Disentangling Accountability and Competence in Elections: Evidence from U.S. Term Limits

Share Your Story

DISENTANGLING ACCOUNTABILITY AND COMPETENCE IN ELECTIONS:

EVIDENCE FROM U.S. TERM LIMITS[1]

James Alt

Frank G. Thomson Professor of Government
Harvard University
1737 Cambridge Street, K307
Cambridge, MA 02138
jalt@gov.harvard.edu

Ethan Bueno de Mesquita

Associate Professor
Harris School of Public Policy Studies
The University of Chicago
1155 E. 60th St.
Chicago, IL 60637
bdm@uchicago.edu

Shanna Rose

Assistant Professor
Wagner School of Public Service
New York University
295 Lafayette St, 2nd Floor
New York, NY 10012
shanna.rose@nyu.edu

Abstract

We exploit variation in U.S. gubernatorial term limits across states and time to empirically estimate two separate effects of elections on government performance. Holding tenure in office constant, differences in performance by reelection-eligible and term-limited incumbents identify an *accountability effect:* reelection-eligible governors have greater incentives to exert costly effort on behalf of voters. Holding term-limit status constant, differences in performance by incumbents in different terms identify a *competence effect:* later-term incumbents are more likely to be competent both because they have survived reelection and because they have experience in office. We show that economic growth is higher and taxes, spending, and borrowing costs are lower under reelection-eligible incumbents than under term-limited incumbents (accountability), and under reelected incumbents than under first-term incumbents (competence), all else equal. In addition to improving our understanding of the role of elections in representative democracy, these findings resolve an empirical puzzle about the disappearance of the effect of term limits on gubernatorial performance over time.

Keywords: accountability, sanctioning, competence, selection, term limits, elections

Elections play two potential roles in representative democracy. First, elections may mitigate moral hazard by creating *accountability*; that is, politicians may take costly actions on behalf of voters because they know that they will only be reelected if their performance exceeds some standard (e.g., Barro 1973; Ferejohn 1986). Second, elections may mitigate adverse selection by allowing voters to select *competent types* who perform better, in expectation, than an unknown challenger. Moreover, in the absence of term limits, elections allow voters to retain incumbents whose competence has increased through experience (Padro i Miquel and Snyder 2006). Thus, over time elections may help voters weed out bad types and retain good types (e.g., Zaller 1998; Fearon 1999; Ashworth 2005; Gordon, Huber, and Landa 2007; Ashworth and Bueno de Mesquita 2008; Gowrisankaran, Mitchell, and Moro 2008).[2]

Empirically distinguishing the accountability and competence effects of elections has proven difficult. While Fearon (1999) emphasizes a tradeoff between accountability and competence (or selecting good types), in many theoretical models, both effects operate in the same direction (Banks and Sundaram 1998; Duggan 2000; Ashworth 2005; Ashworth and Bueno de Mesquita 2006; Besley 2006). Voters' threat to reelect only incumbents believed to be good types gives politicians an incentive to exert effort in order to try to convince voters that they are "good" (that is, more competent than they really are). For this reason, behavior by voters that alleviates adverse selection simultaneously alleviates moral hazard.

We devise an empirical strategy to isolate the two effects by exploiting variation in the length of gubernatorial term limits. The basic argument is as follows. The relative performance of incumbents in the same term, some of whom are eligible to run again and some of whom are not, reflects the accountability effect, since each has survived the same number of elections, but has different incentives to take costly action on behalf of voters. The relative performance of term-limited

incumbents in different terms reflects the competence effect, since each has been reelected a different number of times but has the same incentive to take costly action on behalf of voters.

Our data are from the American states, which have witnessed a major change in gubernatorial term limit laws over the last half century: the gradual shift from one-term limits to two-term limits in nearly one-third of the states. During this same period, other states had two-term limits or no term limits. This variation in the length of term limits across states and time allows us to compare economic and fiscal performance under first-term, term-limited incumbents; first-term, reelection-eligible incumbents; second-term, term-limited incumbents; and second-term, reelection-eligible incumbents.

In regression models with state and year fixed effects, state time trends, and a variety of economic and political controls, we find that economic growth is higher and taxes, spending, and borrowing costs are lower under reelection-eligible incumbents than under term-limited incumbents, holding tenure in office constant (evidence of an accountability effect), and under second-term incumbents than under first-term incumbents, holding term-limit status constant (evidence of a competence effect). We find that these two effects are of comparable magnitudes.

Empirically distinguishing the accountability and competence effects is important for several reasons. First, it improves our understanding of the mechanisms through which elections affect governance. Second, such estimates are critical for thinking about questions of institutional design. For instance, knowing the magnitudes of these two effects would allow institutional designers to better assess whether they should focus on minimizing adverse selection or moral hazard. Third, our estimates help resolve an extant puzzle in the empirical literature on term limits. Term-limited governors in the middle of the twentieth century taxed and spent more than governors who were eligible for reelection; however, this effect has gradually disappeared over time (Besley and Case 2003). Our estimates suggest that when the states relaxed their term limit laws to allow governors to serve a second term in office, voters were able to use elections to weed out less competent incumbents,

thereby mitigating the negative accountability effect of term limits.   Finally, our particular identification strategy also lends some insight into ways in which term limits may affect voter welfare.

Literature Review

The majority of the empirical literature on the role of elections in representative democracy has focused primarily on moral hazard (accountability) to the exclusion of adverse selection (competence). For instance, a sizeable literature investigates whether members of Congress "shirk" in their final terms prior to retirement, where shirking is defined either in terms of reduced effort (i.e., voting less frequently) or ideological congruence with voters (i.e., voting more "sincerely"). Vanbeek (1991) and Lott and Bronars (1993) find that retiring members vote less frequently, but do not find statistically significant last-period effects in the content of members' votes, while Figlio (1995), Tien (2001), and Snyder and Ting (2003) find evidence that members vote more sincerely in the final term. Similarly, McArthur and Marks (1988) find that, in post-election sessions, members who have not been reelected vote systematically differently than those who have been reelected. Ferraz and Finan (2009) investigate last-period effects among Brazilian mayors, and find that those who are eligible for reelection engage in less corruption, on average, than do term-limited mayors.

More recently, a handful of studies have used institutional variation to separate accountability and competence effects. Studying a panel of countries between 1972 and 1990, Johnson and Crain (2004) find that the size of government has expanded more rapidly in countries with one-term limits than in countries with two-term limits, consistent with a competence effect. Gordon and Huber (2007) find that judges in partisan competitive systems sentence more punitively than those in retention systems. Using variation in judges' electoral calendars, they show that these institutional effects are primarily due to changes in judges' incentives, though there is also a less pronounced effect on the selection of judges to office. Gagliarducci and Nannicini (2009) use a combination of term limits and institutional randomness in wages paid to Italian mayors to disentangle accountability and competence.

They find that most of the increased performance associated with higher wages can be attributed to the selection of more competent politicians, rather than increased incentives for reelection.

Our paper contributes to this more recent literature by using institutional variation in U.S. gubernatorial term limits to isolate the accountability and competence effects. In the most widely-cited paper on gubernatorial term limits, Besley and Case (1995a) find evidence of electoral accountability: in the American states between 1950 and 1986, per capita spending and taxes were higher under term-limited governors. However, in a 2003 paper, the authors repeat their analysis on a sample that extends through the mid-1990s and find that the effect of term limits has gradually disappeared over time. The authors conclude that "it seems likely that some omitted variable is responsible for the change in behavior observed for governors working under a term limit. This is an area ripe for future research" (Besley and Case 2003, 55).[3]

We conjecture that the omitted variable that accounts for this puzzling result is a competence effect arising from variation in the length of gubernatorial term limits. In the early postwar years, the majority of term-limited governors served under "no succession" laws limiting them to a single term; today Virginia alone retains the practice. Despite this widespread shift from one- to two-term limits, Besley and Case do not distinguish between term limits of different lengths. As with most other studies, they focus exclusively on accountability, simply including a dummy variable for term-limited governors in their regressions.[4]

Accountability, Competence, and Term Limits: A Model of the Empirical Strategy

The existence of accountability and competence effects in elections is well established in the theoretical literature (e.g., Banks and Sundaram 1998; Fearon 1999; Duggan 2000; Persson and Tabellini 2000; Ashworth 2005; Besley 2006; Banks and Duggan 2008; Besley and Smart 2007; Gordon, Huber, and Landa 2007; Ashworth and Bueno de Mesquita 2008).[5] As such, the purpose of our formal model is not to generate predictions that these effects exist. Rather, our model is intended to

demonstrate that, assuming these two effects exist, our empirical strategy will measure their magnitudes. Showing this is important because the logic of which comparisons isolate which effects is not always straightforward. Indeed, in some circumstances comparisons that, at an intuitive level, seem to isolate one effect or the other turn out not to do so.

