if we ignore sampling variability (accounting for sampling variability takes the upper bound to 3.36). This is consistent with what we saw in the previous section.

[Insert Table 5 about here]

6 Conclusion

We have examined how the absence of reelection incentives affects legislative behavior, specifically the extent to which term-limited legislators engage in participatory shirking. While many scholars have studied last-term effects, ours is the first to experimentally examine the question. Our experimental design overcomes both the static and dynamic methodological challenges that the nonrandom assignment of legislators to lame-duck status has presented previous scholars. Leveraging two natural 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 stark contrast to numerous studies that have uncovered last-term effects.

Although are 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 when electoral accountability is removed casts some doubts on whether the sanctioning model appropriately describes elections’ primary function. 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 pursue the voters’ desired policies 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 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 professionalization 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 “dead-end” legislatures that currently have term limits, which represents almost half of the total number of states with term limits.11
References


Notes

1 See Mansbridge (2009) for a discussion of the sanctioning model and Barro (1973) and Ferejohn (1986) for classical models on the control of politicians via reelection incentives.

2 We consider critiques to the accountability model (e.g., Fearon, 1999; Mansbridge, 2009) in the conclusion.

3 The result that more competent legislators are more likely to survive reelection is standard in agency models of electoral selection (see, for example, Alt et al., 2011, and references therein).

4 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.

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

6 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.

7 Due to the aforementioned 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.

8 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).

9 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.

10 Personal communication with Arkansas senate staff.

11 In addition to Arkansas, the other dead-end legislatures include Arizona, Louisiana, Maine, Montana,
Nebraska and South Dakota (Maestas, 2000).
Table 1: Uncontested Rates and Difference-in-means for Vote Share Between Term-Limited and Non-Term-Limited State Legislators

<table>
<thead>
<tr>
<th>States with Consecutive Bans</th>
<th>Vote Share</th>
<th>Uncontested Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean NTL</td>
<td>Mean TL</td>
</tr>
<tr>
<td>Arizona</td>
<td>49.45</td>
<td>53.74</td>
</tr>
<tr>
<td>Colorado</td>
<td>67.41</td>
<td>71.51</td>
</tr>
<tr>
<td>Ohio</td>
<td>68.19</td>
<td>72.60</td>
</tr>
<tr>
<td>S. Dakota</td>
<td>47.20</td>
<td>52.02</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>States with Lifetime Bans</th>
<th>Vote Share</th>
<th>Uncontested Rates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean NTL</td>
<td>Mean TL</td>
</tr>
<tr>
<td>Arkansas (House)</td>
<td>83.45</td>
<td>93.92</td>
</tr>
<tr>
<td>California</td>
<td>65.73</td>
<td>67.92</td>
</tr>
<tr>
<td>Michigan</td>
<td>66.71</td>
<td>68.31</td>
</tr>
<tr>
<td>Oregon</td>
<td>66.54</td>
<td>71.85</td>
</tr>
</tbody>
</table>

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. Reported p-values are from two-tailed t-tests of null hypothesis that means are equal.
Table 2: Uncontested Rates and Difference-in-means for Vote Share Between Term-Limited and Non-Term-Limited State Legislators, Excluding Freshman Legislators

| States with Consecutive Bans | Vote Share | | Uncontested Race | Mean NTL | Mean TL | p-val | Mean NTL | Mean TL | p-val | N |
|-----------------------------|------------|------------------|------------------|------------|------------------|------------------|
| 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. Reported p-values are from two-tailed t-tests of null hypothesis that means are equal.
Figure 1: Illustration of Research Design

(a) 1990s Design

<table>
<thead>
<tr>
<th>Year</th>
<th>Treatment</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>1992</td>
<td>79th</td>
<td>80th</td>
</tr>
<tr>
<td>1994</td>
<td>80th</td>
<td>81st</td>
</tr>
<tr>
<td>1998</td>
<td>82nd</td>
<td>83rd</td>
</tr>
</tbody>
</table>

2-year term 4-year term 4-year term

(b) 2000s Design

<table>
<thead>
<tr>
<th>Year</th>
<th>Treatment</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>2002</td>
<td>84th</td>
<td>85th</td>
</tr>
<tr>
<td>2004</td>
<td>85th</td>
<td>86th</td>
</tr>
<tr>
<td>2008</td>
<td>86th</td>
<td>87th</td>
</tr>
<tr>
<td>2012</td>
<td>87th</td>
<td>88th</td>
</tr>
</tbody>
</table>

2-year term 4-year term 4-year term
Table 3: Covariate Balance Between Term-Limited and Non-Term-Limited Arkansas Senators, 1992 and 2002 cohorts