Given this goal, the next section provides a nontechnical overview of a simple agency model of elections and lists assumptions that ensure that both the accountability and competence effects exist in the model. We show, within the model, conditions under which term limits allow us to make comparisons that separately identify the magnitudes of those two effects in the data. Formal analysis and proofs of all results are given in the Appendix.

The Model

Consider an infinitely-repeated game in which there are two kinds of players: politicians and a representative voter. The order of play is as follows. At the beginning of each period, the incumbent politician chooses a level of effort $a \in \{\underline{a}, \overline{a}\}$. At the end of each period—after observing a policy outcome, but not the incumbent's level of effort—the voter selects between two candidates (one of whom may be the incumbent).

A t-period term limit means that a politician can only serve in office for t periods. A politician is *term-limited* or a *lame duck* when she can only serve one more period. If a politician is *eligible for reelection*, she runs against a randomly selected challenger (described below). If she is term-limited, there is an open seat election between two randomly selected challengers.

At the beginning of a given period, the incumbent and challenger (or two challengers) can be either of two types $\{\theta_I, \theta_C\}$, for *incompetent* and *competent*, respectively. The probability that a randomly drawn politician is competent is μ_0. A politician's type is private information.

Outcomes, for simplicity, are either good or bad. We interpret the good outcome H as the politician successfully managing economic or fiscal policy in a given year. We assume that

incompetent politicians never achieve this and that competent politicians who choose high effort do so with certainty, while competent politicians choosing low effort achieve H with probability $\gamma < 1$, so that effort (or foregone rents) increases the probability of a good outcome for competent types. We also assume that the voter prefers an incumbent who is competent for certain but who exerts low effort to a randomly drawn challenger who exerts high effort, that is, $\gamma > \mu_0$. If this condition did not hold, the voter would always prefer a challenger and there would be no possibility of selection on competence in a two-term-limit system. This assumption allows us to model elections with both accountability and competence in the most parsimonious way.

There may also be idiosyncratic changes in the strategic environment that cause a previously effective incumbent to become ineffective for reasons beyond her control. The chance that a competent incumbent transitions to incompetent in her t^{th} term is ε^t. The transition occurs after the effort choice but before the outcome is determined.[6]

At the beginning of the game, Nature chooses the cost of high effort, c, and makes it common knowledge. This cost has to be low enough, relative to the benefits of office, so that a competent politician would be willing to exert high effort in exchange for a promise of indefinite reelection as long as she remained competent. (See the upper bound on the support of the distribution of costs in the Appendix.) Without loss of generality, we normalize the cost of low effort to zero. For a benefit of reelection $B > 0$, the politician's payoff from any period in which she is in office is $B - c(a)$, where $c(\overline{a}) = c$ and $c(\underline{a}) = 0$. Politicians discount the future by δ and voters are myopic, caring only about the next period.

We solve the game for pure strategy Perfect Bayesian Equilibria in Markov Strategies (Maskin and Tirole 2001), which we simply refer to as *equilibria*. The solution concept requires that equilibria be stationary in the payoff-relevant state variables: beliefs about ability and terms remaining. We

compare the equilibria under different term limit rules to establish counterfactuals for our empirical analysis.

Evidently, in the absence of any chance of reelection, effort imposes costs without offering any benefits for the politician. Thus, a politician who can only serve one term will choose $\underline{a}$ in that term. Hence the following result (denoted **Proposition 1** in the Appendix): In the unique equilibrium of the game with a one-term limit, the incumbent chooses $\underline{a}$ regardless of type.

Now suppose a politician can serve at most two terms (but that, if not reelected, can never run again). Just as in the one-term-limit case, it is clear that when she is term-limited in her second term, an incumbent will choose $\underline{a}$, regardless of type. The question is, can $\bar{a}$ be sustained in equilibrium in the first term?

The voter only reelects incumbents whom he believes are likely to be competent. He updates his beliefs about competence after observing performance since, as shown in the appendix, better performance suggests that the incumbent is more likely to be competent. Thus, the benefit to the incumbent of effort is that it makes it more likely that she will achieve a good outcome, convince the voter she is competent, and thereby achieve reelection. If, however, the costs of effort are too high, the benefits of office are too low, or the increase in probability of providing a good outcome that is achieved through high effort is too low, then the politician will choose not to invest in a high level of effort, even in her first term. **Proposition 2** in the Appendix formalizes these intuitions. In any equilibrium of the game with two-term limits, the voter reelects an incumbent if and only if the outcome is *H* and a first-term incumbent exerts high effort if and only if the costs of effort lie below a critical value (defined as a function of benefits of holding office, discounting, priors on competence, and the transition rate).

This result implies that first-term outcomes are better for the voter under two-term limits than under one-term limits due to an accountability effect, since reelection-eligible, competent incumbents

exert positive effort if the cost is not too high. Moreover, expected outcomes are also better under second-term term-limited incumbents than under first-term term-limited incumbents due to a competence effect: although no term-limited incumbent has an incentive to exert high effort, only competent types survive the first election and, moreover, due to experience, they are less likely to transition away from competence.

Finally, the absence of term limits enhances the incentive effect even further, since it provides the incumbent with a longer time horizon. Unlike the case with a two-term limit, the voter is now able to provide second- (and later-) term politicians with electoral incentives. Moreover, competence effects continue to operate. **Proposition 3** in the Appendix solves for an equilibrium of the game with no term limits. Again, voters reelect if and only if the outcome is H. In each term, the incumbent chooses high effort if and only if the costs of effort lie below a critical value (now defined as a function of terms served, benefits of holding office, discounting, priors on competence, and the transition rate). First-term incumbents are willing to exert effort for a wider range of costs under no term limits than under two-term limits. This is because the benefit of winning reelection includes not only the immediate payoff but also the possibility of future terms when there are no term limits. Moreover, the cost threshold for high effort is declining with tenure in office. Hence, second-term (and later) incumbents are even more willing to exert effort.

Comparisons across Systems

Recall that only competent incumbents can achieve the outcome H and that they are certain to achieve this outcome if they exert effort. Given this, expected performance in any period is:

Pr(Competent)[Pr(High Effort) + Pr(Low Effort)Pr(*H* with Low Effort)]

Using this formula, we can determine whether differences in expected performance across term-limit systems are due to differences in the probability the incumbent exerts high effort (an accountability effect) or differences in the probability the incumbent is competent (a competence effect).

Table 1 summarizes results that relate to comparisons across one- and two-term-limit systems, which we formalize in **Proposition 4** in the Appendix**.** The basic idea is this**:**

1. *Accountability Effect:* The expected performance of first-term incumbents is lower in a one-term-limit system than in a two-term-limit system.

2. *Competence Effect:* The expected performance of first-term incumbents in a one-term-limit system is lower than the expected performance of second-term incumbents in a two-term-limit system.

Since with a one-term limit no incumbent chooses to exert effort (we can think of this as the cost threshold being zero), the only possibility for a good outcome comes from a competent incumbent who does not choose to exert effort. Hence, the expected performance is: $\mu_0 (1 - \varepsilon)\gamma.$

In the first term under a two-term limit, an incumbent will exert effort if the costs are low enough. Hence, expected performance is: $\mu_0 (1 - \varepsilon)\big(F(\hat{c}) + (1 - F(\hat{c}))\gamma\big)$. The extra term is positive, so expected performance is better in the first term of a two-term-limit system than of a one-term-limit system. The difference between them appears in Table 1 as the accountability effect, since the difference in performance between term-limited and eligible incumbents in their first term is entirely because the former, lacking electoral incentives, is less likely to exert effort.

In the second term under a two-term-limit system, an incumbent will not exert effort, just as in the first term of a one-term-limit system. However, only competent types survive to a second term, so the probability that an incumbent is competent in the second term is $1 - \varepsilon^2$. Expected performance is $(1 - \varepsilon^2)\gamma.$ Since $1 - \varepsilon^2 > \mu_0 (1 - \varepsilon)$, second-term incumbents are more likely to be competent than first-

term incumbents. As shown in Table 1, expected performance in the second term of a two-term-limit system is higher than in the first term of a one-term-limit system due entirely to a competence effect.

[Table 1 about here]

The second set of results relates to comparisons across two-term-limit and no-term-limit systems, which we formalize in **Proposition 5** in the appendix. The basic idea is the following**:**

1. *Accountability Effect:* The expected performance of second-term incumbents is lower in a two-term-limit system than in a no-term-limit system.

2. *Combined Accountability and Competence Effects:* The expected performance of second-term incumbents in a no-term-limit system is higher than the expected performance of first-term incumbents in either term-limit system.

As we have already seen, expected performance in the two-term-limit system is $(1 - \varepsilon^2)\gamma$. In the no-term-limit system, incumbents have the same expected ability but now have incentives to exert effort, so expected performance is $1 - \varepsilon^2$. As recorded in Table 2, the difference between the two reflects this accountability effect.

Comparing the expected performance of second-term incumbents in the no-term-limit system to first-term incumbents in either system is somewhat more complicated. First, consider first-term incumbents in the no-term-limit system. These incumbents will exert effort but, since they have not survived a round of reelection, their expected ability (which, here, equals expected performance) is only $\mu_0 (1 - \varepsilon)$. Hence, the difference between second-term incumbents and first-term incumbents within a no-term-limit system is due entirely to a competence effect (as reported in Table 2).

Now consider first-term incumbents in the two-term-limit system. These incumbents are also of expected ability $\mu_0 (1 - \varepsilon)$. Moreover, since they can only stand for reelection once, they have weaker incentives to exert effort (only doing so if the costs are below the threshold $\hat{c}$). Hence, their expected

10

performance is $\mu_0\left(1-\varepsilon\right)\left(F\left(\hat{c}\right)+\left(1-F\left(\hat{c}\right)\right)\gamma\right)$. Given this, the difference between the performance of

second-term incumbents in a no-term-limit system and first-term incumbents in a two-term-limit

system reflects a combination of accountability and competence effects (as reported in Table 2)

[Table 2 about here]

It is worth noting that what we have called the competence effect actually consists of two

analytically distinct components—selection and experience. Later-term incumbents are more

competent than earlier-term incumbents for two reasons: reelection weeds out incompetent types and

the more experience an incumbent has, the more likely she is to remain competent in the face of

stochastic elements of the political environment. Our results suggest an empirical strategy for

estimating the accountability and competence effects, but not for disaggregating the competence effect

into its constituent parts.

Data and Methodology

Propositions 4 and 5 suggest an empirical strategy for estimating our quantities of interest, the

magnitudes of the accountability and competence effects of elections. We use panel data from the

American states.[7] The sample period is 1950-2000 unless otherwise noted. We adopt a specification

similar to those of Besley and Case (1995a), Besley and Case (2003), and Besley (2006) to show how

our results relate to theirs. All monetary values are reported in constant 1982 dollars.

Dependent Variables

In the existing political agency literature (e.g. Besley and Case 1995a, Persson and Tabellini

2000), incumbents who exert more effort provide services to their constituents at lower cost, as

measured by the size of the public sector. This idea is consistent with the empirical finding that voters

are "fiscal conservatives" and punish incumbent governors for rapid spending and tax growth (e.g.

Peltzman 1992, Besley and Case 1995b). In order to integrate our study with this literature, we use the

log of per capita spending and the log of per capita taxes as dependent variables. The former captures

11

the overall size of the public sector while the latter captures what is arguably the most visible and least politically popular way in which governments raise revenue. Since tax revenue excludes federal grants, user charges, and other non-tax forms of government income, it is quite distinct from and considerably smaller than total expenditures, as shown in Table 3.[8]

In addition to following the existing literature in using these two measures, we propose two additional measures of incumbent performance. The first is the state's financial condition, as measured by relative yield spreads on 20-year general obligation debt.[9] A higher yield spread increases the cost of borrowing and can reflect rising debt, poor economic performance, mismanagement, corruption, and other factors that bear on the state's ability to make its debt service payments (Lowry and Alt 2001). Since higher interest costs translate into higher general fund expenditures that provide no public benefits, they represent poor performance that could affect re-election prospects.

Our second additional measure is the state's economic growth rate relative to the national average. Studies of gubernatorial elections and job approval ratings suggest that voters value and respond to growth (Lowry, Alt, and Ferree 1998) and, in particular, that voters evaluate their governors based on their state's economic performance relative to the performance of the national economy (see for example Wolfers 2002). Data on state economic growth are from the U.S. Department of Commerce's Bureau of Economic Analysis. Descriptive statistics for all four dependent variables appear in Table 3.[10]

[Table 3 about here]

Independent Variables and Controls

Our main independent variables of interest are dummies for the tenure and term limit status of the incumbent governor. We distinguish "lame ducks" who cannot run for re-election from "eligibles" who can, according to whether they are serving in their "first term" or "second term."

We also include a variety of economic and political controls. Like Besley and Case (1995a, 2003), we control for variables that might be expected to affect economic policy outcomes, namely: population, per capita state income, and the proportions of elderly (65+) and school-aged (5-17) in the population. In addition, we follow Alt and Lowry (2000) and Besley (2006) in controlling for the party composition of government, including dummies for Democratic governor, Democratic house, Democratic senate, and divided government, as well as a folded index of legislative party competition (i.e. the absolute value of deviations from 50 percent Democratic control) in each chamber.

The term-limit regime determines the number of terms a successful candidate could serve and thus potentially the value of office. A change in the term-limit regime could therefore affect the average competence of candidates. In that case we confound the competence effects of term limits with pre-existing differences in average quality. In order to control for this possibility, we include a variable reflecting the governor's years of political experience prior to becoming governor, which is a standard proxy for a politician's quality, measured at the time of entering office (Bond et al. 1985; Jacobson 1989; Van Dunk 1997; Hall and Bonneau 2006).[11]

Finally, we include state fixed effects, year fixed effects, and state-specific linear time trends to address any remaining unobserved heterogeneity. Adding state-specific time trends allows "treatment" and "control" states to follow different trends, helping to soak up residual variance attributable to unobservable, time-variant state characteristics (Angrist and Pischke 2009).

Comparing One-Term and Two-Term Limit Systems

The quantities of interest are the magnitudes of the accountability and competence effects. Our first empirical specification estimates these effects—as suggested by Proposition 4—by running the following regression on the sample of states that changed from one- to two-term limits:

$performance_{it} = \alpha_1 first\text{-}term\text{-}eligible_{it} + \alpha_2 second\text{-}term\text{-}lame\text{-}duck_{it} + \alpha_3 X_{it} + \theta_i + \delta_t + \upsilon_{it}$

where performance is as defined above (alternately: log of per capita spending, log of per capita taxes, borrowing costs, and economic growth), X_{it} are control variables (including state-specific time trends), θ_i are state fixed effects, δ_t are year fixed effects, and v_{it} is an error term. This sample of states is disproportionately Southern; we address this issue with a robustness test later in the paper.

The omitted category is first-term lame ducks. As in Table 1, it follows that α_1, the coefficient on *first-term-eligible*, provides an estimate of the accountability effect, since it compares the performance of first-term eligibles to that of first-term lame ducks. In terms of the formal model summarized in Table 1, α_1 estimates the quantity $\mu_0 F(\hat{c})(1-\gamma)$. The coefficient on *second-term-lame-duck*, α_2, provides an estimate of the competence effect, since it compares reelected lame ducks to first-term lame ducks. Since performance is expected to be worst under first-term lame ducks, we expect both of these coefficients to be negative for spending, taxes, and borrowing costs and positive for economic growth.

The odd-numbered columns of Table 4 report the results of these regressions. Robust standard errors are reported in parentheses. The first and third columns show that per capita spending and taxes are three to five percent lower under both first-term eligible governors and second-term lame ducks than under first-term lame ducks, supporting the accountability and competence effects, respectively. This difference, the estimated effect of elections, is of a similar magnitude (but opposite in sign) to the difference made by having a Democratic majority in the state legislature. Column 5 shows that borrowing costs are six to seven basis points lower under both first-term eligible governors and second-term lame ducks, compared to first-term lame ducks. This is similar to the effect of an extra $300 to $400 in real state per capita income. As shown in column 7, the economic growth rate is nearly 0.7 percentage points higher (about a quarter of the average growth rate) under first-term reelection-eligible governors than under first-term lame ducks, reflecting the accountability effect; the positive coefficient on the competence effect goes in the expected direction but falls short of statistical

significance at conventional levels. These are, on the whole, both substantively and statistically significant effects.