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TL</td>
<td>Non-TL</td>
<td>p-val</td>
</tr>
<tr>
<td>Vote share</td>
<td>89.27</td>
<td>86.80</td>
<td>.61</td>
</tr>
<tr>
<td>Married</td>
<td>.88</td>
<td>.88</td>
<td>1.00</td>
</tr>
<tr>
<td>Male</td>
<td>.94</td>
<td>.81</td>
<td>.13</td>
</tr>
<tr>
<td>Democrat</td>
<td>.91</td>
<td>.75</td>
<td>.10</td>
</tr>
<tr>
<td>Black</td>
<td>.13</td>
<td>.06</td>
<td>.39</td>
</tr>
<tr>
<td>Attorney</td>
<td>.25</td>
<td>.34</td>
<td>.42</td>
</tr>
<tr>
<td>Age</td>
<td>50.66</td>
<td>52.78</td>
<td>.44</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TL</td>
<td>Non-TL</td>
<td>p-val</td>
</tr>
<tr>
<td>Sample Size</td>
<td>32</td>
<td>32</td>
<td></td>
</tr>
</tbody>
</table>

Column ‘TL’ reports the mean outcome for those senators assigned a four-year lot (in 1992 or 2002) who are therefore term-limited, and the ‘Non-TL’ reports the mean outcome for those senators assigned a two-year lot who are therefore not yet term-limited.
Table 4: Effects of Term Limits on Legislative Behavior in Arkansas Senate, 1992 and 2002 cohorts

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Non-Term-limited</th>
<th>Term-limited</th>
<th>p-value</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bill introductions</td>
<td>21.48</td>
<td>22.88</td>
<td>.67</td>
<td>[-5.28, 8.10]</td>
</tr>
<tr>
<td>Bills passed</td>
<td>13.35</td>
<td>14.58</td>
<td>.59</td>
<td>[-3.29, 5.74]</td>
</tr>
<tr>
<td>Bills cosponsored</td>
<td>40.85</td>
<td>36.69</td>
<td>.40</td>
<td>[-14.21, 5.90]</td>
</tr>
<tr>
<td>Resolutions</td>
<td>1.57</td>
<td>2.23</td>
<td>.21</td>
<td>[-.39, 1.72]</td>
</tr>
<tr>
<td>Abstentions</td>
<td>1.14</td>
<td>2.36</td>
<td>.34</td>
<td>[-1.34, 3.78]</td>
</tr>
<tr>
<td>Abstentions w/o Steve Faris</td>
<td>1.14</td>
<td>1.21</td>
<td>.84</td>
<td>[-.69, .83]</td>
</tr>
</tbody>
</table>

Sample size 23 26

Note: The non-term-limited column corresponds to the mean outcome for those senators assigned a two-year lot (in 1992 or 2002), and the term-limited column corresponds to the mean outcome for those senators assigned a four-year lot. The reported p-value is from a two-tailed t-test. Cosponsorship only includes 26 observations from 2007—13 senators in the control group and 13 senators in the treated group—as we do not have cosponsorship data from 1997.
Figure 2: Last term effects on bill introduction, passage, cosponsorship and symbolic bills in Arkansas Senate – 81st (1997-1998) and 86th (2007-2008) Legislative Sessions
Figure 3: Last term effects on abstention rates in Arkansas Senate – 81st (1997-1998) and 86th (2007-2008) Legislative Sessions
Table 5: 75th and 25th Percentile Bounds on Last-Term Effects in Arkansas Senate, 1992 and 2002 cohorts

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>ATE Lower Bound</th>
<th>ATE Upper Bound</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bills introduced</td>
<td>-2.75</td>
<td>5.22</td>
<td>[-5.85, 8.28]</td>
</tr>
<tr>
<td>Bills passed</td>
<td>-1.78</td>
<td>3.84</td>
<td>[-3.89, 5.91]</td>
</tr>
<tr>
<td>Bills cosponsored</td>
<td>-5.75</td>
<td>-1.81</td>
<td>[-10.81, 3.36]</td>
</tr>
<tr>
<td>Resolutions</td>
<td>0.03</td>
<td>0.97</td>
<td>[-0.45, 1.43]</td>
</tr>
<tr>
<td>Abstention rate (%)</td>
<td>-0.28</td>
<td>0.44</td>
<td>[-0.63, 0.78]</td>
</tr>
</tbody>
</table>

Note: Calculations for bills introduced and passed include all 64 observations. Calculations for abstention rates exclude Steve Faris, so total sample size is 63. Calculations for cosponsorship only include the 29 observations from the 2000 cohort. 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. 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.