It is worth noting that, in all four cases, the competence and accountability effects are of approximately the same size; in no case does a statistical test reject the hypothesis of equal magnitudes at conventional levels of significance. This fact should not be interpreted to mean that, in some general sense, competence and accountability are of equal importance for the quality of governance. We are estimating the size of a particular accountability effect and a particular competence effect in a particular electoral setting. However, the similarity of the magnitudes in these data is important because it suggests that, by focusing on both accountability and competence, our findings can resolve an extant empirical puzzle.

[Table 4 about here]

Resolving an Empirical Puzzle

As described in the literature review, Besley and Case (1995a) find that per capita taxes and spending were higher under term-limited governors than under eligible governors between 1950 and 1986. However, in a 2003 paper the authors find that the effect of term limits on spending and taxes displays a marked downward trend over the past fifty years. Besley and Case do not distinguish first- and second-term lame ducks; their regression models include a dummy variable for all governors who cannot run for reelection. They conjecture that some unobserved factor has altered gubernatorial behavior over time.

Our results suggest that the changing effect of gubernatorial term limits reflects changes in governors' competence rather than their behavior. As states have gradually switched from one- to two-term limits, voters have increasingly been able to use elections to weed out low-quality incumbents and incumbents have had increased scope for on-the-job learning. As average tenure has increased, performance by term-limited governors has increasingly reflected the effect of greater incumbent

15

competence, offsetting the effect of lower effort over time. Since the estimated competence effect is roughly the same size as the accountability effect, the shift from one- to two-term limits made it appear as though the impact of term limits was declining to zero.

Robustness Tests

As noted above, a possible source of heterogeneity in our empirical specification is that the pool of candidates may change as a result of the relaxation of term limit laws. In particular, one-term limits may do more to dissuade competent candidates from seeking office than two-term limits. If this is the case, we would be overestimating the size of the competence effect, since we would also be picking up an improvement in the quality of the candidates. We do control for the incumbent's prior political experience to attempt to address this possibility. It is worth noting that while the average amount of prior experience does increase in the pool of candidates following the switch from one- to two-term limits, the estimated effect is small—an increase of slightly less than two years—and does not rise to conventional levels of statistical significance.

A second potential endogeneity problem arises from the timing of the transition from one- to two-term limits, which may reflect voters' feelings about the incumbent who was in office at that time. The direction of the potential bias is not clear, *a priori*, as a state might abandon one-term limits either in order to allow a highly competent politician to run for reelection or as a result of heightened concern about incumbent shirking. Many states that switched from one- to two-term limits designed their laws to allow the current incumbent to run for reelection, and in most of those states the incumbent went on to win reelection, suggesting that the bias might be working in the same direction as our results. That is, a highly competent governor who, in the absence of the term limit reform, would have been coded as a first-term lame duck is instead coded as a first-term eligible governor, making the average performance of first-term lame ducks appear relatively worse than if the reform had been exogenous. However, in other states, the current incumbent was either explicitly prohibited from running for

16

reelection or was permitted to run but did not win, suggesting that endogeneity bias might be working in the opposite direction.

To address this potential bias, we conducted a robustness test in which we omitted from the sample the incumbent who was in office at the time of the policy change. The results of this robustness test are shown in the even-numbered columns of Table 4. Our results hold up well: in one case (taxes) the estimates are more or less unchanged, while in the other three cases the estimates become larger in magnitude, more precise, or both. The competence effect in the growth regression is now statistically significant at conventional levels.

Another potential threat to our empirical strategy arises from the fact that a disproportionate number of the states that switched from one- to two-term limits are located in the South. Thus, we may be accidentally picking up a "South" effect rather than (or in addition to) the term limit effect. Our control variables do capture some of the most notable Southern trends over the past fifty years—namely, increased party competitiveness, a declining share of Democratic governors and legislators, and rising incomes, and our state fixed effects, year effects, and state-specific time trends help to absorb some of the residual variation. These controls are not a panacea, of course, so as a robustness test we investigate whether our results differ for the Southern and non-Southern states in the sample. When we interact our term-limit dummy variables with a dummy for Southern states, we find that none of the results are statistically significantly different for the two sets of states with the exception of the coefficient on "second-term lame duck" in the borrowing-cost regression.[12]

Comparing Two-Term to No Term Limits

Our second empirical specification compares states with two-term limits to those with no term limits. According to Proposition 5, second-term eligible governors should perform better than both first-term eligible governors (due to a combination of competence and accountability) and second-term lame ducks (due to accountability). We include in our sample only those states that had either no term

limits or two-term limits during the entire sample period. This allows us to avoid problems with the specification used by Besley and Case (1995a; 2003), in which the estimate of an accountability effect is biased by potentially attributing to it not only the competence effects resulting from the switch from one- to two-term limits in many states but also changes in the sample resulting from the adoption of two-term limits by states that previously did not have any term limits.[13]

We run the following regression:

$$performance_{it} = \beta_1\,first\text{-}term\text{-}eligible_{it} + \beta_2\,second\text{-}term\text{-}lame\text{-}duck_{it} + \beta_3\,X_{it} + \theta_i + \delta_t + \upsilon_{it}$$

The omitted category is second-term governors who are eligible for reelection.[14] Because the omitted category is different than in the previous specification, the interpretation of coefficients is also different. The coefficient on *first-term-eligible* (β_1) now provides an estimate of a combination of competence and some accountability, since it compares first- to second-term eligible incumbents. The coefficient on *second-term-lame-duck* (β_2) provides an estimate of the accountability effect, since it compares second-term lame ducks to second-term eligible incumbents.[15]   Since performance is expected to be worse under both of these types of governors than under second-term eligibles, we expect the coefficients to be of the opposite signs as in the previous regression: positive for spending, taxes, and borrowing costs and negative for economic growth. The dependent variables and control variables are the same as in the previous specification and, as before, we control for state-specific time trends, state fixed effects, and year fixed effects.

[Table 5 about here]

Column 1 reveals that per capita spending is nearly three percent higher under first-term eligible governors than under second-term eligible governors, who have survived reelection and have more experience (competence) as well as stronger incentives to exert effort (accountability), all of which pull in the same direction. The coefficient on second-term lame duck (accountability) has the expected sign but falls short of statistical significance. Column 3 shows that per capita taxes are three

percent higher under second-term lame ducks than under second-term eligible governors (accountability), and more than four percent higher under first-term eligible governors than under second-term eligible governors (competence and accountability).

Column 5 shows that borrowing costs are about 12.2 basis points higher under second-term lame ducks than under second-term eligible governors (accountability) and about 10.9 basis points higher under first-term eligible governors than under second-term eligible governors (competence and accountability). In the economic growth regression (column 7), the coefficients have the expected signs but fall short of statistical significance at conventional levels.

In the even-numbered columns of Table 5 we separately estimate the performance of first-term governors who won reelection and those who lost or withdrew. As a basic test of the plausibility of our account, we expect first-term governors who won reelection to have performed better than those who were not reelected. The coefficient on *first-term eligible* captures the performance of first-term eligible winners relative to second-term eligible governors. The coefficient on *first-term eligible non-winner* captures the performance of the non-winners *relative* to the winners (thus, the total effect for non-winners is the sum of these two coefficients).

We find, as expected, that governors who won reelection performed better than governors who did not win reelection. Columns 4 and 8 reveal that per capita taxes are two and a half percent higher, and economic growth is nearly 0.7 percentage points lower, under non-winners than under winners. However, we do not find statistically significant differences between winners and losers in either spending or borrowing costs.

Conclusions

This paper attempts to shed light on the role of elections in representative democracy, using an agency model of elections that encompasses the effects of both moral hazard or "accountability" (voters' ability to induce politicians to take costly actions) and adverse selection or "competence"

19

(voters' ability to select "good types"). We incorporate term limits of varying lengths into the model and use it to map out an empirical strategy that treats variation in the length of term limits as a natural experiment. This allows us to extract the information necessary to disentangle accountability from competence in a common set of elections. While both of these effects have long existed in theory, this paper is among the first to empirically estimate the size of each of these effects separately.

Our empirical findings suggest that both accountability and competence play important roles in elections. First, we have found evidence of an accountability effect: per capita taxes, per capita spending, and borrowing costs are higher and economic growth is lower under term-limited incumbents than under those who are eligible for reelection, controlling for tenure in office. Second, we have found that performance improves with the incumbent's tenure in office. Per capita taxes, per capita spending, and borrowing costs are lower and economic growth is higher under second-term incumbents than under first-term incumbents, controlling for term-limit status. This is consistent with differences in competence that result from both electoral selection and experience in office. These effects were estimated most consistently when comparing one-term and two-term limit regimes. Comparisons of states with and without two-term limits further confirm the existence and scale of the two effects. We found the magnitudes of the accountability and competence effects to be similar to each other, resolving a puzzle about the changing effect of gubernatorial term limits over time.

Our finding of a sizeable competence effect suggests an important direction for future research: empirically disentangling the components of competence into electoral selection, the process by which bad incumbents are weeded out over time, and experience, the increase in competence that comes with serving in office. The key constraint in our approach that prevented us from doing so is that one cannot become an experienced governor without having survived reelection, so selection and experience are confounded. One way to drive a wedge between inherent ability and experience is to compare winners to losers within a given term (as we did in Table 5) since within a term there is no

20

experience effect. However, since competent types exert higher effort in the theoretical model, this comparison conflates inherent ability (i.e., selection) and accountability. If one could find a set of elections in which incumbents were surprised by the opportunity to run again (e.g., due to court decisions striking down term limit laws), then it might be possible to empirically separate selection from both experience and accountability.

Our findings also have implications for the normative debate on term limits. Opponents often argue that term limits not only cause incumbents to shirk but also interfere with voters' ability to choose the best-qualified representatives. Our results suggest that concerns about competence are relevant, and that the length, not just the existence, of term limits matters. Ultimately, the debate over term limits entails weighing these disadvantages against the potential advantages, such as limiting opportunism (e.g. Smart and Sturm 2006) and reducing the incumbency advantage. We do not claim to settle this debate, but rather to clarify why and how some of these effects arise and to estimate their magnitudes. Our results may also help inform analyses (e.g. Snyder and Ting 2008) of why voters and interest groups take the positions they do on term limits.

Appendix

The probability of H is given by $g(\bar{a},\theta_C)=1$, $g(\underline{a},\theta_C)=\gamma$, and $g(\bar{a},\theta_I)=g(\underline{a},\theta_I)=0$. Recall the following from the main text. First, $\mu_0<\gamma<1$. Second, the cost is drawn from an absolutely continuous, increasing cumulative distribution function, F, with associated density function f whose support is $(0,\bar{c})$, with $\bar{c}=\dfrac{(1-\gamma)B\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^{i}(1-\varepsilon^j)}{1+(1-\gamma)\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^{i}(1-\varepsilon^j)}$. We normalize the cost of low effort to zero.

Let $\mu(O,a,\mu',t)$ be the voter's posterior belief that the incumbent will be competent in her t^{th} term, given the outcome O in her t-1th term, the voter belief that the politician's action in her t-1th term was a, and the voter belief that the incumbent was competent in her t-1th term with probability μ':

$$\mu(H,\bar{a},\mu',t)=\mu(H,\underline{a},\mu',t)=1-\varepsilon^t$$
$$\mu(L,\bar{a},\mu',t)=0$$
$$\mu(L,\underline{a},\mu',t)=\frac{(1-\varepsilon^t)\mu'(1-\gamma)}{(1-\gamma)\mu'+1-\mu'}.$$

We write *equilibrium* for pure strategy PBE in Markov Strategies (Maskin and Tirole 2001).

Proposition 1. *In the unique equilibrium of the game with a one term limit, any incumbent chooses $\underline{a}$.*

Proof. For all types, the payoff from high effort is B – c and the payoff from low effort is B.∎

Proposition 2. Let $\hat{c}\equiv(1-\gamma)(1-\varepsilon)\delta B$. *In any equilibrium of the game with two-term limits:*

- *If $c<\hat{c}$, first-term incumbents choose $\bar{a}$ if competent and $\underline{a}$ if incompetent; all second-term incumbents choose $\underline{a}$. A second-term incumbent is competent with probability $1-\varepsilon^2$.*

- *If $c>\hat{c}$, both first- and second-term incumbents choose $\underline{a}$, regardless of type. The probability that a second-term incumbent is the competent type is $1-\varepsilon^2$.*

- *The voter reelects first-term incumbents if and only if the outcome is H.*

Proof. In any equilibrium incompetent types choose low effort in every period, since outcomes are unresponsive to action. In the second term, the incumbent will choose low effort even if competent, since there is no reelection. Now, consider voter behavior when faced with a first-term incumbent.

Lemma 1. *In any equilibrium of the game with two-term limits, the voter reelects a first-term incumbent following an outcome of H and does not reelect following an outcome of L.*

Proof. If the voter expects high effort from first-term competent types, then the voter reelects if $\mu(O, \bar{a}, \mu_0, 1)\gamma \geq \mu_0(1-\varepsilon)$. Following outcome H, $\mu(H, \bar{a}, \mu_0, 2) = 1 - \varepsilon^2$, so the voter reelects since $\gamma > \mu_0$. Following the outcome L the voter is certain the incumbent is incompetent, so does not reelect.

Now suppose the voter expects low effort from competent types in the first term. Then the voter reelect the incumbent when his posterior beliefs are μ if $\mu\gamma \geq \mu_0(1-\varepsilon)\gamma$. The voter will reelect following H, since his beliefs are $1 - \varepsilon^2 > \mu_0(1-\varepsilon)$. Following L, the voter's beliefs are

$$\mu(L, \underline{a}, \mu_0(1-\varepsilon), 2) = \frac{(1-\varepsilon^2)\mu_0(1-\varepsilon)(1-\gamma)}{(1-\gamma)\mu_0(1-\varepsilon) + 1 - \mu_0(1-\varepsilon)} < \mu_0(1-\varepsilon),$$ so the voter will not reelect. ∎

Now consider first-term effort by competent types.

Case 1: $c < (1-\gamma)(1-\varepsilon)\delta B$

From Lemma 1 the expected payoff to a competent type from $\bar{a}$ in the first term is $B - c + (1-\varepsilon)\delta B$, while the payoff from $\underline{a}$ is $B + (1-\varepsilon)\gamma\delta B$. Competent types choose $\bar{a}$ in the first term if $c \leq (1-\gamma)(1-\varepsilon)\delta B$, which is confirmed by the hypothesis of this case. Since incompetent types are never reelected, the probability that second-term incumbents are competent is $1 - \varepsilon^2$.

Case 2: $c > (1-\gamma)(1-\varepsilon)\delta B$

An argument identical to the one above shows that there is an equilibrium where first-period effort is $\underline{a}$ and that there is no equilibrium where first-period effort is $\bar{a}$. Again, by Lemma 1, incompetent

types are never reelected, so the probability that second-term incumbents are competent is $1 - \varepsilon^2$. The equilibrium is not unique because of the measure zero case where $c = (1 - \gamma)(1 - \varepsilon)\delta B$, making the first-term incumbent is indifferent. Whatever the incumbent does in this case, in any equilibrium play is as described in the proposition whenever $c \neq (1 - \gamma)(1 - \varepsilon)\delta B$. ∎

Proposition 3 *There is an infinite, increasing sequence* $\{\tilde{c}_1, \tilde{c}_2, ...\}$ *satisfying* $F(\tilde{c}_t) = 1$ *for all t, such that in any equilibrium of the game with no term limits* t^{th} *term competent incumbents choose* $\bar{a}$ *if* $c \leq \tilde{c}_t$, *incompetent incumbents choose* $\underline{a}$ *in any term, the voter reelects if and only if the outcome is H.*

In such an equilibrium, the probability that a tth*-term incumbent is competent is* $1 - \varepsilon^t$.

Proof. Consider the following strategy profile and belief system: Competent incumbents choose $\bar{a}$. Incompetent incumbents choose $\underline{a}$. The voter reelects if and only if his posterior beliefs are greater than $\mu_0(1 - \varepsilon)$. After outcome H in an incumbent's t^{th} term, the voter believes the incumbent is competent with probability $1 - \varepsilon^t$. After outcome L the voter is certain the incumbent is incompetent.

Given politician strategies, competent types succeed and incompetent types fail. Competent types transition with probability $1 - \varepsilon^t$, so voter beliefs are consistent. Since outcomes (and thus voter beliefs and reelection decisions) are not sensitive to effort, incompetent types optimally choose $\underline{a}$.

A competent type gets payoff $B - c + (1 - \varepsilon^t)\delta\left(B - c + (1 - \varepsilon^{t+1})\delta\left(B - c + (1 - \varepsilon^{t+2})\delta\right)..\right)$ in her t^{th} term if she follows the strategy. A one shot deviation to $\underline{a}$ yields expected payoff $B + (1 - \varepsilon^t)\gamma\delta\left(B - c + (1 - \varepsilon^{t+1})\delta\left(B - c + (1 - \varepsilon^{t+2})\delta\right)..\right)$. There is no profitable if

$$c \leq \frac{B(1 - \gamma)\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^{i}(1 - \varepsilon^{t+j})}{1 + (1 - \gamma)\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^{i}(1 - \varepsilon^{t+j})} \equiv \tilde{c}_t.$$ $\tilde{c}_t$ is minimized at $t = 1$ and $\tilde{c}_1 = \bar{c}$, so $F(\tilde{c}_t) = 1$ for all t.

24

Given politician strategies, the voter's payoff from reelection is the probability the incumbent is competent. His payoff from not reelecting is $\mu_0(1-\varepsilon),$ so his strategy is sequentially rational.

For uniqueness. the argument above shows that incompetent types cannot exert effort in equilibrium. The only possibility is for competent types not to exert effort. Notice, however, if all types choose low effort, then the voter's beliefs and strategy must be exactly as in the proposition. But, then, the argument above establishes that sequential rationality requires competent types to exert effort. ∎

Proposition 4.

1. *The expected performance of first-term incumbents in a one-term-limit system is lower than in a two-term-limit system, due entirely to differences in effort.*

2. *The expected performance of first-term incumbents in a one-term-limit system is lower than the expected performance of second-term incumbents in a two-term-limit system, due entirely to differences in expected competence.*

Proof.

1. From Proposition 1 in a one-term-limit system effort is $\underline{a}$, expected ability is $\mu_0(1-\varepsilon)$, and so expected performance is $\mu_0(1-\varepsilon)\gamma.$ From Proposition 2, effort by a first-term incumbent in a two-term limit system is $\bar{a}$ if $c \le \hat{c},$ expected ability is $\mu_0(1-\varepsilon),$ and so expected performance is $\mu_0(1-\varepsilon)\big(F(\hat{c})+(1-F(\hat{c}))\gamma\big)> \mu_0(1-\varepsilon)\gamma.$

2. From Proposition 2, effort by a second-term lame duck is $\underline{a}$, expected ability is $1-\varepsilon^2$, and so expected performance is $(1-\varepsilon^2)\gamma > \mu_0(1-\varepsilon)\gamma.$ ∎

Proposition 5.

1. *The expected performance of second-term incumbents in a two-term-limit system is lower than in a no-term-limit system, due entirely to differences in effort.*

25

2. *The expected performance of second-term incumbents in the no-term-limit system is higher than the expected performance of first-term incumbents in either term limit system. Comparing to no term limits, the difference is due entirely to differences in effort. Comparing to two-term limits, the difference is due to difference in effort and expected ability.*

Proof.

From Proposition 3, effort by a second-term incumbents in a no-term-limit system is $\bar{a}$, expected ability is $1 - \varepsilon^2$, so expected performance is $1 - \varepsilon^2$.

1. From Proposition 2, effort by a second-term lame duck is $\underline{a}$, expected ability is $1 - \varepsilon^2$, and so expected performance $\left(1 - \varepsilon^2\right)\gamma < 1 - \varepsilon^2$.

2. From Proposition 3, effort by a first-term incumbent in a no-term-limit system is $\bar{a}$, expected ability is $\left(1 - \varepsilon\right)\mu_0$, so expected performance is $\left(1 - \varepsilon\right)\mu_0 < 1 - \varepsilon^2$. From Proposition 2, effort by a first-term incumbent in a two-term limit system is $\bar{a}$ if $c \le \hat{c}$, expected ability is $\mu_0\left(1 - \varepsilon\right)$, and so expected performance is $\mu_0\left(1 - \varepsilon\right)\left(F(\hat{c}) + \left(1 - F(\hat{c})\right)\gamma\right) < 1 - \varepsilon^2$. To show an effort difference in addition to an expected competence difference, it suffices to show $\hat{c} < \bar{c}$. This follows from:

$$(1-\gamma)(1-\varepsilon)\delta B < \frac{B(1-\gamma)\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^{i}\left(1-\varepsilon^j\right)}{1+(1-\gamma)\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^{i}\left(1-\varepsilon^j\right)} \Leftrightarrow (1-\varepsilon)\delta + (1-\varepsilon)\delta(1-\gamma)\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^{i}\left(1-\varepsilon^j\right) < \sum_{i=1}^{\infty}\delta^i\prod_{j=1}^{i}\left(1-\varepsilon^j\right) \Leftrightarrow$$

$$(1-\varepsilon)\delta(1-\gamma)\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^{i}\left(1-\varepsilon^j\right) < \sum_{i=2}^{\infty}\delta^i\prod_{j=2}^{i}\left(1-\varepsilon^j\right) \Leftrightarrow -(1-\varepsilon)\delta\gamma\sum_{i=1}^{\infty}\delta^i\prod_{j=1}^{i}\left(1-\varepsilon^j\right) < \sum_{i=1}^{\infty}\delta^{i+1}\left(\varepsilon - \varepsilon^{i+1}\right)\prod_{j=1}^{i}\left(1-\varepsilon^j\right)$$

The left-hand side is clearly negative while the right-hand side is clearly positive. ∎

References

Alt, James E., and Robert C. Lowry. 2000. "A Dynamic Model of State Budget Outcomes Under
Divided Partisan Government." *Journal of Politics* 62 (4): 1035-69.

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

Ashworth, Scott. 2005. "Reputational Dynamics and Political Careers." *Journal of Law, Economics
and Organization* 21 (2): 441-46.

Ashworth, Scott, and Ethan Bueno de Mesquita. 2006. "Delivering the Goods: Legislative
Particularism in Different Electoral and Institutional Settings." *Journal of Politics* 68 (1):168-
79.

Ashworth, Scott, and Ethan Bueno de Mesquita. 2008. "Electoral Selection, Strategic Challenger
Entry, and the Incumbency Advantage." *Journal of Politics* 70 (4): 1006-25.

Banks, Jeffrey S., and John Duggan. 2008. "A Dynamic Model of Democratic Elections in
Multidimensional Policy Spaces." *Quarterly Journal of Political Science* 3 (3): 269-99.

Banks, Jeffrey S., and Rangarajan K. Sundaram. 1998. "Optimal Retention in Agency Problems."
Journal of Economic Theory 82 (2): 293-323.

Barro, Robert. 1973. "The Control of Politicians: An Economic Model." *Public Choice* 14 (1): 19-42.

Besley, Timothy. 2006. *Principled Agents? The Political Economy of Good Government*. Oxford:
Oxford University Press.

Besley, Timothy, and Anne Case. 1995a. "Does Electoral Accountability Affect Economic Policy
Choices? Evidence from Gubernatorial Term Limits." *Quarterly Journal of Economics* 110
(3): 769-98.

Besley, Timothy, and Anne Case. 1995b. "Incumbent Behavior: Vote-Seeking, Tax-Setting, and
Yardstick Competition." *American Economic Review* 85 (1): 25-45.

Besley, Timothy, and Anne Case. 2003. "Political Institutions and Policy Choices: Evidence from the
United States." *Journal of Economic Literature* 41 (1): 7-73.

Besley, Timothy, and Michael Smart. 2007. "Fiscal Restraints and Voter Welfare." *Journal of Public Economics* 91 (3-4): 755-73.

Beyle, Thad, Richard G. Niemi, and Lee Sigelman. 2002. "Gubernatorial, Senatorial, and State-level Presidential Job Approval: The U.S. Officials Job Approval Ratings (JAR) Collection." *State Politics & Policy Quarterly* 2 (3): 215-29.

Bond, Jon R., Cary Covington, and Richard Fleisher. 1985. "Explaining Challenger Quality in Congressional Elections." *Journal of Politics* 47 (2): 510–29.

Canes-Wrone, Brandice, and Kenneth W. Shotts. 2007. "When Do Elections Encourage Ideological Rigidity?" *American Political Science Review* 101 (2): 273-288.

Canes-Wrone, Brandice, Michael C. Herron, and Kenneth W. Shotts. 2001. "Leadership and Pandering: A Theory of Executive Policymaking." *American Journal of Political Science* 45 (3): 532-50.

Carey, John, Gary F., Moncrief, Richard G. Niemi, Lynda W. Powell. 2006. "Term Limits in the State Legislatures: Results from a New Survey of the 50 States." *Legislative Studies Quarterly*, 31 (1): 105-36.

Duggan, John. 2000. "Repeated Elections with Asymmetric Information." *Economics and Politics* 12 (2): 109-135.

Fearon, James. 1999. "Electoral Accountability and the Control of Politicians." In *Democracy, Accountability and Representation*, eds. Adam Przeworski, Susan C. Stokes, and Bernard Manin. New York: Cambridge University Press, 55-97.

Ferejohn, John. 1986. "Incumbent Performance and Electoral Control." *Public Choice* 50 (1-3): 5-26.

Ferraz, Claudio, and Frederico Finan. 2009. "Electoral Accountability and Corruption: Evidence from the Audits of Local Governments." *American Economic Review*, forthcoming.

Figlio, David N. 1995. "The Effect of Retirement on Political Shirking: Evidence from Congressional Voting." *Public Finance Review* 23 (2): 226-41.

Fox, Justin. 2007. "Government Transparency and Policymaking." *Public Choice* 131 (1/2): 23-44.

Gagliarducci, Stefano, and Tommaso Nannicini. 2009. "Do Better Paid Politicians Perform Better? Disentangling Incentives from Selection." IZA-Bonn Discussion paper 4400.

Gordon, Sanford C., and Gregory A. Huber. 2007. "The Effect of Electoral Competitiveness on Incumbent Behavior." *Quarterly Journal of Political Science* 2 (2): 107-38.

Gordon, Sanford C., Gregory A. Huber, and Dimitri Landa. 2007. "Challenger Entry and Voter Learning." *American Political Science Review* 101 (2): 303-20.

Gowrisankaran, Gautam, Matthew F. Mitchell, and Andrea Moro. 2008. "Electoral Design and Voter Welfare from the U.S. Senate: Evidence from a Dynamic Selection Model." *Review of Economic Dynamics* 11 (1): 1-17.

Hall, Melinda Gann, and Chris Bonneau. 2006. "Does Quality Matter? Challengers in State Supreme Court Elections." *American Journal of Political Science* 50 (1): 20-33.

Hou, Yilin, and Daniel L. Smith. 2006. "A Framework for Understanding State Balanced Budget Requirement Systems: Reexamining Distinctive Features and an Operational Definition." *Public Budgeting and Finance* 26 (3): 22-45.

Jacobson, Gary C. 1989. "Strategic Politicians and the Dynamics of U.S. House Elections, 1946–86." *American Political Science Review* 83 (3): 773–93.

Johnson, Joseph M., and W. Mark Crain. 2004. "Effects of Term Limits on Fiscal Performance: Evidence from Democratic Nations." *Public Choice* 119 (1/2): 73-90.

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

Kurtz, Karl T., Bruce Cain, and Richard G. Niemi, eds. 2007. *Institutional Change in American Politics: The Case of Term Limits*. Ann Arbor: University of Michigan Press.

Lott, John R., and Stephen G. Bronars. 1993. "Time Series Evidence on Shirking in the U.S. House of Representatives." *Public Choice* 76 (1/2): 125-49.

Lowry, Robert C., and James Alt. 2001. "A Visible Hand? Bond Markets, Political Parties, Balanced Budget Laws, and State Government Debt." *Economics and Politics* 13 (1): 49-72.

Lowry, Robert C., James E. Alt, and Karen Ferree. 1998. "Fiscal Policy and Elections in American States." *American Political Science Review* 92 (4): 759-74.

Maskin, Eric, and Jean Tirole. 2001. "Markov Perfect Equilibrium I: Observable Actions." *Journal of Economic Theory* 100 (2): 191-219.

Maskin, Eric, and Jean Tirole. 2004. "The Politician and the Judge: Accountability in Government." *American Economic Review* 94 (4): 1034-54.

Matthews, Donald R., and James W. Prothro. 1962. "Southern Racial Attitudes: Conflict, Awareness, and Political Change." *The ANNALS of the American Academy of Political and Social Science* 344 (1): 108-21.

McArthur, John, and Steven V. Marks. 1988. "Constituent Interest vs. Legislator Ideology: The Role of Political Opportunity Cost." *Economic Inquiry* 26 (3): 461-70.

Meirowitz, Adam. 2007. "Probabilistic Voting and Accountability in Elections with Uncertain Policy Constraints." *Journal of Public Economic Theory* 9 (1): 41-68.

Padro i Miquel, Gerard, and James M. Snyder. 2006. "Legislative Effectiveness and Legislative Careers." *Legislative Studies Quarterly* 31 (3): 347-81.

Peltzman, Sam. 1992. "Voters as Fiscal Conservatives." *The Quarterly Journal of Economics* 107 (2): 327-61.

Persson, Torsten, and Guido Tabellini. 2000. *Political Economics: Explaining Economic Policy*. Cambridge: MIT Press.

Poterba, James M. 1995. "Balanced Budget Rules and Fiscal Policy: Evidence from the States." *National Tax Journal* 48(3): 329–36.

Schwabe, Rainer. 2009. "Reputation and Accountability in Repeated Elections." Typescript: Princeton University.

Smart, Michael, and Daniel M. Sturm. 2006. "Term Limits and Electoral Accountability." *LSE PSPE Working Paper No. 3*.

Snyder, James M., and Michael M. Ting. 2003. "Roll Calls, Party Labels and Elections." *Political Analysis* 11 (4): 419-44.

Snyder, James M., and Michael M. Ting. 2008. "Interest Groups and the Electoral Control of Politicians." *Journal of Public Economics* 92 (3-4): 482-500.

Tien, Charles. 2001. "Representatives, Voluntary Retirement, and Shirking in the Last Term." *Public Choice* 106 (1/2): 117-30.

Van Dunk, Emily. 1997. "Challenger Quality in State Legislative Elections." *Political Research Quarterly* 50 (4): 793–807.

Vanbeek, James R. 1991. "Does the Decision to Retire Increase the Amount of Political Shirking?" *Public Finance Review* 19 (4): 444-56.

Wolfers, Justin. 2002. "Are Voters Rational? Evidence from Gubernatorial Elections." Stanford GSB Working Paper #1730.

Zaller, John. 1998. "Politicians as Prize Fighters: Electoral Selection and the Incumbency Advantage." In *Politicians and Party Politics,* ed. John G. Geer. Baltimore, MD: Johns Hopkins University Press, 128-85.

Table 1. Comparing Systems with One- and Two-Term Limits

	Expected Effort by Competent Types	$\Pr(\theta_C)$	Comparison Identifies
First Term Eligible	$\bar{a}$ only if $c \le \hat{c}$	$\mu_0(1-\varepsilon)$	Accountability Effect: $\mu_0(1-\varepsilon)F(\hat{c})(1-\gamma)$
First Term Lame Duck	$\underline{a}$	$\mu_0(1-\varepsilon)$	
Second Term Lame Duck	$\underline{a}$	$1-\varepsilon^2$	Competence Effect: $\gamma\left(\left(1-\varepsilon^2\right)-(1-\varepsilon)\mu_0\right)$

Table 2. Comparing Systems with Two- and No Term Limits.

		Expected Effort by Competent Types	$\Pr(\theta_C)$	Comparison Identifies
Second Term Lame Duck		$\underline{a}$	$1-\varepsilon^2$	Accountability Effect: $\left(1-\varepsilon^2\right)(1-\gamma)$
Second Term Eligible		$\bar{a}$	$1-\varepsilon^2$	
First Term Eligible	No Term Limits	$\bar{a}$	$\mu_0(1-\varepsilon)$	Competence Effect $\left(1-\varepsilon^2\right)-\mu_0(1-\varepsilon)$
	2 Term Limit	$\bar{a}$ only if $c \le \hat{c}$	$\mu_0(1-\varepsilon)$	Competence and Accountability $\left(1-\varepsilon^2\right)-\mu_0(1-\varepsilon)\left(F(\hat{c})+(1-F(\hat{c}))\gamma\right)$

Table 3. Summary Statistics

	States that switched from 1- to 2-term limits (1)		States with no term limits (2)		States with 2-term limits (3)	
	Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
Per capita spending	966	453	1,318	481	1,305	522
Log of per capita spending	6.74	0.55	7.11	0.40	7.08	0.47
Per capita taxes	477	215	684	260	614	237
Log of per capita taxes	6.05	0.53	6.45	0.41	6.33	0.47
Borrowing cost	4.95	22.72	15.46	26.91	12.51	22.21
Economic growth (%)	2.44	3.13	2.04	2.86	1.95	2.81
First-term lame duck	0.46	0.50	-	-	-	-
First-term eligible	0.35	0.48	0.48	0.50	0.62	0.49
Second-term lame duck	0.19	0.39	-	-	0.38	0.49
Second-term eligible	-	-	0.52	0.50	-	-
State income per capita	8,896	3,100	12,117	3,508	12,545	3,057
Population ('000)	4,735	2,796	6,475	6,196	5,099	3,619
Percent elderly	0.11	0.03	0.11	0.02	0.11	0.02
Percent school-aged	0.23	0.04	0.23	0.04	0.21	0.04
Democratic Governor	0.72	0.45	0.52	0.50	0.56	0.50
Democratic House	0.89	0.31	0.51	0.50	0.61	0.49
Democratic Senate	0.87	0.34	0.41	0.49	0.70	0.46
Political competition folded index: House	0.26	0.16	0.11	0.08	0.15	0.12
Political competition folded index: Senate	0.28	0.16	0.11	0.08	0.16	0.10
Divided government	0.26	0.44	0.51	0.50	0.38	0.49
Governor's years of prior political experience	9.16	7.53	11.37	6.99	11.43	7.56
States in sub-sample (Year of adoption)	AL (1968), FL (1968), GA (1976), IN (1972), KY (1992), LA (1966), MS (1988), MO (1965), NC (1977), OK (1966), PA (1967), SC (1980), TN (1978), WV (1970)		CT, ID, IL, NY, UT, WA		DE, MD, NJ, OH, OR	

Table 4. One-Term Limits vs. Two-Term Limits

Dependent variables: Expected signs on coefficients:	Log of per capita spending –		Log of per capita taxes –		Borrowing cost –		Economic growth +	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First-term eligible	-0.048**	-0.065**	-0.039**	-0.039**	-5.81**	-14.04**	0.66**	0.82**
(Accountability)	(0.012)	(0.015)	(0.014)	(0.018)	(2.18)	(3.45)	(0.27)	(0.33)
Second-term lame duck	-0.041**	-0.050**	-0.030**	-0.029*	-6.75**	-14.54**	0.45	0.54*
(Competence)	(0.012)	(0.015)	(0.015)	(0.018)	(2.47)	(3.44)	(0.29)	(0.32)
Sample includes governors in office at time of two-term-limit adoption?	Yes	No	Yes	No	Yes	No	Yes	No
Observations	686	622	686	622	286	261	686	622
R-squared	0.98	0.98	0.98	0.98	0.72	0.75	0.69	0.68

Note: The omitted category is first-term lame ducks. Controls: state income, population, percent elderly and school-aged, Democratic governor, Democratic house, Democratic senate, divided government, political competition in the house and senate, governor's years of prior political experience, state-specific time trends, state fixed effects, and year fixed effects.
Robust standard errors in parentheses.
* Significant at 10% level.
** Significant at 5% level.

Table 5. Two-Term Limits vs. No Term Limits

Dependent variables: Expected signs on coefficients:	Log of per capita spending +		Log of per capita taxes +		Borrowing cost +		Economic growth −	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Second-term lame duck	0.018	0.016	0.032*	0.030*	12.21**	12.76**	-0.28	-0.22
(Accountability)	(0.016)	(0.016)	(0.017)	(0.017)	(3.21)	(3.20)	(0.40)	(0.41)
First-term eligible	0.027**	0.022**	0.045**	0.038**	10.89**	11.89**	-0.21	-0.01
(Competence/accountability)	(0.011)	(0.011)	(0.011)	(0.012)	(2.62)	(2.75)	(0.31)	(0.32)
First-term eligible non-winner		0.017		0.025*		-2.98		-0.68**
		(0.012)		(0.014)		(2.94)		(0.33)
Observations	440	440	440	440	344	344	440	440
R-squared	0.97	0.97	0.97	0.97	0.50	0.50	0.58	0.58

*Note:*The omitted category is second-term reelection-eligible governors. Controls: state income, population, percent elderly and school-aged, Democratic governor, Democratic house, Democratic senate, divided government, political competition in the house and senate, governor's years of prior political experience, state-specific time trends, state fixed effects, and year fixed effects.

For the shorter period (1973-95) for which borrowing cost data are available, results in columns 5 and 6 include IA, MA, MN, ND, TX, and WI (no term limits) and KS, ME, NE, NV, and SD (two-term limits) in addition to the states listed in Table 3. TX is included from 1979 to avoid two-year terms.

Robust standard errors in parentheses.

* Significant at 10% level.

** Significant at 5% level.

Notes

[2] For theoretical models of elections with moral hazard and adverse selection, see, among others, Banks and Sundaram 1998, Persson and Tabellini 2000, Canes-Wrone, Herron, and Shotts 2001, Maskin and Tirole 2004, Ashworth 2005, Ashworth and Bueno de Mesquita 2006, Besley 2006, Besley and Smart 2007, Canes-Wrone and Shotts 2007, and Snyder and Ting 2009.

[3] A few recent studies examine the impact of the introduction of legislative term limits on the composition of state legislatures and the behavior of term-limited legislators, and find mixed results (Kousser 2005; Carey et al. 2006; Kurtz, Cain, and Niemi 2007). However, since legislative term limits took effect quite recently—between 2000 and 2002 in most states—there is only limited scope for a systematic analysis of shirking and competence effects. Gubernatorial term limits, by contrast, date from as early as 1787 (Delaware), and therefore offer greater scope for empirical analysis.

[4] This specification has generated some confusion in the literature. For example, Johnson and Crain (2004) state that "Besley and Case (1995a) find evidence of cyclical activity in policy variables for two-term limit states…. Besley and Case do not analyze the behavior of single-term limited governors, a relatively rare institution in the United States" (75). In fact, more than half of the term-limited governors in Besley and Case's sample were serving under a one-term limit.

[5] However, some agency models of elections with adverse selection and moral hazard generate somewhat different hypotheses. Meirowitz (2007) and Schwabe (2009) identify mixed strategy equilibria of a repeated model of elections with moral hazard and adverse selection in which voters do not select "good types." Several models generate behavior where high-type politicians pander to voters' interests (Canes-Wrone, Herron, and Shotts 2001, Maskin and Tirole 2004, Canes-Wrone and Shotts 2007, Fox 2007). Snyder and Ting 2009 show how the presence of lobbying can complicate the relationship between accountability and competence.

[6] We assume that the transition probability decreases over time to reflect the idea that incumbents become more effective due to experience, but all qualitative results continue to hold if the transition probability is not decreasing with terms in office.

[7] We thank Tim Besley for generously sharing the data from Besley and Case (2003). The data set has been extended and amended to include some variables from additional sources, as noted below.

[8] Another distinction between these variables arises from divergences between spending and total revenues. Although some states have strict balanced budget rules that prevent them from carrying a deficit from one year to the next, the majority do not; instead, they have weaker balanced budget rules that require the legislature to pass or the governor to propose or sign a balanced budget, but place no limits on deficits that emerge thereafter (see for example Poterba 1995 and Hou and Smith 2006).

[9] This Chubb Relative Value Survey generates biennial expert opinions of the difference in yields between each other state and, arbitrarily, New Jersey. We thank Jim Poterba and Kim Rueben for supplying the data. The data are only available for states that issue 10-year general obligation debt, and only go back to 1973, so the sample period begins in that year for the borrowing cost regressions. This enables us to include in the regression several additional states that adopted two-term limits between 1950 and 1973.

[10] In addition, we considered using several other dependent variables, including consumer confidence and voter approval ratings. However, historical state-level consumer confidence data are not available, and approval data (e.g. Beyle et al. 2002) do not go back far enough to encompass the transition from one-term to two-term limits.

[11] We return to this issue in the robustness-test section of the paper and find that, in fact, the switch from one- to two-term limits is not associated with a significant increase in the incumbent's political experience.

[12] We define the South as including Alabama, Arkansas, Georgia, Louisiana, Mississippi, North Carolina, and South Carolina. We exclude Florida and Tennessee because they are usually considered part of the "peripheral South" (Matthews and Prothro 1962).

[13] Many states that switched from no term limits to two-term limits during the sample period had two- rather than four-year gubernatorial terms prior to the adoption of term limits, potentially confounding estimates of the effects of term limits. To include Maryland and Ohio, which adopted two-term limits in the 1950s, our sample begins in 1960. We find similar results for the sample beginning in 1950 and excluding these two states.

[14] It also includes a handful of governors in their third term or higher. All results are robust to excluding these governors.

[15] Although β_1 estimates competence plus accountability and β_2 estimates pure accountability, $\beta_2 - \beta_1$ does not estimate pure competence, because the size of the accountability component in β_1 is not expected to be the same as the size of the accountability effect measured by β_2. See Table 2.